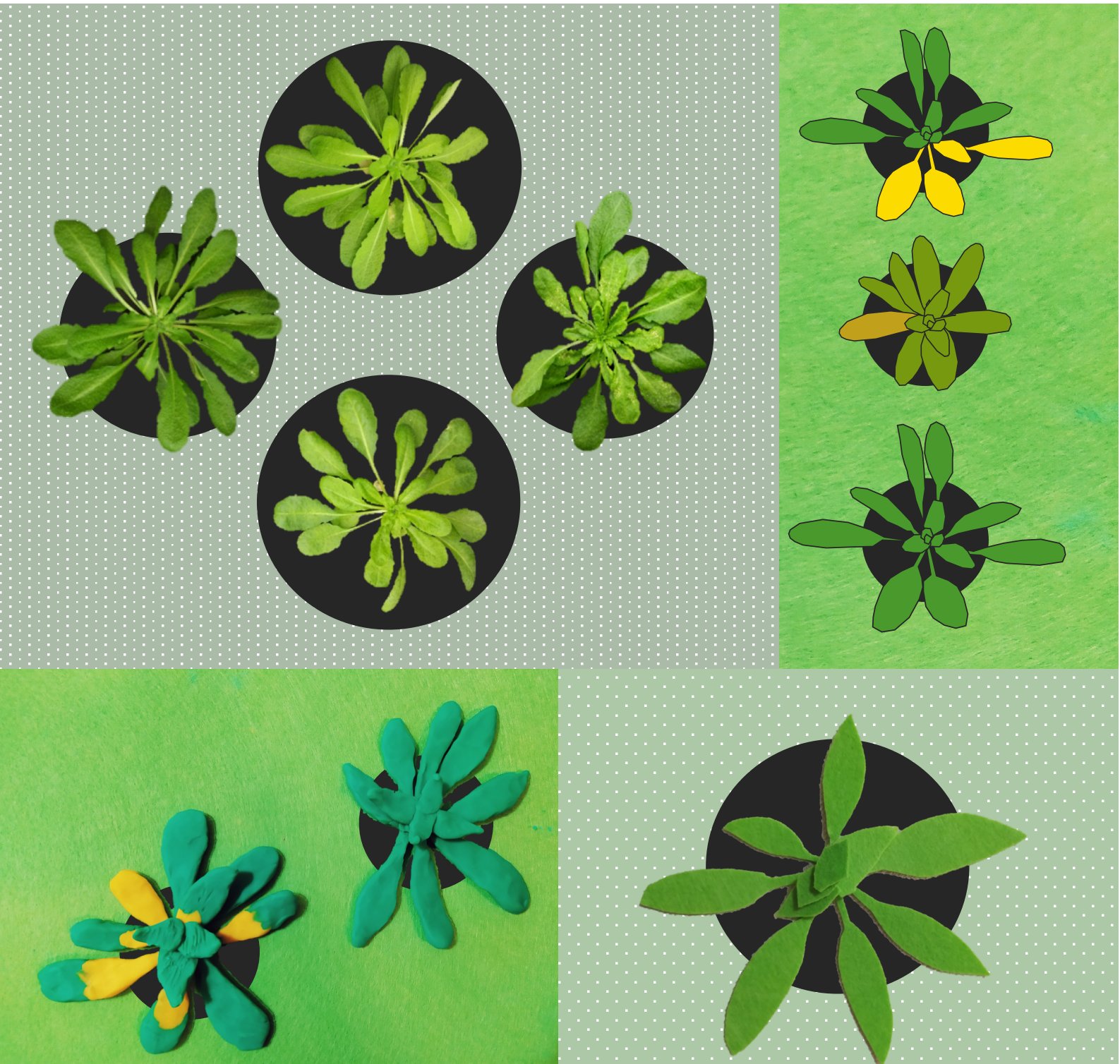




# VIRUS ADAPTATION AT DIFFERENT LEVELS

Study on the evolutionary effects of mutations,  
host population genetic structure and  
environmental factors in potyviruses



Rubén González Miguélez

Mentored by Prof. Santiago F. Elena Fito







UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

PhD Program in Biotechnology

---

## VIRUS ADAPTATION AT DIFFERENT LEVELS

Study on the evolutionary effects of mutations,  
host population genetic structure and  
environmental factors in potyviruses.

---

*October 2021*



**S.F. Elena Lab**  
Evolutionary Systems Virology

**Rubén González Miguélez**

Mentor: **Prof. Santiago F. Elena Fito**

Tutor: **Prof. Carmelo López Del Rincón**



**CSIC**

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

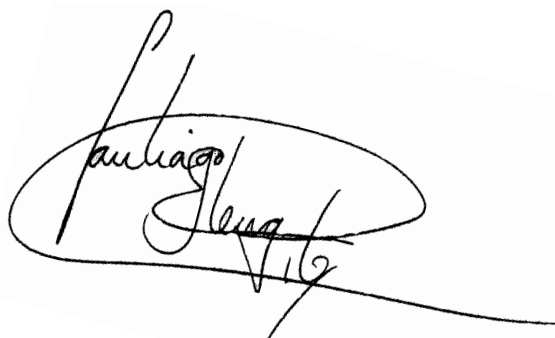


**Santiago F. Elena Fito**, doctor en Ciencias Biológicas y Profesor de Investigación del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Biología Integrativa de Sistemas (I2SysBio), centro mixto del CSIC y de la Universitat de València (UV).

CERTIFICA:

Que Don Rubén González Miguélez, master en Biología Molecular y Celular de plantas por la Universitat Politècnica de València, ha realizado bajo su supervisión la Tesis Doctoral titulada “Evolution in plant RNA viruses: study on the evolutionary effects of mutations, host population genetic structure and environmental factors in potyviruses”.

Para que así conste, en cumplimiento de la legislación vigente, firman el presente certificado en Paterna a 13 de septiembre del 2021



Doctoral thesis of



Rubén González Miguélez

Vigo, 1993

## INDEX

---

	Page
Acknowledgments	1
Abbreviations	5
Abstract / Resum / Resumen	7
General introduction	11
Objectives	25
Articles	27
Chapters	
Chapter I: Mutagenesis scanning uncovers evolutionary constraints on <i>Tobacco etch potyvirus</i> membrane-associated 6K2 protein.	29
Chapter II: Role of host genetic diversity for susceptibility-to-infection in the evolution of virulence of a plant virus.	73
Chapter III: From foes to friends: Viral infections expand the limits of host phenotypic plasticity.	103
Chapter IV: Plant virus evolution under strong drought conditions results in a transition from parasitism to mutualism.	145
General discussion	173
General conclusions	177
Supplementary material	181
References	187





## ACKNOWLEDGMENTS

---

This thesis has been completed thanks to:

- The funding provided by Spanish Government, FISABIO and EMBO.
- All people working for the correct functioning of the IBMCP, I2SysBio and IBENS.
- Project collaborators.
- Evolutionary Systems Virology group members.
- People that made it complicated, “calm seas never made a good sailor”
- Marie-Anne Félix team.
- My tutor.
- My mentor. As there are no words enough, I will acknowledge you by becoming the best scientist I can.
- All my good friends.
- My family, specially:
  - My wife. This has been fun, let's continue! *Volim te.*
  - My parents. *Gracias por todo, meus queridos pais. O voso esforzo permitiume chegar ata aquí.*

To all of you, I own you the completion of this work.

Thanks a lot.

“Do the best experiments you can, and always tell the truth. That's all.”

**Sydney Brenner**



To Anamarija.

## ABBREVIATIONS

---

<b>(-)ssRNA</b>	negative single-stranded RNA
<b>(+)ssRNA</b>	positive single-stranded RNA
<b>ABA</b>	abscisic acid
<b>AMV</b>	<i>Alfalfa mosaic alfamovirus</i>
<b>BMV</b>	<i>Brome mosaic bromovirus</i>
<b>BYDV</b>	<i>Barley yellow dwarf luteovirus</i>
<b>CaMV</b>	<i>Cauliflower mosaic caulimovirus</i>
<b>CMV</b>	<i>Cucumber mosaic cucumovirus</i>
<b>CP</b>	capsid protein
<b>CThTV</b>	<i>Curvularia thermal tolerance virus</i>
<b>CVM-1</b>	<i>Chlorella chlorovirus Marburg-1</i>
<b>CYDV</b>	<i>Cereal yellow dwarf polerovirus</i>
<b>DNA</b>	deoxyribonucleic acid
<b>dsDNA</b>	double-stranded DNA
<b>dsRNA</b>	double-stranded RNA
<b>Fig</b>	figure
<b>IAA</b>	indole-3 acetic acid
<b>ICTV</b>	International Committee on Taxonomy of Viruses
<b>JA</b>	jasmonic acid
<b>JA-Ile</b>	jasmonoyl isoleucine
<b>KYMV</b>	<i>Kennedya yellow mosaic tymovirus</i>
<b>mRNA</b>	messenger RNA
<b>OPDA</b>	oxo-phytodienoic acid

<b>PA</b>	phaseic acid
<b>PCV-1</b>	<i>Pepper cryptic deltapartitivirus 1</i>
<b>PPV</b>	<i>Plum pox potyvirus</i>
<b>PVX</b>	<i>Potato potexvirus X</i>
<b>RdRp</b>	RNA-dependent RNA polymerase
<b>RNA</b>	ribonucleic acid
<b>RT</b>	retrotranscriptase
<b>ssDNA</b>	single-stranded DNA
<b>ssRNA</b>	single-stranded RNA
<b>TEV</b>	<i>Tobacco etch potyvirus</i>
<b>TMV</b>	<i>Tobacco mosaic tobamovirus</i>
<b>TRSV</b>	<i>Tobacco ringspot nepovirus</i>
<b>TRV</b>	<i>Tobacco rattle tobravirus</i>
<b>TSWV</b>	<i>Tomato spotted wilt orthospovirus</i>
<b>TuMV</b>	<i>Turnip mosaic potyvirus</i>
<b>TYLC</b>	<i>Tomato yellow leaf curl begomovirus</i>
<b>VRC</b>	virus replication complex
<b>WT</b>	wild-type
<b>YTMMV</b>	<i>Yellow tailflower mild mottle tobamovirus</i>

## ABSTRACT / RESUM / RESUMEN

---

Experimental evolution allows us to test theoretical postulates and make observations that help increase our knowledge about Evolution. This work aims to use experimental approaches to study the evolution of viruses. Viruses have a high degree of evolvability, which makes them perfect subjects to address evolutionary questions quite rapidly. The underlying processes of pathogen evolution are governed by many factors. These factors can be affecting the virus adaptation at different levels: from the intrinsic virus nature to environmental factors affecting the host, the pathogen and the interaction between both. In this thesis, we used a pathosystem formed by a plant and a potyvirus (+ssRNA virus). Using this pathosystem we have explored how different factors modulate the virus evolution. First, we explored the biological effects of mutations in a potyvirus protein that is an essential component of the virus replication complex. We unveiled the evolutionary constraints on this viral protein, with an evolutionary tradeoff between within-host accumulation and severity of symptoms. Second, we examined the effects of the host population genetic structure on virus evolution: we evolved viruses in homogeneous populations of plants with different viral susceptibilities and in a heterogeneous population. With this work we illustrated how the genetic diversity of hosts in an ecosystem affects virus adaptation, as viruses specialized faster in homogeneous populations but were more pathogenic in heterogeneous ones. Finally, we studied the impact of the environment. For this part we first reviewed the possible beneficial effects of virus infection under certain environments. Afterwards we studied the effect of drought, an environmental condition with a predicted increased incidence and known to affect the host physiology. Therefore, we evolved a virus in host under either well-watered or drought conditions. The viruses adapted under drought conditions conferred



an increased drought tolerance to the host plant through specific alterations in host gene expression and hormonal signaling. Overall, this thesis contributed to the increase in knowledge in evolutionary biology of plant RNA viruses.

## ABSTRACT / RESUM / RESUMEN

---

L'evolució experimental ens permet comprovar postulats teòrics i realitzar observacions que ajuden a incrementar el nostre coneixement sobre l'evolució. Aquest treball té com a objectiu estudiar l'evolució dels virus utilitzant enfocaments experimentals. Els virus mostren una alta capacitat d'evolució, la qual cosa els converteix en models perfectes per a abordar qüestions evolutives amb bastant rapidesa. Els processos subjacents a l'evolució i adaptació dels patògens es regeixen per molts factors: des de la naturalesa intrínseca del virus fins a components ambientals que afecten l'hoste, al patogen i la interacció entre tots dos. En aquesta tesi utilitzem un patosistema format per una planta i un potyvirus (virus de (+)ssRNA) per a explorar com diferents factors biòtics i abiòtics modulen l'evolució del virus. Primer, explorem els efectes biològics de les mutacions en una proteïna del potyvirus, la qual és un component essencial del complex de replicació viral. Revelem les limitacions evolutives que operen sobre aquesta proteïna, i que són conseqüència d'un equilibri evolutiva entre l'acumulació dins de l'hoste i la gravetat dels símptomes. En segon lloc, examinem els efectes de l'estructura genètica de la població de l'hoste sobre l'evolució del virus: evolucionem virus en poblacions genèticament homogènies de plantes amb diferents susceptibilitats a la infecció i en una població heterogènia. Amb aquest treball il·lustrem com la diversitat genètica d'hostes en un ecosistema afecta l'adaptació del virus, ja que els virus es van especialitzar més ràpidament en poblacions homogènies però van ser més patògens en poblacions

heterogènies. Finalment, estudiem l'impacte de l'ambient sobre la interacció virus-planta. Per a aquesta part, primer revisem els possibles efectes beneficiosos de la infecció per virus en uns certs entorns hostils per a la planta. Posteriorment estudiem l'efecte de la sequera, una condició ambiental amb una incidència cada vegada major i que se sap afecta la fisiologia de l'hoste. Per tant, evolucionem un virus en hostes sota condicions de sequera o de reg abundant. Els virus adaptats en condicions de sequera conferien una major tolerància a la sequera a la planta hoste a través d'alteracions específiques en l'expressió gènica de l'hoste i la senyalització hormonal. En general, aquesta tesi contribueix a l'augment del coneixement en biologia evolutiva dels virus d'ARN de plantes.

#### **ABSTRACT / RESUM / RESUMEN**

---

La evolución experimental nos permite comprobar postulados teóricos y realizar observaciones que ayuden a incrementar nuestro conocimiento sobre la evolución. Este trabajo tiene como objetivo estudiar la evolución de los virus utilizando enfoques experimentales. Los virus muestran una alta capacidad de evolución, lo que los convierte en modelos perfectos para abordar cuestiones evolutivas con bastante rapidez. Los procesos subyacentes a la evolución y adaptación de los patógenos se rigen por muchos factores: desde la naturaleza intrínseca del virus hasta componentes ambientales que afectan al hospedador, al patógeno y la interacción entre ambos. En esta tesis utilizamos un patosistema formado por una planta y un potyvirus (virus de (+)ssRNA) para explorar cómo diferentes factores bióticos y abióticos modulan la evolución del virus. Primero, exploramos los efectos biológicos de las mutaciones en una proteína del potyvirus, la cual es un componente esencial del complejo de replicación viral. Revelamos las limitaciones

evolutivas que operan sobre esta proteína, y que son consecuencia de un equilibrio evolutiva entre la acumulación dentro del huésped y la gravedad de los síntomas. En segundo lugar, examinamos los efectos de la estructura genética de la población del huésped sobre la evolución del virus: evolucionamos virus en poblaciones genéticamente homogéneas de plantas con diferentes susceptibilidades a la infección y en una población heterogénea. Con este trabajo ilustramos cómo la diversidad genética de huéspedes en un ecosistema afecta la adaptación del virus, ya que los virus se especializaron más rápidamente en poblaciones homogéneas pero fueron más patógenos en poblaciones heterogéneas. Finalmente, estudiamos el impacto del ambiente sobre la interacción virus-planta. Para esta parte, primero revisamos los posibles efectos beneficiosos de la infección por virus en ciertos entornos hostiles para la planta. Posteriormente estudiamos el efecto de la sequía, una condición ambiental con una incidencia cada vez mayor y que se sabe afecta la fisiología del huésped. Por lo tanto, evolucionamos un virus en huéspedes bajo condiciones de sequía o de riego abundante. Los virus adaptados en condiciones de sequía conferían una mayor tolerancia a la sequía a la planta huésped a través de alteraciones específicas en la expresión génica del huésped y la señalización hormonal. En general, esta tesis contribuye al aumento del conocimiento en biología evolutiva de los virus de ARN de plantas.

## GENERAL INTRODUCTION

---

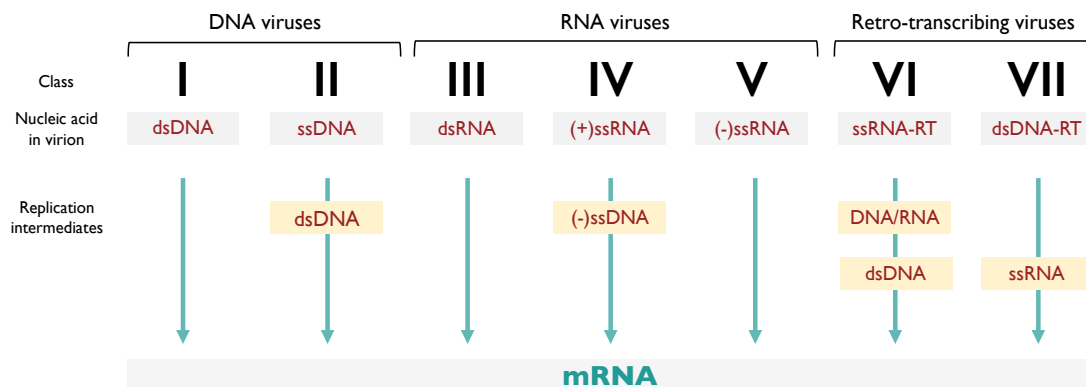
### 1. A general View Into Viruses

Dmitrii Ivanowski was studying tobacco leaves with mosaic symptoms when in 1892 he demonstrated that, even after passing through a filter that excluded bacteria, sap from these leaves remained infective. Six years later, Martinus Beijerinck demonstrated that this pathogen was able to multiply in the plant host and therefore it was not any bacterial toxin; what he was observing was a new infectious agent that he described as *contagium vivum fluidum* or *virus* (Knipe and Howley, 2013). Our knowledge about viruses have notably expanded since then and virus research has help other science fields to progress.

Nowadays we know that viruses are intracellular infectious agents that consist of a monopartite or fragmented, circular or linear, single-stranded or double-stranded nucleic acid (DNA or RNA) that is surrounded by a shell formed by one or more proteins. They are entities with autonomous replication, capable of transmitting their genome from one cell to another by using their host's enzymatic machinery for intracellular multiplication and transmission between host cells. The consequences of the infection for the host can go from dead to such a benefit that the host-virus interaction would evolve into a symbiosis (Roossinck and Bazán, 2017). With the development of sequencing techniques and a change in the perspective with which viruses are viewed, we are starting to realize that most of the organisms are infected with viruses and only few of them cause actual harm (Roossinck, 2011). We are also starting to realize about the important role that viruses have in global ecosystems. The interactions viruses establish need to be studied further, but so far it is known that viruses can alter its host behavior, its physiology and how they interact with other organisms. This may induce changes in the host niche and

therefore a modification in how components of the system interact (French and Holmes, 2019). The impact of viruses in ecosystems must be taken into consideration as viruses are present at high abundance and diversity. It has been estimated to be up to  $10^{31}$  virus particles in our planet (Mushegian, 2020), which is larger than the number of stars in the known Universe ( $10^{22}$  to  $10^{24}$ ) and the diversity within viruses is so big that it makes their classification a difficult task. There are many approaches trying to classify them but two main classifications are mostly used: Baltimore and hierarchical.

The Baltimore system (Baltimore, 1971) classifies viruses depending on the nature of the nucleic acid present in virion particles and its intermediates until being expressed as messenger RNA (mRNA). According to this system viruses can be classified in seven groups (Fig. 1):



**Fig. 1.** Schematic representation of Baltimore's classification. Adapted from the illustration of Thomas Splettstoesser ([www.scistyle.com](http://www.scistyle.com)).

Viruses having DNA genome mostly belong to class I or II depending if they have a double-stranded (ds) or single-stranded (ss) genome respectively. Their replication depends of the use of host cellular proteins and it usually occurs in the nuclei of the infected host cells. Viruses with RNA genome are also classified based on their ds-

genomes (class III, replicating in the cytoplasm and encoding their own replication enzymes) or ss-genomes. Single-stranded RNA viruses are the most abundant and they can be divided into viruses where the genome encapsidated can directly be translated by the cellular translation machinery (positive sense, group IV) or if it has to be first transcribed into its complementary strand which will later be translated by the cellular ribosomes (negative sense, group V). The other two groups are conformed by retro-transcribing viruses: group VI are positive sense ssRNA viruses with dsDNA intermediate meanwhile group VII correspond to dsDNA viruses with ssRNA intermediate.

The International Committee on Taxonomy of Viruses (ICTV) has the task of establishing a taxonomy for viruses. For this task it has adopted a hierarchical system with the taxonomic levels of order, family, subfamily, genus, and species. As the study of virus diversity progress, new viruses are discovered and/or classified. Thanks to the new techniques and tools available, in the past years an increasing number of virus species has being classified by the ICTV (Table 1).

To define a virus specie, the ICTV has adopted the concept proposed by Van Regenmorte (1990): a virus species is a “a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche”. This concept was updated in 2013 (Adams et al. 2018) to:

“A specie is the lowest taxonomic level in the hierarchy approved by the ICTV. A species is a monophyletic group of viruses whose properties can be distinguished from those of other species by multiple criteria. [...] These criteria may include, but are not limited to, natural and experimental host range,

cell and tissue tropism, pathogenicity, vector specificity, antigenicity, and the degree of relatedness of their genomes or genes”

**Table 1.** Viruses diversity classified by ICTV through the years. Adapted from Lefkowitz et al. (2017).

Year	Release Information	Number of Taxa				
		Order	Family	Subfamily	Genus	Species
2016	EC 48, Hungary	8	122	35	735	4404
2009	ICTV 9th Report	6	87	19	349	2285
2005	ICTV 8th Report	3	73	11	289	1899
1999	ICTV 7th Report	3	64	9	234	1551
1995	ICTV 6th Report	1	50	9	166	2220
1991	ICTV 5th Report	1	40	9	142	1674
1982	ICTV 4th Report	0	29	8	97	1209
1979	ICTV 3rd Report	0	24	8	84	1008
1976	ICTV 2nd Report	0	17	3	67	754
1971	ICTV 1st Report	0	2	0	43	290

Virus species can infect organisms from all the life domains [prokaryotes (Weinbauer, 2004), archaeas (Adebon and Murphy, 2013), eukaryotes] and even other viruses (Desnues et al., 2012). Within the eukaryotes viruses infects all its kingdoms: algae (Short, 2012), fungi (Xie and Jiang, 2014), animals, and plants.

## 2. Plant Viruses

Viruses infecting plant hosts have to penetrate the plant cuticle and cell wall in order to reach the cell. As far as plant viruses do not have known mechanisms to surpass these barriers they enter passively through wounds. Therefore, in order to get horizontally transmitted plant viruses, need an agent that causes a mechanical damage to the host so they can make it through the cell and try to stablish an infection.

The infection is not always successful as the virus may not be adapted to infect the host. Even if the virus would be able to infect the host they are inoculated in, plants present a complex defense system against pathogens that obstructs the success of viral infection: (1) RNA interference (RNAi) or RNA silencing, a mechanism triggered by dsRNA formed during viral replication that leads to the degradation of invading viruses (Lewsey et al., 2018). (2) Translation repression, where the plant represses their translation activity at different levels in order to impede the use of the host machinery to generate viral proteins (Machado et al., 2017). (3) Ubiquitination-mediated and autophagy-mediated protein degradation, where ubiquitins are covalently attached to viral proteins so they are degraded by the plant machinery (Verchot, 2016). (4) resistance genes, virus-specific genes that impede the infection as they recognize a virus avirulence (Avr) gene and activate a response that induce apoptosis (so-called hypersensitive response) around the cells where the virus first entered, avoiding virus spread and infection. Plants may also have genes that provide a partial resistance (tolerance genes) that allow low viral levels in the host (de Ronde, 2014). (5) DNA methylation and histone modification may play a role in host antiviral response, the magnitude of this viral defense depending on the degree of viral adaptation to the host (Corrêa et al., 2020). (6) Phytohormone regulation, as plant hormones (mainly salicylic acid (SA), jasmonic acid (JA), ethylene (Et) and abscisic acid (ABA) mediate the response to pathogens as they regulate the plant defense. Viruses also have a set of mechanisms to counter the host defense (Wu et al., 2019). For example, viral proteins can hijack the host ones to create complex centers wherein they can replicate or move more protected from host defense. Viruses can also reprogram their host to their benefit by manipulating its translational activity, hormonal levels or ubiquitination and autophagy pathways. Even the resistance



genes can be overcome thanks to a single amino acid change in the virus protein recognized by the resistance gene.

If the virus is able to infect the plant host, in its infection cycle (Bowman, 2019) the first step after entering the cell will be to get rid of the virion coat protein in the cytoplasm. After this the virus will hijack the translation machinery in order to produce new viral proteins. It is important to point that due to its nature DNA viruses will also need to hijack the transcription machinery, so these viruses must access the cell nucleus. The new viral proteins will be vital for the virus as they will (1) have counter-defense functions, (2) associate with host proteins to form replication complexes that will replicate the viral genetic material, (3) serve as structural components of new assembled virions, and (4) allow the virus to move to other cells by modifying plasmodesmata and launch the formation of movement complexes by association with other viral and host proteins. Viruses first movements are limited to neighbor cells until it reaches the phloem. This is known as local infection and the diffusion of the virus during this phase is as low as 0.01-0.03 mm<sup>2</sup>/h (Hillung et al., 2013). If the plant has a hypersensitive response to the virus, the cells surrounding the local infection will suffer necrosis and this will prevent the virus to spread. If this does not happen and the virus reaches the phloem it will spread through the plant and colonize new tissues, making the infection systemic.

The virus may reach reproductive tissues of the plant and accumulate in the seeds, this allowing vertical transmission to the host offspring and infection. This way of transmission is thought to be associated with viruses that end up having a non-pathogenic relationship with their hosts and it is a really infrequent way of transmission as meristematic tissues are often virus-free. The most common transmission for plant viruses

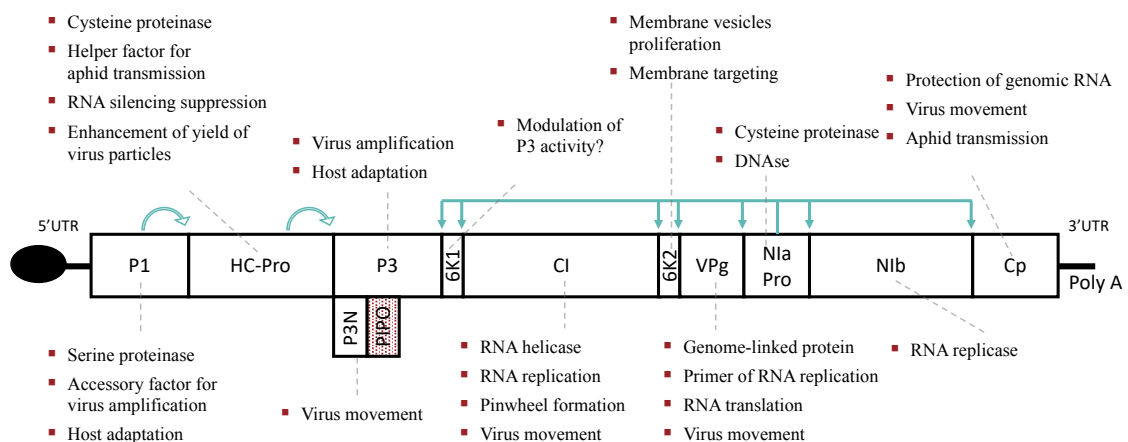
is the horizontal one, where they are transmitted from the host they are infecting to a non-infected one. As there are few viruses reported to remain long-term infective in the environment most plant viruses need to be actively transmitted from their current host to a healthy one. This is done by other organisms called vectors.

Some plant RNA viruses (reoviruses, rhabdoviruses, orthospoviruses, marafiviruses, and tenuiviruses) are even more particular, since they can infect and replicate in a large number of plant species but they also are able to generate viral inclusion and replicate in the digestive system tissues of the thrips that carry them from plant to plant (Hogenhout et al. 2008; Oliver and Whitfield, 2016). These viruses increase their genome and evolve in both plants and thrips, being even processed differentially depending which one they are infecting (Fletcher et al., 2016) so it is complicated to define what organism is the main host of those viruses. The commonly-used term plant virus may not be really accurate as most of the so-called plant viruses probably have evolved from viruses that once had non-plant organisms as hosts (Lefeuvre et al., 2019). Here we will consider whatever virus that is able to replicate in plants as a plant virus. Even though plant viruses are known for being strongly pathogenic to certain plant species of interest, causing mayor losses in agricultural activities (Hull, 2011), there is increasing evidence that most of the plant viruses are non-pathogenic to their hosts. The new sequencing technologies, the change in the perspective used to study viruses and the exploration of natural systems will shed light onto this. More than 900 plant virus species have been reported, but the number of plant viruses to be discovered is probably much higher (Roosinck, 2012). Between all these plant viruses, the most common are the RNA ones and the (+)ssRNA viruses, that alone represent the 65% of known viruses (Bowman, 2019). In this work we have studied

potyviruses, a genus of (+)ssRNA viruses responsible of the largest virus-caused losses in agriculture.

### 2.1 Potyviruses

Potyvirus is one of the eight genera within the family *Potyviridae*. They are the largest genus of plant RNA viruses, containing around the 90% of the identified species of their family (King, 2012). Viruses of this genus belong to the Blatimore's class IV, they are (+)ssRNA viruses. Their genomes are around 10 Kb long and contain one major open reading frame (ORF) that is translated to a polyprotein that is self-proteolytically processed into 10 proteins (Fig. 2). There is an additional ORF in the middle of the P3 coding region that originates by the slippage of the viral polymerase at a specific GA6 motif. This ORF generates the P3N-PIPO protein (Chung et al, 2008; White, 2015). Therefore, most potyvirus genomes encode for eleven proteins. Nevertheless, some sweet-potato-infecting potyviruses have an extra ORF originated by slippage downstream of a GA6 motif in the P1 coding region (Mingot et al., 2016). This ORF encodes for P1N-PISPO, resulting in some sweet-potato-infecting potyviruses having twelve proteins.



**Fig. 2.** Schematic representation of a potyvirus genome, adapted from Revers and García, 2015.

Each potyvirus virion contains a single copy of the genome and they have the shape of a flexuous filament with a length of between 700 and 750 nm. In nature potyviruses are propagated by vectors in a non-persistent manner (they do not replicate within the insect), being aphids their main propagators. The potyviruses interact with aphids through their CP protein but require another virus helper protein for the interaction (Dombrovsky et al., 2005). In some potyviruses seed transmission has been described (Simmons and Munkvold, 2014) but how the virus reaches the germ line is not known yet. It is important to note that viruses may reach only a portion of the host seeds and from there only a few may be transmitted to the seedlings.

Potyviruses are important pathogens causing heavy losses in a wide range of horticultural species, which makes them really well studied (Revers and García, 2015). This is due to the fact that almost all infected plants will show symptomatology that may affect the quantity and/or the quality of the crop. Infected plants will show different symptoms with diverse degrees of intensity depending on the environment, the host and which potyvirus is infecting it but general potyvirus-infection symptoms are: developmental arrest, organ deformation (specially leaves and caulinar apex), a range of vein bleaching and leaf chlorosis that can show a pattern (mosaic, etch, etc.) and partial or total necrosis. In some cases, infected plants can also show enormous yield losses as the virus can cause silique abortions in the host. The infection also causes within-cell visible symptoms as potyvirus-infected host cells typically present pinwheel inclusion bodies in their cytoplasm (Singh et al. 2014)

The research done in this thesis uses two well know potyviruses: turnip mosaic virus (TuMV; species *Turnip mosaic potyvirus*, genus *Potyvirus*, family *Potyviridae*) and the tobacco etch virus (TEV; species *Tobacco etch potyvirus*, genus *Potyvirus*, family *Potyviridae*). TuMV has a very wide host range, as it has spread all over the world as an important pathogen of *Brassicaceae* crops (broccoli, brussels sprouts, cabbage, cauliflower, mustard seed, rapeseed, etc.). TEV has a moderately wide host range, with most of its natural hosts being species of the family *Solanaceae*. Their wide host range makes it possible to study them in plants traditionally used in basic research.

## 2.2 Plant Hosts Employed in this Work

Research done in this thesis mainly uses *Arabidopsis thaliana* (L.) Heynh, a flowering plant that became of interest for molecular genetic experiments due to its small size (around 40 cm tall), relatively small genome for a eukaryote (135 Mbp) and a short generation time (four to five weeks) (Meyerowitz and Pruitt, 1985). *A. thaliana* has been broadly study by scientists in the past decades (Krämer, 2015) becoming a model organism in plant research. As a widely used and thoroughly studied organism, there is thousands of studies done in *A. thaliana* tackling its genetics, evolution, physiology, development, interaction with pathogens, etc. *A. thaliana* also shows a notable natural variation and thousands of natural accessions have been identified, allowing a deep study of the ecology and genetics of the plant (Weigel, 2012). Furthermore, more than a thousand of these natural accession genomes have been sequenced (The 1001 Genomes Consortium, 2016). For some of these natural accessions (mostly *Ler-0* and *Col-0*, the most studied ones) a large catalogue of gene knock-outs plants is available for researchers. In virology, *A. thaliana* is a suitable host for a wide range of plant viruses: so far at least 46 different species (with multiple viral strains in each specie) from 16

genera have been described to infect Col-0 (Shukla et al., 2019). Despite being able to infect *A. thaliana*, certain potyvirus species can show difficulties to infect this model organism. However, *A. thaliana* plants can end up being infected as viral adaptation to more susceptible accessions can expand the number of accessions a virus can infect (Lalić et al., 2010). Even certain potyviruses that poorly infect *A. thaliana* (as is the case for TEV) can improve their infection through adaptation to susceptible accessions, the adapted strain is still not performing as well as it would in other hosts. Therefore, our research used TEV and *Nicotiana benthamiana* Domin and *Nicotiana tabacum* L. cv Xanthi NN as hosts, as these plants are widely used in plant virology studies, showing a systemic infection for TEV and are easily agroinfiltrated. Furthermore, the big biomass of this plant compared to *Arabidopsis* makes it an ideal host to prepare stocks of virus-infected tissue. To obtain local infections *Chenopodium quinoa* Willd was the selected host. Having a host with local infection allows isolation of virus subclonal populations (Cervera and Elena, 2016) and by using serial-diluted inoculums one can calculate the infectious viral load of the original inoculum (Kleczkowski, 1950).

In summary, plants are organisms with low-cost maintenance and no bioethics concerns that can be subject to studies whose results could be extrapolated to other organisms.

### 2.3. Plants Science

Plant science is important not only for the implications in agricultural systems but it greatly contributes to other fields of science (Takanami, 2006). Next, some of the most relevant contributions of plant science will be highlighted. The term cell was coined by Robert Hooke as he observed them in 1655 when examining cork, the phellem layer of

*Quercus suber* L-, through magnifying glasses (Hooke, 2007). In 1853 Heinrich Anton de Bary proposed the role of microorganisms in disease in his book “Researches Concerning Fungal Blights”, pointing that a fungal mycelium was the agent of the wheat rust. Years later, Gregor J. Mendel performs experiments in thousands of pea plants whose results lead him to coin the concept of gene and the principles of Genetics. The demonstration that new organism variants are caused by spontaneous heritable variations of the heritable material (coined mutations) was done by Hugo de Vries (1901-03; The Mutation Theory) after studying primrose species of *Oenothera*. The maize research of Barbara McClintock led her to the discovery in the late 1940s of bits of DNA (transposons) that move and influence the expression of other genes (Ravindran, 2012). Folke K. Skoog and Carlos Miller showed in 1957 that tobacco cell cultures could differentiate to shoot or root tissues depending on the hormonal composition of the medium, proving the botanist Gottlieb Haberlandt’s theory of totipotentiality: ‘all plant cells are able to give rise to a complete plant’. The study of plant pathogens allowed not only the birth of virology but also helped us understand the chemical nature of viruses, visualize the first electron microscopy and X-ray crystallography of virus particles, obtain the first complete genome sequence of a eukaryotic pathogen and other discoveries as the existence of the viroids or the viral suppressors of RNAi-mediated response (Knipe and Howley, 2013). In summary, the study of plants can lead to breakthrough discoveries that can be applied to other fields of knowledge.

But apart from the discovery of the cell, the principles of genetics, the rise of virology, the discovery of transposons, the totipotency of cells and the post-transcriptional gene silencing, what have plant research ever done for science? Apart from being interesting

entities for studies in many fields, plants are infected by viruses that can be used as models for understanding basic mechanisms of evolution (Roossinck et al., 2003).

### **3. Virus Evolution**

Viruses are complex adaptive systems (Solé and Elena, 2019). In order to replicate in the host, viruses establish complex interaction networks with their host. With these interactions, viruses regulate the host to be able to replicate in it. If these interactions are suboptimal the virus will not be able to replicate (no infection) or it will replicate poorly (usually associated with milder infections). Therefore, viruses evolve adapting their interactions with their hosts in order to find an optimal environment for its replication.

Viruses have high mutations rates:  $10^{-8}$  to  $10^{-6}$  substitutions per nucleotide per cell infection for DNA viruses and from  $10^{-6}$  to  $10^{-4}$  for RNA viruses (Sanjuán et al., 2010). The high mutation rates limit the genome length, which results in compacted genomes (Elena and Sanjuán, 2005). Therefore, the small genomes of viruses, where proteins are multifunctional, are more prone to have mutations with a deleterious (or even lethal) effect in fitness (Sanjuán, 2010). Viruses deal with this high mutations rates owed to population-level strategies: their large populations provide robustness to harmful mutations (Elena and Sanjuán, 2005). The high mutation rate and the large population size leads to heterogenous viral populations, originating a swarm of sequences called quasispecies. Viral quasispecies refers to a population structure consisting of a swarm of variant genomes where mutants arise continually and their relative frequency changes as viral replication occurs (Domingo and Perales, 2019). In this heterogeneous population low-fitness genomes will be removed from the population due to natural selection, the genomes with the highest fitness are the ones replicating and adapting to the host.



In summary, viruses are highly evolvable: they have small compacted genomes, high mutation rates, short generation times and reach very large and heterogeneous populations. Therefore, viruses are perfect for evolutionary studies.

## OBJECTIVES

---

The main goal of this thesis is to explore, using potyviruses, the evolutive dynamics of RNA viruses at different levels. The specific objectives are to study:

1. The evolutionary constraints of a virus protein fundamental for its replication and movement.
2. The differences in virus evolutionary dynamics caused by the genetic structure of host populations.
3. The effects of environmental stress conditions in virus evolution.



## ARTICLES

---

- The following articles are part of this thesis:

### Chapter I

González, R.\* , Wu, B.\* , Li, X.\* , Martínez, F.†, Elena, S.F.† 2019. Mutagenesis scanning uncovers evolutionary constraints on tobacco etch potyvirus membrane-associated 6K2 protein. *Genome Biology and Evolution* **11**: 1207-1222.

### Chapter II

González, R., Butković, A., Elena, S.F.† 2019. Role of host genetic diversity for susceptibility-to-infection in the evolution of virulence of a plant virus. *Virus Evolution* **5**: vez024.

### Chapter III

González, R.\*†, Butković, A.\* , Elena, S.F.† 2020. From foes to friends: viral infections expand the limits of host phenotypic plasticity. *Advances in Virus Research*. **106**: 85-121.

### Chapter IV

González, R., Butković, A., Escaray F.J., Martínez-Latorre, J., Melero, I., Pèrez-Parets, E., Gómez-Cadenas, A., Carrasco, P., Elena, S.F.† 2021. Plant virus evolution under strong drought conditions results in a transition from parasitism to mutualism. *Proceedings of the National Academy of Sciences of the United States of America*. **118**: e2020990118.

- During the development of this thesis, his author also contributed to the following articles:

Butković, A., González, R., Cobo, I., Elena, S.F.† 2020. Adaptation of *Turnip mosaic potyvirus* to a specific niche reduces its genetic and environmental robustness. *Virus Evolution* **6**: veaa041.

González, R.\*†, Butković, A.\*, Rivarez, M. P. S., Elena, S.F. 2020. Natural variation in *Arabidopsis thaliana* rosette area unveils new genes involved in plant development. *Scientific Reports* **10**: 17600.

Butković, A., González, R., Rivarez, M.P.S., Elena, S.F.† 2020. *Arabidopsis thaliana* genes contributing to differences in the outcome of infection with generalist and specialist strains of turnip mosaic virus identified by genome-wide association studies. *bioRxiv* 397661 (<https://doi.org/10.1101/2020.11.25.397661>)

Butković, A., González, R., Elena, S.F.† 2021. Revisiting *Orthospovirus* phylogeny using full-genome data and testing the contribution of selection, recombination and segments reassortment in the origin of new species. *Archives of Virology* **166**: 491-499.

\* equal contribution

† corresponding author

## CHAPTER I

---

### Mutagenesis scanning uncovers evolutionary constraints on TEV membrane-associated 6K2 protein

---

#### I. ABSTRACT

RNA virus high mutation rate is a double-edged sword. At the one side, most mutations jeopardize proteins functions; at the other side, mutations are needed to fuel adaptation. The relevant question then is the ratio between beneficial and deleterious mutations. To evaluate this ratio, we created a mutant library of the 6K2 gene of TEV that contains every possible single- nucleotide substitution. 6K2 protein anchors the virus replication complex to the network of endoplasmic reticulum membranes. The library was inoculated into the natural host *N. tabacum*, allowing competition among all these mutants and selection of those that are potentially viable. We identified 11 nonsynonymous mutations that remain in the viral population at measurable frequencies and evaluated their fitness. Some had fitness values higher than the wild-type (WT) and some were deleterious. The effect of these mutations in the structure, transmembrane properties, and function of 6K2 was evaluated *in silico*. In parallel, the effect of these mutations in infectivity, virus accumulation, symptoms development, and subcellular localization was evaluated in the natural host. The  $\alpha$ -helix H1 in the N-terminal part of 6K2 turned out to be under purifying selection, while most observed mutations affect the link between transmembrane  $\alpha$ -helices H2 and H3, fusing them into a longer helix and increasing its rigidity. In general, these changes are associated with higher within-host fitness and development of milder or no symptoms. This finding suggests that in nature selection upon 6K2 may result from a tradeoff between within-host accumulation and severity of symptoms.

## **I. 1. INTRODUCTION**

### **I. 1.1. Mutation, Selection, and Mutant Swarms in RNA Virus Populations**

RNA viruses exist as complex mutant swarms that result from the combination of three factors: high mutation rates, very short generation times, and very large population sizes (Domingo et al., 2012). These mutant swarms are commonly referred in the virological literature as viral quasispecies (Domingo et al., 2012). Quasispecies theory represent a twist of the classic mutation-selection balance concept from population genetics (Wilke 2005) in which high mutation rates ensure the coupling between genotypes and thus selection operates not on the individual genotype but on the cloud of mutants that are all linked by one or few mutational steps (Bull et al., 2005). The mutant swarm is dominated by a master sequence with the highest fitness surrounded by a cloud of mutants in frequencies that rank according to their fitness. One of the principal tenants of the quasispecies theory is that viral populations replicate close to the so-called error threshold, that is, the highest mutation rate compatible with maintaining genetic information and which is usually proportional to the inverse of the genome length (Bull et al., 2005; Domingo et al., 2012). Increases in mutation rate over the error threshold push the viral populations into a regime known as the error catastrophe in which the frequency of genotypes in the population is not proportional anymore to their fitness, the master sequence disappears and the integrity of genetic information vanishes out (Bull et al., 2005; Domingo et al., 2012). Indeed, this principle provides the theoretical background for antiviral therapies based on lethal mutagenesis (Bull et al., 2007; Perales et al., 2011). The quasispecies population structure is supposed to bestow great adaptive

potential to viral populations as the mutant swarm contains genetic variability to respond to environmental fluctuations (Domingo et al., 2012). However, whether such high mutation rates are adaptive *per se* or a side effect of a parasitic life style that favors fast replication at the cost of low fidelity has been a subject of debate (Elena and Sanjuán, 2005; Belshaw et al., 2008). Indeed, the evolutionary fate of mutant swarms and whether they may transit or not to error catastrophe by pushing up mutation rates depends on the proportion of all possible mutations that are lethal, deleterious, neutral, or beneficial (Sardanyés et al., 2014) and on the topography of the underlying fitness landscape (*i.e.*, epistasis; Elena et al., 2010). A considerable amount of effort has been devoted to characterizing the distribution of mutational fitness effects across the genome of RNA viruses by generating collections of random single-nucleotide substitution mutants (*e.g.*, Sanjuán et al., 2004; Carrasco et al., 2007; Visher et al., 2016). However, given the cost of creating such collections, they are limited in size.

To circumvent this limitation and to explore the selective constraints upon an essential viral protein in a high-throughput manner, here we have taken an approach inspired by recent advances in experimental evolution that allow quantitatively tracking the evolutionary dynamics of individual lineages at high resolution (*e.g.*, Levy et al., 2015). Instead of generating mutants by site-directed mutagenesis of a handful of candidate positions, we created *in vitro* a library of mutants that contains all possible single-nucleotide substitutions plus a large number of variants containing more than one mutation of a particular coding gene. The mutant library was then inoculated in a susceptible host and allowed natural selection to do the work of fishing out viable mutants and washing out highly deleterious and lethal ones.



The choice of the viral protein to be studied was based on two principles: (1) It has to play a fundamental role during infection so we expect large fitness effects and (2) it has to be of small enough size so every possible mutation can be created and then the entire coding sequence determined in a single amplicon using Illumina next-generation sequencing, hence linkage among mutations can be directly preserved. We have chosen the 6K2 protein encoded in the genome of TEV. TEV genome structure is shown in Fig. 2. The 6K2, which receives its name from its molecular weight of 6 kDa, is a small protein of only 53 amino acids with a transmembrane region encompassing residues about 20 - 42 (Schaad et al. 1997). There is another protein with similar characteristics, named 6K1. Both proteins are indispensable for viral replication (Merits et al., 2002; Cui and Wang, 2016).

### **I. 1.2. Potyvirus 6K2: An Essential Component of the Virus Replication Complex**

Positive-strand RNA viruses replicate on intracellular vesicles that often result from extensive rearrangements of the endoplasmic reticulum (ER) membrane network (Salonen et al., 2005; Miller and Krijnse-Locker, 2008; Den Boon and Ahlquist, 2010). The role of these membranous vesicles is manifold. First, they provide the scaffold for anchoring the so-called virus replication complex (VRC) to cellular membranes. Second, by restricting the replication process to a small and closed region of the cytoplasm, the concentration of relevant molecules can be high (*e.g.*, replication factors, co-opted host factors, and nucleotides). Third, these vesicles provide a safe environment to protect double-stranded RNA replication intermediates from the action of the DICER endonucleases that trigger the RNAi-mediated cellular antiviral system. Interestingly, some had argued that vesicles containing VRCs composed by viral proteins interacting with a large number of cellular factors, altogether dubbed as virus replication factories

(Den Boon and Ahlquist, 2010), should be considered as the real alive virus (the virion being just a transmission inert stage) (Claverie, 2006) and that such virus factories may even be the remote ancestor of the eukaryotic nucleus (Bell, 2001; Koonin and Dolja, 2013).

6K2 is able to create vesicles by itself even though it needs to interact with other viral and cellular proteins to create functional VRC (Löhmus et al., 2016; Geng et al., 2017). Purifying high-molecular-weight complexes involving 6K2 and using proteomic techniques to identify components, Löhmus et al. (2016) found that 6K2 associates with the potyviral NIb replicase, CI helicase, HC-Pro suppressor of RNA silencing, and the VPg that attaches covalently to the 5' end of the genomic RNA. In addition, they identified a number of host proteins within these complexes: for example, the isoform 4E of the eukaryotic initiation factor, the poly(A)-binding protein, the eukaryotic elongation factor 1A, the ATP-dependent Clp protease, and the heat-shock cognate 70-1 and 70-3 proteins (Löhmus et al., 2016). Using direct protein-protein interaction techniques [*e.g.*, yeast two-hybrid, coimmunoprecipitation or bimolecular fluorescence complementation (BiFC)], several other host proteins have been identified as 6K2 interactors, including the photosystem II oxygen evolution complex protein NtPsbO1 (Geng et al., 2017), the COPII coatomer component Sec24a protein (Jiang et al. 2015), and the dynamin-related proteins AtDRP2A and AtDRP2B that detach mature clathrin-coated vesicles from plasma membrane (Wu et al., 2018) that are all recruited to the potyvirus VCR.

6K2 anchors the VRC vesicles to the ER membranes (Restrepo-Hartwig and Carrington, 1994; Schaad et al., 1997; Cotton et al., 2009) and then migrates to the Golgi apparatus in a COPII-dependent manner (Wei et al., 2010; Agbeci et al., 2013; Jiang et

al., 2015). Then, the Golgi-associated 6K2 vesicles move along the actomyosin microfilaments until docking on the outer chloroplast envelopes, wherein they induce membrane invaginations (Cotton et al., 2009; Wei et al., 2010). Cabanillas et al. (2018) have recently shown that impairing the traffic between ER and Golgi by overexpression of the SNARE Sec22 protein resulted in enhanced intracellular virus movement and concluded that TuMV replication vesicles bypass the Golgi and take an unconventional pathway for infection that may involve prevacuolar components. Furthermore, 6K2-induced vesicles are involved in cell-to-cell (Grangeon et al., 2013) and long-distance systemic movement of virus-containing vesicles (Rajamäki and Valkonen, 1999; Spetz and Valkonen, 2004; Wan et al., 2015) reminiscent of the egression of animal viruses from cells (Den Boon and Ahlquist, 2010). Finally, it has been shown that vesicles are generated from single viral genomes and grow in size as RNA replication and translation continues within the vesicle and viral proteins accumulate in situ (Miller and Krijnse-Locker, 2008; Cotton et al., 2009).

Despite the relevance of 6K2 in the infectious cycle of potyviruses, very little is known about the selective constraints operating upon this small protein. In the case of potato virus A (PVA), Rajamäki and Valkonen (1999) have shown that mutation M5V in the N-terminal region of the protein is enough to revert the nonsystemic infection phenotype of isolate PVA-M into a virus capable of spreading systemically. The N-terminal sequence of the 6K2 from TEV contains a conserved diacidic D(X)E motif (Lerich et al., 2011) that is involved in interactions with the COPII coat protein Sec24 in the ER and with the Golgi protein Man1, suggesting a role in trafficking between these two organelles (Hanton et al., 2005). Likewise, mutations in the N-terminal region 6K2 affect the union with Sec24a, blocking export to the ER and thus precluding intra- cellular and cell-to-cell

movement (Jiang et al., 2015). Finally, the 6K2 transmembrane GxxxG motif is key for TuMV infection, and replacement of glycine by valine precludes virus accumulation as a consequence of the relocation of the VRC to the Golgi and the plasma membrane (Cabanillas et al., 2018).

### **I. 1.3. Short Overview of This Study**

The 6K2 mutant library was cloned into an infectious plasmid containing the TEV genome. Then this library was inoculated (by itself or in combination with a fraction of WT genomes) into tobacco plants and infection started. As said above, we expected natural selection to fish out all possible mutations that were viable and not too different in fitness from WT. Nine days post-inoculation (dpi), the resulting viral progenies were characterized by Illumina next-generation sequencing and a number of amino acid replacements identified and ranked according to their frequency in the population; in vivo fitness was predicted from the observed changes in frequency. The impact of these mutations in the 6K2 structure and functionality has been evaluated in silico. Finally, we have determined the effect of some of these mutations in TEV infectivity, accumulation, symptomatology, and in 6K2 subcellular location in *N. tabacum*.

## **I. 2. MATERIALS AND METHODS**

### **I. 2.1. Synthesis of 6K2 Variants Library**

A 206-nt long oligonucleotide was synthesized by TriLink Biotechnologies Inc. with the following design: 159 nucleotides (nt) mutated region of 6K2 gene with the 98.8% probability of having WT sequence and 0.4% equal probability of having all three other

possible substitutions per position. The immediate upstream 25-nt and down-stream 22-nt sequences of mutated regions were synthesized as the WT sequence. The synthesized oligo sequence is 5'-TCACCTGGAAACTATCTATCTCCAAtcagatagcgaagtggttaagcatctgaagcttaaagtcactggaataaaagccaaatcactagggacatcataatagctttgtctgtgtaattggtggtggatggatgcttgaacgtacttcaagacaagtcaatgaaccagtctatttccaaGGGAAGAAGAATCAGAAGCACA-3', with the uppercase in the sequence indicating the constant flanking regions of the mutated sequence. In between, the lowercase shows the mutated region. Lowercase "a" indicates 98.8% A, 0.4% G, 0.4% C, and 0.4%T; lowercase "g" indicates 98.8% G, 0.4% A, 0.4% C, and 0.4% T; lowercase "c" indicates 98.8% C, 0.4% A, 0.4% G, and 0.4% T; and lowercase "t" indicates 98.8% T, 0.4% A, 0.4% G, and 0.4% C.

## **I 2.2. Plasmid Library Preparation**

To start with, 1 mg of dry mutagenic oligos were dissolved in 500  $\mu$ l Milli-Q water as a stock, and further diluted to about 0.5 ng/ $\mu$ l as a working stock. The mutagenic oligo was mixed equimolar with two short DNA fragments overlapping with the mutagenic oligo constant regions for an overlap-extension polymerase chain reaction (PCR). The purpose of this overlap-extension PCR was to introduce *Eco*NI and *Bam*HI restriction enzyme recognition sites for the following ligation step as well as increasing the insert length for the higher efficiency of ligation (Fig. 3). The PCR amplification was performed using Phusion hot-start polymerase (Thermo Scientific), first 10 cycles without adding any primers and the next 10 cycles with primers to amplify the full-length products combining all three fragments (Supplementary file 1 at [doi.org/10.1093/gbe/evz069](https://doi.org/10.1093/gbe/evz069)). The PCR product was purified with MinElute PCR Purification kit (Qiagen) after confirming the size of PCR product is correct on a 2% agarose gel. Then, the PCR amplicon was double-

digested with *Eco*NI and *Bam*HI-HF enzymes for 1.5 h at 37 °C. Correct size band (449nt), as an insert for later ligation, was retrieved from agarose gel under the blue light illumination. QIAEX II Gel Extraction kit (Qiagen) was used to purify the DNA from the agarose. The linearized vector was prepared from pMTEV plasmid (Bedoya and Daròs, 2010) by cutting and recovering the 10,806-nt fragment from the gel after *Eco*NI and *Bam*HI- HF enzymes double digestion, the same way as the insert preparation. ElectroLigase (New England BioLabs Inc.) was used for ligating the two fragments following the manufacturer's instructions. Ligated products were directly transformed into NEB 10-beta Electrocompetent *Escherichia coli* cells (New England BioLabs Inc.). After recovery in super optimal complete (SOC) for 1 h at 37 °C, 200 µl of each SOC media with cells was plated on Luria-Bertani broth (LB) + ampicillin plates. To estimate the total number of transformants, a range of small aliquots of SOC was plated as well on LB þ ampicillin plates. We collected an estimate of 0.78-1.2 million transformants from the experiment. Cells were collected from the LB + ampicillin plates after an overnight incubation at 37 °C, with phosphate-buffered saline (PBS) with 1 mM ethylenediaminetetraacetic acid (EDTA), and plasmids were prepared directly from the cell pellets collected from the overnight plates using Plasmid Midiprep kit (Qiagen). Plasmid concentration was quantified using NanoDrop 2000 (Thermo Scientific).

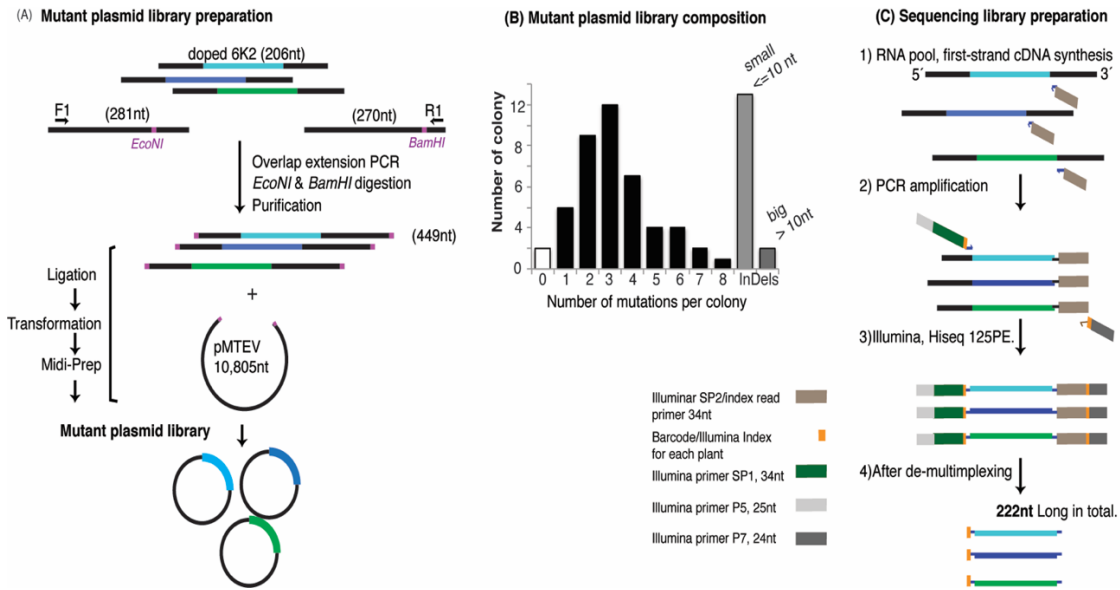
To verify the quality of the library, 96 colonies were picked randomly for Sanger sequencing after amplifying the mutated region by colony PCR (primers provided in Supplementary file 1 at [doi.org/10.1093/gbe/evz069](https://doi.org/10.1093/gbe/evz069)). Sanger sequencing was performed at UPF Genomics Core Facility ([www.upf.edu/ web/sct-genomics](http://www.upf.edu/web/sct-genomics); last accessed May 4, 2016). The number of mutations and their distributions per colony were examined (Fig. 3B).

### **I 2.3. Plant Inoculation**

The mutant plasmid library was linearized by digestion with *Bgl*III prior to *in vitro* RNA synthesis using the mMESSAGE mMACHINE SP6 Transcription Kit (Ambion Inc.) following the manufacturer's instructions to obtain 50-capped infectious RNA, as described in Carrasco et al. (2007). The third true leaf of 4-week-old *N. tabacum* plants was mechanically inoculated with 20 µg of transcribed RNA suspended in 500 µl inoculation buffer (50 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.0, 3% PEG6000, 100 mg/mL Carborundum). All symptomatic tissue was collected 9 dpi and stored at -80 °C until analyzed. RNA was extracted from 100-mg homogenized infected tissue using the InviTrap Spin Plant RNA Mini Kit (Strattec Molecular GmbH).

### **I 2.4. Illumina Library Preparation**

RNA was treated with Turbo DNase and made into first-strand cDNA with SuperScriptIII kit (Thermo Scientific). The primer for this first step is a target-specific primer with an overhang of SP2 Illumina sequence for cluster generation for HiSeq (Illumina Inc., San Diego, CA) (Fig. 3C). After treatment with RNaseH, cDNA was amplified with Phusion polymerase (Thermo Scientific) for 25 cycles with Illumina-barcoded primers. In this step, the same forward primer (with an overhang of 6-nt barcode and Illumina P5/SP1) but a different reverse primer (Illumina P7 sequence and sample-specific Illumina barcodes) were used for each sample. Correct size PCR products were collected using the 2% size-select e-gel system, and further desalted with MiniElute PCR Purification kit (Qiagen). Sample libraries were multiplexed into one HiSeq 2500 System (Illumina Inc.) lane at CRG Genomics Core Facility for the 125 pair-end sequencing, for a total read count of 3.1-4.2 millions per sample.



**Fig. 3.** Strategy for cloning and sequencing the “doped” 6K2 library into the infectious clone pMTEV by replacing the WT version of 6K2 by the mutant library. (A) Cloning process of the mutated amplicons into the linearized pMTEV plasmid. (B) Distribution of number of mutations per plasmid transformant. (C) Sequencing library preparation strategy.

According to the phiX spike-in error rate given by the CRG Genomics Core Facility, the technical error of the HiSeq varies between 0.3% and 1% per site.

## I 2.5. Bioinformatics Pipeline for Analysis of NGS and TEV 6K2 Quasispecies Reconstruction

Read artifact filtering and quality trimming (30 minimum Q28 and minimum read length of 50 bp) was done using FASTX- Toolkit version 1.01 as implemented in Galaxy (Afgan et al., 2018) and with default parameters. In addition, FASTQ datafiles were also transformed into SAM formatted files using FastqToSam version 2.7.1.0 and BAM-to-SAM version 2.0 as implemented in Galaxy (Afgan et al., 2018) and with default parameters. In any case, reads containing undefined nucleotides (N) were discarded. Two



different algorithms were used to reconstruct TEV 6K2 quasispecies. First, QuRe version 0.99971 (Prosperi and Salemi, 2012), with default parameters, that uses a heuristic algorithm which matches multinomial distributions of distinct viral variants overlapping across the 6K2 sequence. QuRe incorporates a built-in Poisson error correction method and a postreconstruction probability clustering. The input for QuRe is a FASTA file containing all aligned reads and a reference sequence (also in FASTA format). Second, aBayesQR (Ahn and Vikalo, 2018), also with default parameters, that uses a maximum-likelihood framework to infer individual sequences in a mixture via agglomerative hierarchical clustering and has been proved to be particularly efficient detecting low frequency variants. The input for aBayesQR is a SAM file and a reference sequence (in FASTA format). Only variants detected by both methods were retained for subsequent studies. As 6K2 reference sequence to align the reads in both algorithms, we used GenBank accession DQ986288, which corresponds to the TEV-7DA strain used in this experiment.

The fitness value of each selected variant,  $W_i$ , relative to the rest of the variants in the quasispecies was evaluated as described in Carrasco et al. (2007). In short,

$$W_i \approx \left\{ \frac{p_{i,t}(1 - [fp_{i,lib,0} + (1-f)p_{i,WT,0}])}{(1-p_{i,t})[fp_{i,lib,0} + (1-f)p_{i,WT,0}]} \right\}^{1/t} \quad (\text{Eq. 1})$$

where  $p_{i,t}$  is the frequency of the  $i$ -variant  $t$  dpi,  $p_{i,lib,0}$  the frequency of the same variant in the mutant library,  $p_{i,wt,0}$  the frequency of the variant in the WT population, and  $f$  the fraction of the library in the inoculated mixture.

## **I 2.6. *In Silico* Structure Predictions and Probability of Membrane Association**

Three-dimensional structure predictions were created for WT and mutant 6K2 proteins using the RaptorX server (Källberg et al., 2012; raptorx.uchicago.edu; last accessed May 4, 2019) and then visualized and annotated using tools available at the Jena3D server (Hühne et al., 2007; jena3d.leibniz-fli.de; last accessed May 4, 2019). In addition, transmembrane helices were predicted for the different 6K2 sequences using the methods implemented in the Trans-Membrane Hidden Markov Model (TMHMM) server version 2.0 (Sonnhammer et al., 1998; www.cbs.dtu.dk/services/TMHMM/; last accessed May 4, 2019).

Comparisons between predicted structures for the WT and the mutants 6K2 proteins were performed in the MArkovian TRAnSition of Structure evolution server (Kawabata, 2003; strcomp.protein.osaka-u.ac.jp/matras; last accessed May 4, 2019). Two different measures of structural similarity were obtained.  $R_{dis}$  represents the normalized structural similarity index and ranges between 0% (no overlap in structures) and 100% (complete overlap).  $dRMS$  represents the root-mean square deviation (in Å) of distances between C $\beta$  atom positions of aligned residues; the larger the  $dRMS$  value, the less overlap among structures (Kawabata and Nishikawa, 2000).

The functional effects of the different 6K2 sequence variants detected in this study were explored in silico using the machine learning tools provided in the Screening for Nonacceptable Polymorphisms (SNAP2) web server (Hecht et al., 2015; rostlab.org/services/snap2web/; last accessed May 4, 2019). Using information about the biophysical amino acid properties, sequence, predicted secondary structure, residue

flexibility, solvent accessibility, PFAM, PROSITE and SWISS-PROT annotations, predicted binding residues, predicted disordered and low-complexity regions, proximity to N- and C-termini, and statistical contact potentials, SNAP2 provides a score for all possible variants at each residue of the 6K2 protein. The score ranges from -100 in the case of no effect to 100 in the case of maximal effect on the function, regardless it is positive or negative in terms of TEV fitness (Hecht et al., 2015).

### **I 2.7. Site-Directed Mutagenesis and *In Vitro* Transcription**

The infectious clone pMTEV contains a full copy of the genome of a TEV-7DA strain isolated from tobacco (GenBank accession DQ986288; Bedoya and Daròs, 2010). Eleven mutant genotypes were constructed by site-directed mutagenesis starting from template pMTEV plasmid. Mutagenesis was done using the *Pfu*Turbo DNA polymerase (Stratagene), and following the manufacturer's instructions using the pairs of mutagenic primers listed in Supplementary Table S1. After Sanger sequencing the mutant genotypes, infectious 50-capped RNAs were generated in vitro as described above. RNA integrity and quantity were assessed by gel electrophoresis.

### **I 2.8. Transient Expression of Mutant 6K2 Proteins in *Nicotiana benthamiana* Leaves and Confocal Laser-Scanning Microscopy**

To express 6K2 fused to the yellow fluorescent protein (YFP), the 6K2 cDNA was amplified from pMTEV plasmid with *Pfu*Turbo DNA polymerase (Stratagene) and specific primers including Gateway adapters, and recombined into pDONR207 using BP ClonaseMixII kit (Invitrogen). After sequencing, 6K2 cDNA was recombined into pEarleyGate101 vector (Invitrogen) using LR ClonaseMixII kit (Invitrogen). The same

direct site mutagenesis described above was done to obtain all 6K2 mutant genotypes in this plasmid.

*Agrobacterium tumefaciens* C58 cultures harboring relevant binary constructs were centrifuged and suspended in 10-mM MES pH 5.6, 10-mM MgCl<sub>2</sub>, 150-mM acetosyringone and OD<sub>600</sub> was adjusted to 1. Transient expression was performed by agroinfiltration into *N. benthamiana* leaves. After 2 days, fluorescence was analyzed in infiltrated leaves using an inverted Zeiss LSM780 inverted confocal microscopy with a CAPO 40x/1.2 objective (Carl Zeiss MicroImaging GmbH). YFP-derived fluorescence was monitored by excitation with 488-nm argon laser, and detection windows of 520-550nm. Imaging processing was performed by ImageJ version 1.8.0\_172 (Schneider et al., 2012; [image-j.nih.gov/ij](http://image-j.nih.gov/ij); last accessed May 4, 2019).

### **I 2.9. Plant Inoculations, Virus Purification, Phenotyping of Infections, and Quantification of Infectious Viral Load**

All the inoculations were performed in the virus natural host *N. tabacum* plants. Batches of ten 8-week-old plants were inoculated with ~5 mg RNA of each viral genotype by abrasion of the third true leaf with 10% carborundum (100 mg/ml). Plants were maintained in a BSL-2 greenhouse chamber at 25 °C under a 16 h natural sunlight (supplemented with 400 W high-pressure sodium lamps as needed to ensure a minimum light intensity of PAR 50 μmol/m<sup>2</sup>/s) and 8-h dark photoperiod.

After inoculation, plants were visually observed every day for the presence and severity of symptoms and the number of symptomatic plants recorded. A plant was considered as infected if it showed visible symptom of TEV infection. Two different

pathogenicity-related traits were estimated from these data. First, the percentage of symptomatic plants at the end of the experiment (*i.e.*, 20 dpi) was used as an estimate of infectivity (*i*). This estimate should be taken as a lower limit of the real *i*, as asymptomatic plants yet infected plants would be missed. However, it has been confirmed in many previous studies that there is a one-to-one association between infection and symptoms development (*e.g.*, Lafforgue et al., 2012), henceforth, we are confident our estimate based on symptoms is not largely deviating from the actual value. Second, infectivity time-series data were submitted to a Kaplan-Meier regression analysis of survival times and the median time to the appearance of symptoms ( $ST_{50}$ ) estimated.

Nine dpi virus-infected leaves and apices from each symptomatic plant were collected and this tissue was frozen with liquid N<sub>2</sub> and homogenized using a Mixer Mill MM 400 (Retsch GmbH). Sap was prepared by adding 1mL of 50 mM potassium phosphate buffer (pH 7.0) per gram of homogenized tissue. For each sap sample 1:1, 1:5, 1:10, 1:50, 1:100, 1:500, and 1:1,000 serial dilutions were done, inoculating with 10  $\mu$ l of each dilution four replicates in independent leaves of 4-week-old *C. quinoa* plants. Extra care was taken to always inoculate leaves of the same developmental stage. Infectious viral load, measured as the number of lesion-forming units (LFUs) per  $\mu$ L of inoculum, were inferred from the regression of the observed number of local lesions to the dilution factor (Kleczkowski, 1950). These experiments were reproduced in two independent blocks in consecutive years (started October 2, 2017 and January 29, 2018, respectively). All mutants were assayed simultaneously in both blocks.

All raw data are available at LabArchives under doi:10.25833/9kqb-e545.

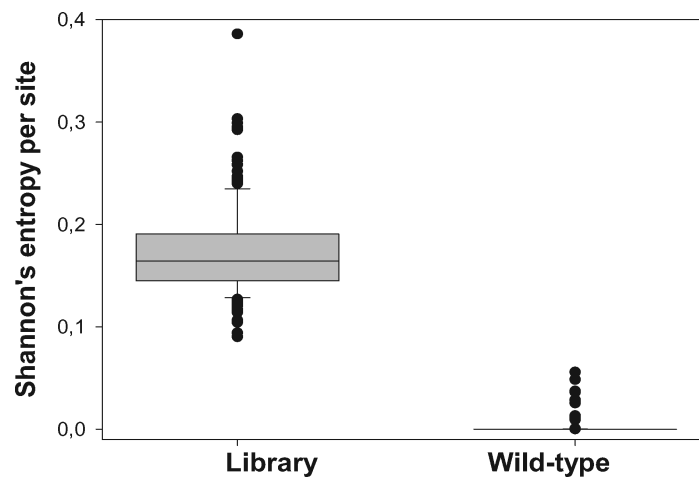
## **I 3. RESULTS**

### **I 3.1. Variants Library Characterization and Bulk Selection Experiments**

First, we assessed the 6K2 allelic composition present in a TEV population resulting after infecting a *N. tabacum* plant with the WT virus. To do so, we sequenced it using a HiSeq 2500 System (Illumina Inc.). We found 331 different 6K2 variants (haplotypes) out of 2,216,431 valid reads: 98.67% were of the WT sequence, five other variants had a frequency  $\geq 0.10\%$  and two more had a frequency  $> 0.01\%$ . The per site values of Shannon's entropy (Shannon, 1948) ranged between 0 (completely monomorphic site) and 0.0557 with a median value of 0 (Fig 3).

As expected, this genetic composition was in sharp contrast with the composition and variability of the mutant library, wherein we have identified 1,029,895 different genetic variants out of 2,100,648 valid reads obtained. Twenty-five percent of variants contained indels, and the number of point mutations per molecule ranged between 0 (3% of molecules) and 8 (1% of molecules), with a mean of 3.19 (median of 3) and a standard deviation of 1.75 mutations per variant. Individual variants ranged in frequency between 1.62% and  $9.71 \times 10^{-5}\%$ ; 42 variants had a frequency  $\geq 0.10\%$ , 112 a frequency  $\geq 0.05\%$ , 797 a frequency  $\geq 0.01\%$ , and all the rest had lower frequencies. In this case, the per site values of Shannon's entropy ranged between 0.0903 and 0.3860 with a median value of 0.1644 (Fig. 4), being the difference with the WT variability highly significant (Wilcoxon signed rank test  $V = 12246$ ,  $P < 0.0001$ ).

When the library was inoculated in 25 tobacco plants, none resulted in symptoms, whereas all 10 plants inoculated with the WT virus developed the normal symptoms associated with TEV infection 9 dpi. Given the tremendous genetic diversity in the library, this lack of infectivity supports the concept of the quasispecies error catastrophe (Domingo et al., 2012) and suggests that the frequency of the WT 6K2 sequence, plus any other viable sequences, was too low in the mutant library to establish infection at all or to sustain population growth and hence the virus extinguished as a result of entering into the error catastrophe regime (Bull et al., 2007; Domingo et al., 2012).



**Fig. 4.** Average per site variability, measured as Shannon's entropy, for the variants library and the WT TEV population.

In an attempt to avoid this error catastrophe effect, we generated mixtures of the library and RNA transcribed from the WT infectious clone pMTEV with increasing ratios of the library: 50%, 75%, and 95%. Inoculation with all three mixtures resulted in symptomatic infections (ten plants each). After purifying the resulting viral populations 9 dpi and deep sequencing them as above, the number of valid reads were 3,342,356, 3,840,036 and 2,907,325, respectively. In the three populations, the frequency of the WT 6K2 were very similar: 90.4%, 90.6%, and 90.5%, respectively, and significantly lower than observed in

the WT population (see previous paragraph), meaning that a number of alternative variants coexisted in these mixed populations at noticeable frequencies. Using two different quasispecies-reconstruction algorithms (QuRe and aBayesQR; see the Materials and Methods section for details), we identified 23 variants with frequencies in the range 0.32-0.01% depending on the particular inoculation experiment (Table 2). Some of the mutations (though not necessarily the variants) were recurrently observed in different inoculated plants (*e.g.*, nonsynonymous mutations U95G, G101A, G104U, and C132A and synonymous mutation C157A). Another remarkable observation is that 18 of these 23 viable mutant alleles affect the stretch of amino acid residues from 32 to 39, all belonging to the predicted transmembrane domain of the 6K2 protein.

Using the observed change in frequency of all 23 variants between the inoculation mixture and the viral population recovered 9 dpi, it was possible to calculate their relative fitness values in the system (see the Materials and Methods section for details). Estimated fitness values were in the range 0.455-1.826 (median = 0.780), with a mean value of 0.8456 0.398 (Table 2). Five variants have relative fitness higher than WT: G34S/A39S, I32S/G33D, G34R/G35A, G34D, and D44E/F46L (Table 2). Four variants only contain synonymous changes. Eleven of the variants containing nonsynonymous changes, including the five with beneficial fitness effects, were introduced by site-directed mutagenesis into the WT infectious clone pMTEV and their biological properties further characterized.



**Table 2.** Twenty-Three 6K2 Variants Identified by Illumina Sequencing after Infection

<b>Amino Acid Substitution</b>	<b>Nucleotide Substitution</b>	<b>Protein Localization</b>	<b>Type of Amino Acid Substitution</b>	<b>% of Library in Inocula</b>	<b>Relative Fitness</b>
<i>I32S</i>	U95G	Transmembrane	Nonpolar to polar	95	0.5122
<i>I32S</i>	U95A	Transmembrane	Nonpolar to polar	75	0.5837
<i>I32S/G33D</i>	U95G/G98A	Transmembrane	Nonpolar to acid	50	1.5669
<i>I32T</i>	U95C	Transmembrane	Nonpolar to polar	50	0.8056
<i>Sy<sup>a</sup></i>	IJ99G			95	0.5495
<i>Sy<sup>a</sup></i>	U99G/U102G		Conservative acid/conservative	50	0.9726
<i>Sy<sup>a</sup>/Sy<sup>a</sup></i>	U99G/U102G/ <i>D44/EF461</i>	C-terminal	hydrophobic	95	$\infty^b$
<i>G34C</i>	G100U	Transmembrane	Nonpolar to polar	50	0.8869
<i>G34D</i>	G101A	Transmembrane	Nonpolar to add Conservative	95	1.1500
<i>G34V</i>	G101U	Transmembrane	nonpolar Nonpolar to basic/conservative	50	0.7607
<i>G34R/G35A</i>	G101C/G104C	Transmembrane	nonpolar Nonpolar to polar/conservative	95	1.3716
<i>G34S/A39S</i>	G101NG116U	Transmembrane	polar Conservative	50	1.8255
<i>G35V</i>	G1041J	Transmembrane	nonpolar	75	0.5239

				95	0.5319
			Hydrophobic to		
W36C	G108U	Transmembrane	polar	95	0.5194
<i>A39E</i>	C116A	Transmembrane	Nonpolar to add	95	0.4547
			Conservative		
<i>A39V</i>	C116U	Transmembrane	nonpolar	95	0.4897
D44E	C132A	C-terminal	Conservative acid	50	0.8467
			Conservative		
			acid/hydrophobic		
<i>D44E/F46L</i>	C132A/C138A	C-terminal	to nonpolar	95	1.2397
Sy <sup>a</sup>	A147U			50	0.7986
Sy <sup>a</sup>	C157A			75	0.5053

Note: *Italic entries* indicate 11 haplotypes constructed by site-directed mutagenesis for further experiments.

<sup>a</sup> Sy = Synonymous

<sup>b</sup> Infinity results from an apparent zero frequency of the haplotype in the inoculation mix.

### I 3.2. Effect of Mutations on Predicted Protein Structure and Functionality

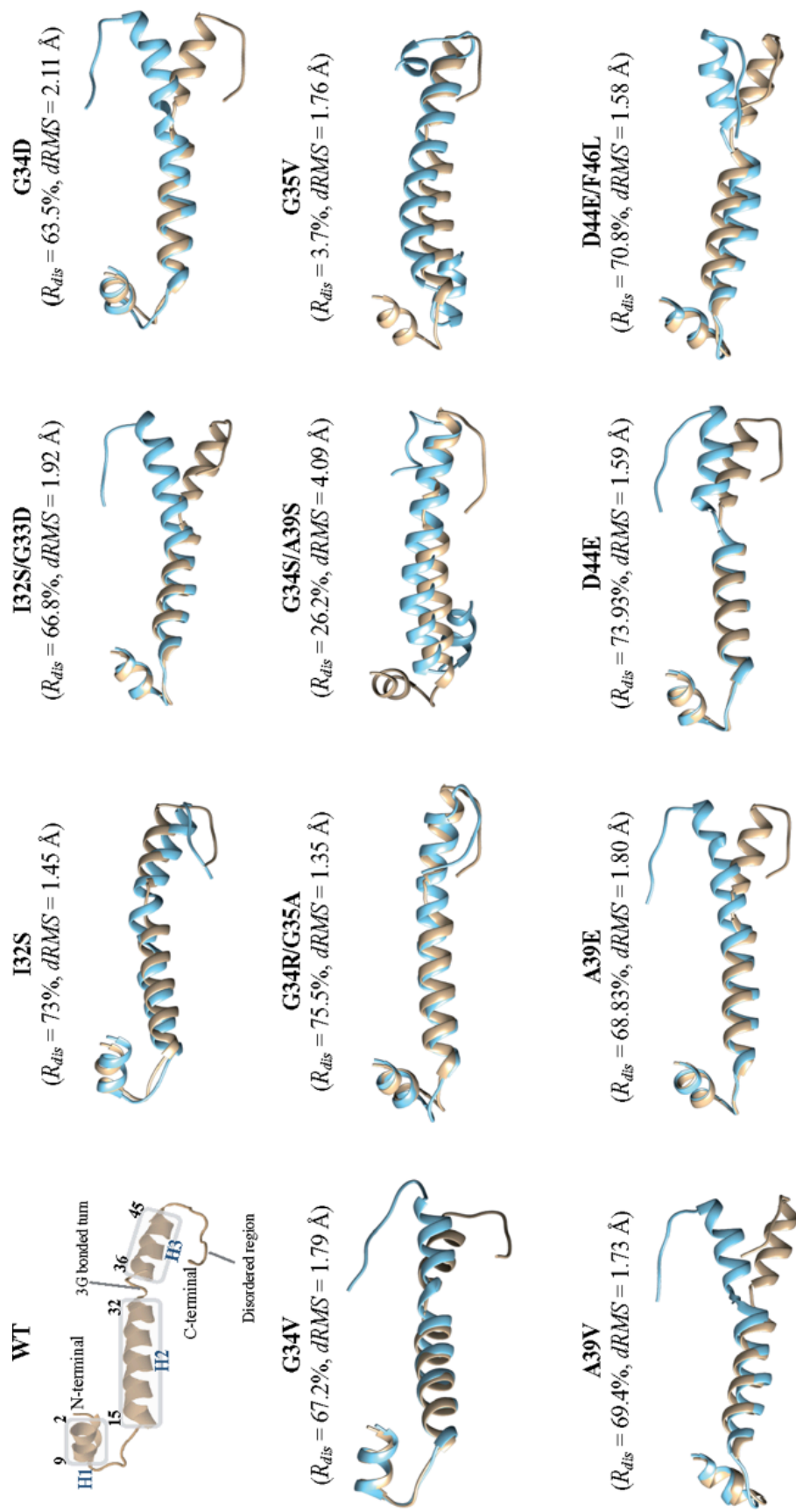
The ternary structure of the WT and mutant 6K2 proteins was evaluated using the RaptorX server (Källberg et al., 2012) and structural similarities evaluated using the MArkovian TRAnSition of Structure evolution server (Kawabata, 2003). Figure 5 shows the results of these studies. The predicted folding of WT 6K2 is characterized by the existence of three  $\alpha$ -helices, H1 (residues 2-9), H2 (residues 15-32), and H3 (residues 36-45); H1 and H2 are separated by a stretch of five amino acids with low structural complexity (residues 10-14) and H2 and H3 are separated by a bonded turn of three glycine (residues 33-35). The C-terminal part of H1 ends with amino acids K10 involved in a hydrogen bonded turn and L11 bending out. The C-terminal part of H3 ends with F46 also bending out. Residues 46-53 are a disordered region that may confer flexibility to

the C-terminal region of the protein. Besides some minor details that affect the C-terminal parts of H1 and H3, the 11 mutants can be classified into two categories according to the separation between H2 and H3: Those that preserve the three  $\alpha$ -helices and those that fuse H2 and H3 into a long helix. Mutants A39V, D44E, and D44E/F46L belong to the first category, whereas all others belong to the second (Fig. 5). Structural similarities were quantified using the  $R_{dis}$  and  $dRMS$  scores (Kawabata and Nishikawa, 2000). A large overlap corresponds to  $R_{dis} = 100\%$  and  $dRMS = 0$ , decreasing structural overlaps translate into smaller  $R_{dis}$  and larger  $dRMS$  values. G35V shows an inconsistency among the two values, having the smallest  $R_{dis}$  but not the largest  $dRMS$  values observed. Besides this particular case, the most dissimilar fold corresponded to the double mutant G34S/A39S and the most similar one to the double mutant G34R/G35A (Fig. 5).

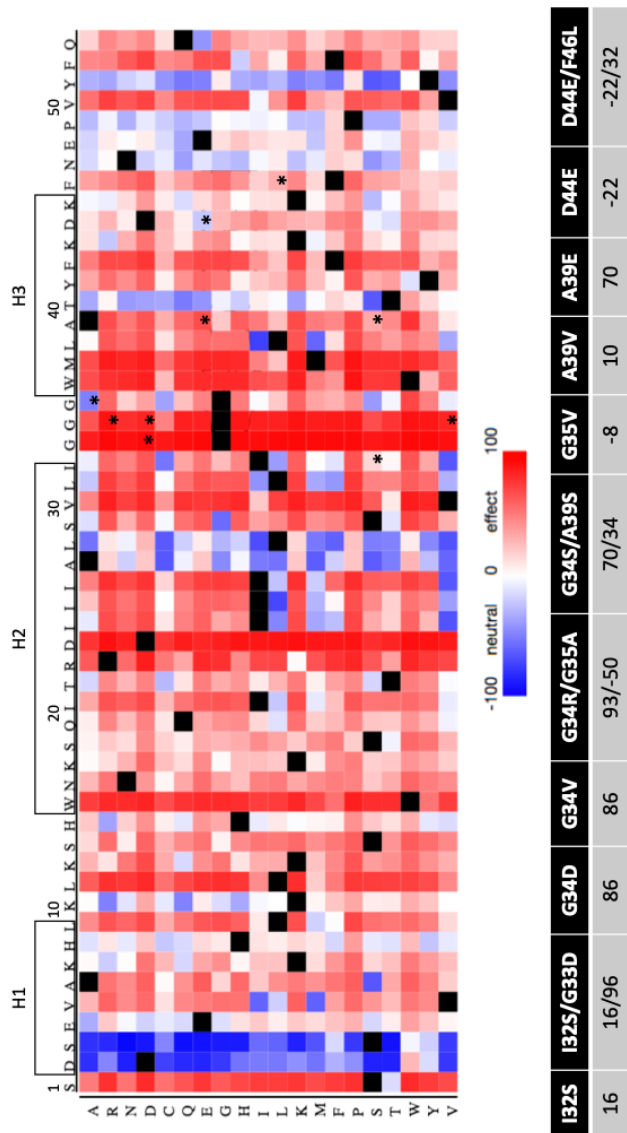
The TMHMM algorithm (Krogh et al., 2001) predicts that the WT 6K2 has a transmembrane domain encompassing amino acids 20-42. This transmembrane domain coincides with H2 and H3 predicted above and leaves H1 outside membranes. Nine of the eleven viable genotypes we examined contain mutations located in this transmembrane domain. TMHMM also shows that all eleven variants had WT-like predicted trans-membrane domains, thus confirming that the changes in folding discussed above do not affect the trans-membrane properties of the protein. Therefore, we reason that this domain is essential for the activity of the protein and thus a change in its conformation may have a strong impact in its functioning.

Next, we sought to evaluate *in silico* the expected effect of the observed mutations in the 6K2 function. To this end, we used the neural network-based classifier implemented in the SNAP2 webserver (Hecht et al., 2015). The aim of this computational approach is

not to address the question of whether or not an allele improves the fitness of TEV but just if it has a possible impact on 6K2 function. Figure 6 shows the results of this study. Interestingly, any amino acid replacement affecting residues W15 and D23 (H2), G33 and G34 (bond between H2 and H3), and W36 (H3) is predicted to have a strong effect on 6K2 function. By contrast, positions D2 and S3 (N-terminal part of H1), T40 (center of H3), and Y51 (disordered C-terminal region) are those with more tolerance to changes.



**Fig. 5.** Comparison of the predicted ternary structures for the WT (in gold) and the mutant (in blue) 6K2 peptides. As measures of structural similarity, the  $R_{dis}$  and  $dRMS$  values are reported. The upper most left panel shows the predicted structure of the WT 6K2 sequence; the three  $\alpha$ -helices and other relevant motifs are marked with boxes.



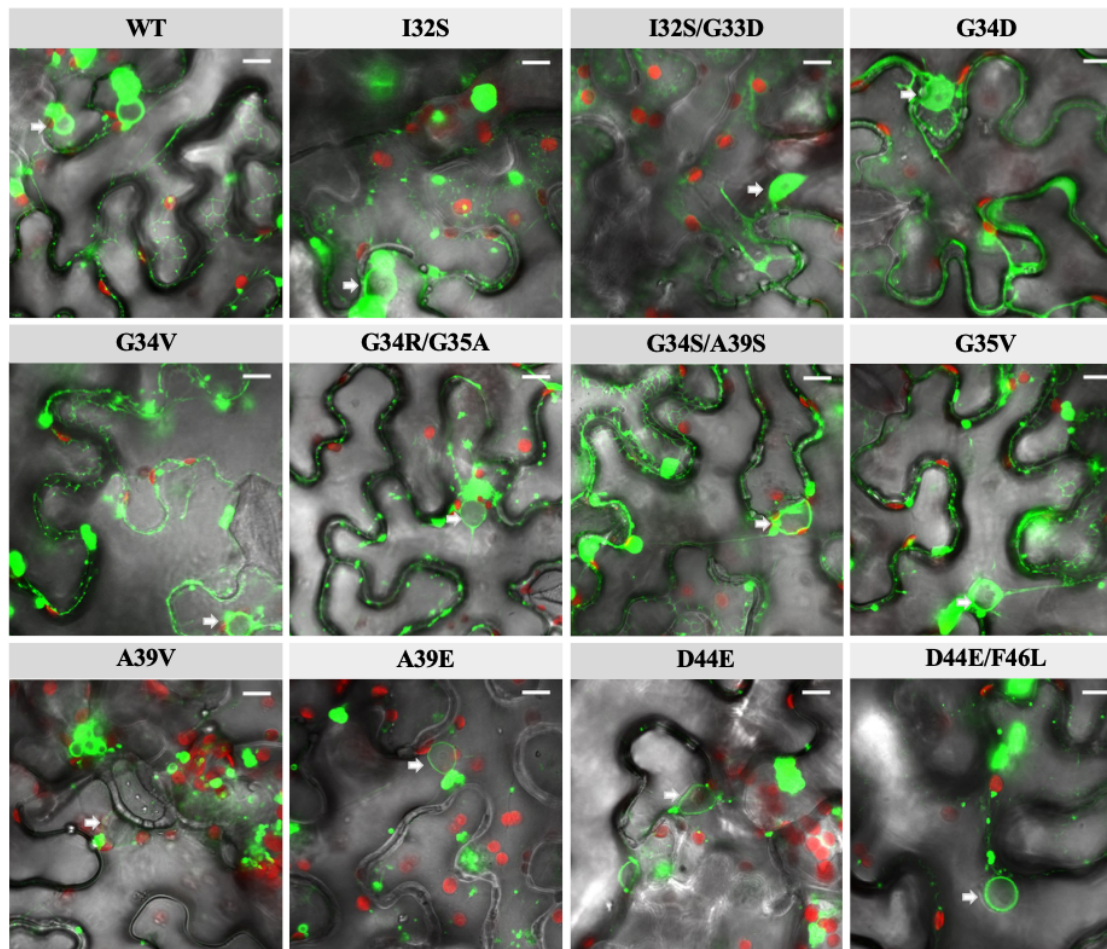
**Fig. 6.** *In silico* evaluation of the potential functional effect of every possible amino acid replacement on each residue of 6K2. Columns represent the 6K2 residues (indicated in the top) and rows the possible changes. Hot colors (red) represent strong functional effects, whereas cold colors (blue) represent neutral changes. Black squares represent no amino acid change. Mutations studies in this work are indicated with asterisks. The table below indicates the scores for each of the 11 6K2 mutants studied. Amino acids involved in  $\alpha$ -helices H1, H2, and H3 are indicated with boxes.

### I 3.3. Subcellular Localization of 6K2 Mutants

As mentioned in the Introduction, 6K2 has the ability to induce vesicles by itself and these vesicles form VRC by associating with other viral host proteins. As we have illustrated in the previous section, 6K2 is predicted to be hydrophobic and transmembrane associated and hence the vesicles it creates are localized in the ER membranes and around the nucleus and organelles (Cotton et al., 2009; Wei et al., 2010). We have engineered C-terminus YFP-tagged versions of WT and the eleven mutant proteins to explore, using confocal microscopy, whether mutations have an effect on the subcellular localization of

the proteins. 6K2-YFP was transiently expressed by agroinfiltration (see Materials and Methods). Figure 7 shows representative images for all the mutant genotypes. We found that nine of the mutants have the same subcellular localizations as the WT protein, localizing in perinuclear ER membranes (nuclei are pointed with arrows) and generating vesicles of cytoplasmic localization. By contrast, mutants I32S/ G33D and G34D show a different intracellular distribution: In addition to distribute homogeneously along the periplasmic membrane, they are also highly concentrated in the nucleoplasm (Fig. 7). Interestingly, these two genotypes have a predicted fused H2 and H3, whereas inducing a change in the angle of the region corresponding the H3 and the C-terminus of the protein respect to the WT configuration (Fig. 5). In concordance, these two genotypes also obtained large scores of functional changes in the SNAP2 analyses.

As an interesting corollary of these experiments, all mutations rescued from the mutant library are able to form vesicles and anchor to membranes, thus suggesting that strong selection for these two characteristics is at play.



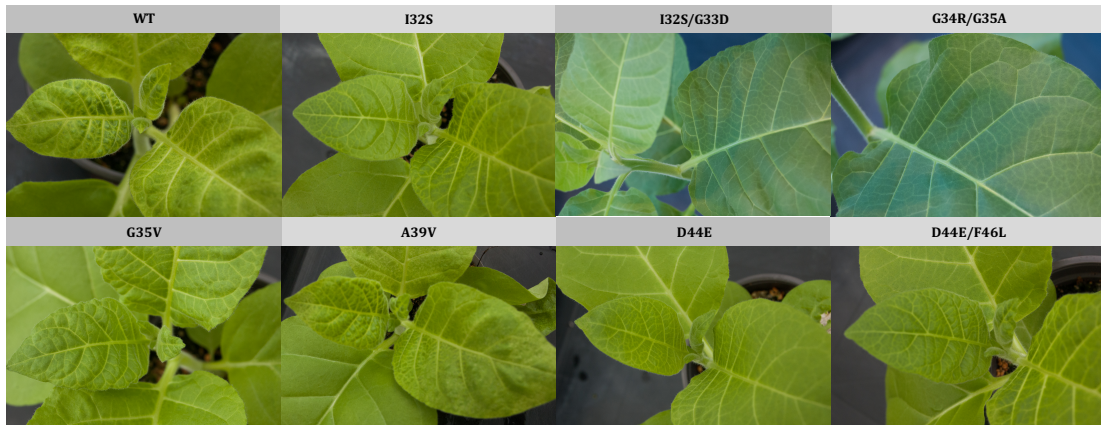
**Fig. 7.** Confocal microscopy imaging of *Nicotiana benthamiana* leaf epidermal cells expressing 6K2::YFP. Each image panel shows the localization of the WT protein and its variants (1-13). The position of nuclei is indicated with an arrow. Red objects are chloroplasts. White scale bar represents 10  $\mu\text{m}$ .

### I 3.4. Phenotypic Properties of 6K2 Mutants

To phenotype the 11 genetic variants of 6K2 studied in the previous sections in the natural host *N. tabacum*, two independent infectivity assays were performed using RNA transcripts generated *in vitro* as inocula. Five of the 6K2 variants (I32S, G35V, A39V, D44E, and D44E/F46L) were infectious in both assays. However, these variants showed differences among them in terms of symptomatology and progression of infection. As can be seen in Fig. 8, the TEV carrying different 6K2 mutants induced symptoms that were



distinguishable from those induced by the WT virus, all inducing symptoms that were equal or milder and that were visible at earlier or later dpi than those characteristic of the WT TEV.

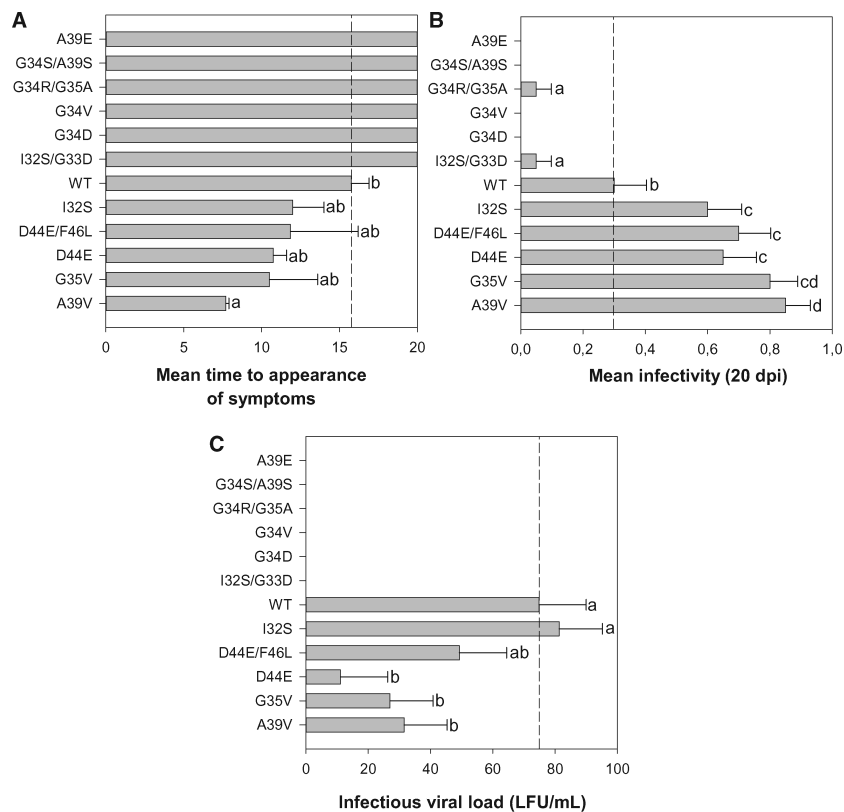


**Fig. 8.** Representative symptoms induced by TEV WT and mutant clones carrying the different 6K2 mutants. Plants were inoculated with equal amounts of RNA transcribed *in vitro* (see Material and Methods section). Pictures were taken 8 dpi for all the variants except for I32S/G33D and G34R/G35A that were taken 17 dpi.

Besides these differences in the symptomatology, the 6K2 mutants also induced visible symptoms at different dpi (Fig. 9A). A Kaplan-Meier survival analyses was performed to evaluate differences in the dynamics and timing of symptom development. This analysis showed that genotype A39V induced symptoms faster than the WT: on average, a plant infected with the A39V mutant developed visible symptoms ~8 dpi compared with the average of ~16 dpi necessary for the WT (Fig. 9A). The dynamics of symptoms development for mutants I32S, G35V, D44E, and D44E/F46L were equivalent to that observed for WT (Fig. 9A; post hoc Bonferroni test  $P > 0.05$ ). The rest of mutants developed symptoms much slower than WT or had not developed symptoms at all 20 dpi when the experiment was concluded (Fig. 9A). A generalized linear model (GLM) was

fitted to the mean time to the appearance of symptoms data. The model used genotypes as a random factor, experimental blocks as replicates, and a Normal probability distribution with an identity link function. Highly significant differences in the time to symptoms appearance among genotypes were observed ( $\chi^2 = 156.597$ , 11 d.f.,  $P < 0.001$ ). A *post hoc* sequential Bonferroni test classified the genotypes into three groups (Fig. 9A).

As a second phenotypic trait, we evaluated the infectivity, that is, the proportion of infected plants 20 dpi. A GLM with genotypes as random factor, experimental blocks as replicates, and a binomial probability distribution with a logit link function found highly significant differences in infectivity among genotypes (Fig. 9B;  $\chi^2 = 151.034$ , 11 d.f.,  $P < 0.001$ ). Genotypes G34D, G34V, and A39E did not induce any symptom 20 dpi and thus we conclude that they were not infectious. Genotypes I32S/G33D and G34R/G35A show a low infectivity (10%) compared with WT. A *post hoc* sequential Bonferroni test classified the rest of genotypes into three categories, with five of the mutants being significantly more infectious than WT, with the A39V variant being the most infectious (85%, almost three times more infectious than WT). Genotypes I32S, D44E, and D44E/F46L had intermediate infectivity values between WT and A39V variant. The infectivity results negatively correlate with the time required to develop symptoms (Pearson's  $r = -0.991$ , 10 d.f.,  $P < 0.001$ ): suggesting that those genotypes inducing symptoms faster were also more infectious (*e.g.*, A39V).



**Fig. 9.** Phenotypic characterization of the different TEV 6K2 mutants. (A) Mean time to symptoms development estimated from the Kaplan-Meier regression. For those genotypes that did not showed symptoms at the end of the experiment, 20 dpi represents the lower bound of the estimated mean time (upper bound being  $+\infty$ ). (B) Mean infectivity 20 dpi ( $n = 10$  plants inoculated). (C) Infectious viral load estimated by means of *C. quinoa* local lesion assay method. In all cases, the dashed vertical line corresponds to the mean phenotypic value of the WT TEV. Mutants are ordered in the ordinate axis to better illustrate the statistically homogeneous groups.

Finally, the third phenotypic trait associated with infection that we characterized was the infectious viral load, measured as the number of LFUs per given amount of infected tissue in a local-lesions assay in fully expanded leaves of *C. quinoa*. LFUs were inferred from the regression of the observed number of local lesions to the dilution factor (Kleczkowski, 1950). For six genotypes (Fig. 9C), we found no local lesions produced in

the *C. quinoa* leaves and thus concluded that the amount of infectious viral particles produced in the infected source tobacco plants was null or below the detection limit of the local-lesion technique. Among those genotypes for which lesions were found, *N. tabacum* plants inoculated with the WT produced, on average, 76.81 LFU/ml, whereas mutant D44E shows the lowest value (11.09 LFU/ml) and mutant I32S the highest one (81.29 LFU/ml) (Fig. 9C). A GLM with genotypes as random factor, experimental blocks as replicates, and a Normal probability distribution with an identity link function revealed highly significant differences in infectious viral load among 6K2 mutant genotypes ( $\chi^2 = 25.665$ , 11 d.f.,  $P = 0.007$ ). A *post hoc* sequential Bonferroni test classified the viable genotypes into two categories, those that accumulate similar amount of infectious viral particles than WT (I32S and D44E/F46L) and those that accumulate significantly less (G35V, D44E, and A39V). Comparing the three panels in Fig. 9, we see that infectious viral load was weakly though significantly correlated with the other two phenotypic traits analyzed. First, it was negatively correlated to the mean time to symptoms development (Pearson's  $r = -0.582$ , 10 d.f.,  $P = 0.047$ ), meaning that the more infectious viral particles accumulated, the faster the symptoms might appear. Second, infectious viral load was positively correlated to infectivity (Pearson's  $r = 0.584$ , 10 d.f.,  $P = 0.046$ ), meaning that more infectious genotypes might also result in a larger accumulation of viral particles.

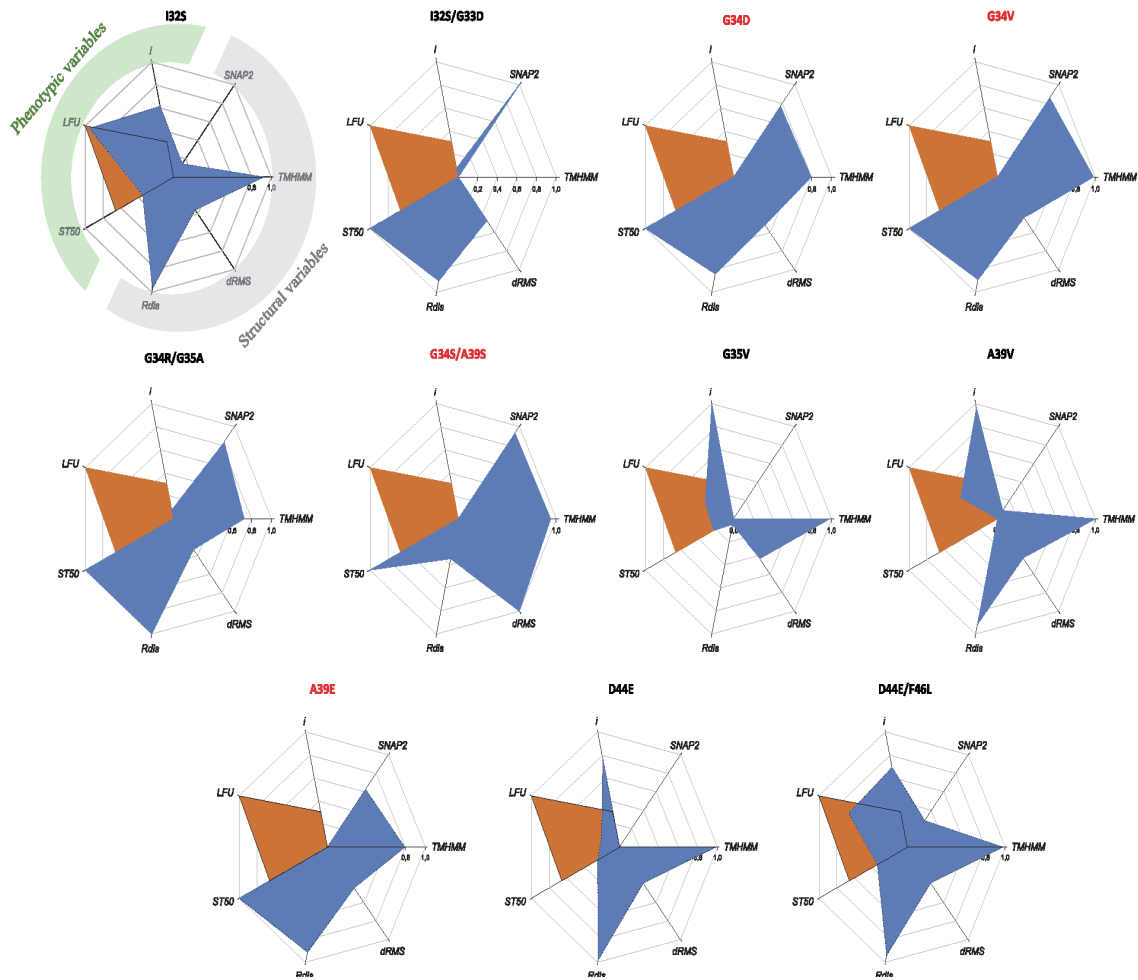
### **I 3.5. Integrating Structural Changes with Phenotypic Effects**

Figure 10 shows two groups of genotypes: those inducing symptomatic infections and those resulting in asymptomatic or no infection. Indeed, using “presence of symptoms” as a binary factor, we can ask whether differences exist between these two classes in terms of SNAP2 scores (in the case of the double mutants, assuming dominance and using as score the maximum value of the pair) and *in vivo* replicative fitness. For the case of

SNAP2 scores, the average score for the genotypes resulting in asymptomatic or no infections was  $86.20 \pm 4.5$  ( $\pm 1$  SEM), whereas it was  $14.0 \pm 11.4$  for those inducing symptomatic infections, the difference highly significant (2- samples t-test with equal variances:  $t_{10} = 5.093$ ,  $P < 0.001$ ), thus confirming that larger functional changes of 6K2 are associated with milder or no infection, whereas small effect or neutral changes are phenotypically closer to the WT infection. For the case of competitive fitness values, the results are equivalent. The average fitness value for genotypes not inducing a symptomatic infection (or more likely being not infectious) was  $1.335 \pm 0.182$ , whereas the value for symptomatic infections was  $0.726 \pm 0.117$ , being the difference also significant ( $t_{10} = 2.959$ ,  $P = 0.014$ ).

Finally, to better integrate the four structural variables and the three phenotypic traits, we generated the radar plots shown in Fig. 10. The four structural variables are placed in the right side of the heptagon (highlighted in gray in the first panel) and the three phenotypic ones in the left side (highlighted in green). In all panels, the combination of WT normalized variables has been indicated by the brown area. Six similar patterns can be readily identified: (1) mutants G34D, G34V, G34R/G35A, and A39E all show the strongest structural perturbations ( $R_{dis}$  and  $dRMS$ ) and functional effects (SNAP2), which correlate with very low or no infectivity ( $i$ ), undetectable viral loads (LFU) and the longest times to symptoms development ( $ST_{50}$ ; if any at the end of the experiments). Indeed, all but one (G34R/G35A) have been classified as nonviable. (2) I32S and D44E/F46L with strong effects on transmembrane properties (TMHMM), large structural distances with WT, yet small functional effects that translate into moderate infectivity, high viral load and short times to symptoms development. (3) A39V and D44E show strong effects on transmembrane properties and structural differences with WT that are

not translated into predicted functional effects yet had no negative effect in infectivity, significantly reduce the time to symptoms development and have a strong negative effect on accumulation. (4) I32S/G33D mutant has no effect on the transmembrane properties but an overall effect on folding similarity with WT that translates into a strong predicted functional effect. This effect is seen in terms of no infectivity, undetectable viral loads and no symptoms development. (5) Mutant G34S/A39S does not differ much from WT folding though it has a stronger transmembrane prediction and large functional effects which result is low infectivity and accumulation and long times to symptoms to be developed. Finally, (6) mutant G35V shows the most distinctive pattern: a strong effect on transmembrane properties which has no predicted functional effect though it increases infectivity, fastens the developments of symptoms and produce moderate accumulations (lower than WT).



**Fig. 10.** Radar plot summarizing all the structural and phenotypic variables estimated for each 6K2 mutant. Trait values have been normalized to a zero to one scale. In all plots, the WT values are included in brown as reference. The three phenotypic traits (infectivity  $i$ , viral load LFU, and mean time to symptoms  $ST_{50}$ ) are placed at the left side of the heptagon (highlighted in green in the first panel). For those genotypes that did not showed symptoms at the end of the experiment,  $ST_{50} \rightarrow +\infty$ ). The four structural traits (SNAP2, transmembrane score TMHMM,  $dRMS$ , and the normalized structural similarity index  $R_{dis}$ ) are placed at the right side of the heptagon (highlighted in gray in the first panel). Nonviable mutants are indicated in red.

## I 4. DISCUSSION

*Potyvirus* 6K2 protein is an essential component for the formation of VRC and hence a key element to successfully complete the infection cycle (Restrepo-Hartwig and Carrington, 1994; Rajamäki and Valkonen, 1999). Consequently, it has attracted a wealth of attention during recent years. These studies have mainly focused on disentangling the mechanisms by which 6K2 forms vesicles by modifying the structure of ER and other endomembrane complexes (Agbeci et al., 2013; Jiang et al. 2015), how these vesicles move along the cytoskeleton microfilaments (Cotton et al., 2009) to neighboring cells (Grangeon et al. 2013) and long distance (Spetz and Valkonen 2004; Wan et al. 2015), and identifying viral and cellular components that are required for 6K2 function (Löhmus et al., 2016; Geng et al., 2017; Movahed et al., 2017). However, no attention has been paid to the evolutionary constraints that should be operating upon this essential small protein. To tackle this question for the first time, we have taken an experimental evolution approach. We began our work with the generation of a library of TEV mutants carrying almost every possible mutation in 6K2. Then, we inoculated the mutant virus' library into a plant host and allowed all variants to compete with each other and with different proportions of the WT. It should be expected that natural selection would fish out those variants that are viable and, in the absence of other selective forces such as random drift or spatial stratification, rank their population abundance according to fitness. Finally, using high-throughput Illumina sequencing, we characterized the composition of the evolved populations and estimated the fitness values of each surviving 6K2 haplotype. By doing so, we have identified 23 different variants some of which were as fit as the WT TEV, some beneficial and some slightly deleterious. Some of the mutations pervasively appearing in different experiments (*e.g.*, I32T, I32S, D44E, and G35V) or different



mutations affecting the same amino acids sites (*e.g.*, I32, G34, G35, A39, and F46). Out of these, and based on their fitness estimates, we selected 11 6K2 mutants for further biological characterization.

As mentioned, some amino acid residues were more tolerant to mutations than others. Interestingly, most of the identified amino acid replacements were located in the transmembrane domain of 6K2 that encompasses residues 20-42 (Schaad et al., 1997). Out of 17 unique amino acid substitutions (Table 2), none was located in the N-terminal domain that contains  $\alpha$ -helix H1, 14 were located in the transmembrane domain that contains  $\alpha$ -helices H2 and H3 and three affected the C-terminal disordered domain. This distribution of amino acid replacements per domain significantly deviates from what should be expected by chance (Fisher's exact test,  $P < 0.001$ ), showing an enrichment in mutations in the trans- membrane domain and a depletion of mutations in the two extra-membrane domains, most significantly in the N-terminal one. This provides evidences for purifying selection upon the N-terminal region being stronger than in the other protein domains. This hypothesis is further backed up by previous observations by Rajamäki and Valkonen (1999) and Spetz and Valkonen (2004), indicating that mutations in the N-terminal domain strongly affect the ability of PVA to move systemically, and by Jiang et al. (2015) that found the N-terminus domain of 6K2 essential for ER export and vesicle formation of TuMV.

It is remarkable that several of the mutations found affected the run of three glycines (residues 33-35) involved in the bonded turn that separates H2 and H3. These mutations induced a conformational change resulting in the fusion of H2 and H3 into a longer  $\alpha$ -helix, which may increase the rigidity of the transmembrane domain. In addition, this

conformational change may contribute to better solubilization of 6K2, allowing entry to the nucleus and reducing to some extent their ability to become integral membrane proteins. Other mutants with large SNAP2 scores (*e.g.*, G34R/G35A and G34V; Fig. 6) by contrast, do not suffer the angle torsion affecting H3 and C-terminus. Furthermore, mutations G33D and G34R are the two predicted to have the largest functional effect by the SNAP2 classifier (Fig. 6), they both replace the small nonpolar radical of glycine with a large charged radical (negative and positive, respectively). They both appeared linked to other mutations I32S/G33D and G34R/G35A that also affect the bonded turn. I32S contributed an additional polar radical and G35A retains a small nonpolar radical. These mutations affecting the rigidity of the transmembrane  $\alpha$ -helix have a negative impact in the dynamics of virus accumulation and symptoms development (Fig. 7), yet surprisingly I32S/G33D and G34R/G35A mutants have some of the largest beneficial fitness effects measured *in vivo* from the change in frequency data (Table 2). Supporting the importance of this glycine-rich motif, Cabanillas et al. (2018) have recently shown for TuMV that mutations in this motif cause 6K2 to accumulate in the Golgi apparatus and plasma membrane. Indeed, glycine by valine mutants accumulate in the apoplastic side of the plasma membrane, in contrast to the WT protein that accumulates in the cytoplasmic side.

Mutations A39V and G35V showed the most virulent phenotypes (more infectious and shorter timing to symptom development), though they accumulated less infectious units than WT (Fig. 9) and had estimated fitness values approximately half of the WT (Table 2). A39V induces a torsion in the molecule that affects the region from H3 to the C-terminus (Fig. 5), although this apparently major structural change is associated with a relatively small functional effect according to SNAP2 (Fig. 6). Likewise, mutation G35V induces a torsion that affects the orientation of the H1 at the N-terminal part of the

molecule and fuses H2 and H3; this structural change was scored as relatively neutral (Fig. 6) by SNAP2.

At a first glimpse, the results obtained for predicted functional effects explain quite well the observed disease phenotypes but are at odds with the estimated fitness values. However, three factors should be taken in consideration to draw a complete picture: (1) SNAP2 estimates refer to potential functional changes in the protein itself out of its biological contexts (*i.e.*, interactions with other viral and host factors). (2) Our estimates of within-host fitness from the Illumina data are based on changes in frequency from the inoculum to the sampling time. We cannot rule out that some of mutations in 6K2 rose in frequency by reasons other than an inherent beneficial fitness effect (*e.g.*, drift or a selective hitchhiking). (3) The fitness estimates were obtained in the context of a complex quasispecies in which the master 6K2 sequence corresponded to the WT one. The effect of these mutations thus is modulated by and depends on the composition of the cloud of mutants. By contrast, the virulence assays were done with different quasispecies compositions resulting from the replication of the corresponding infectious clone and thus the master sequences correspond to the mutant sequence.

#### **I 4.1. Some Considerations about the Bulk Selection Experiments and Fishing Out Potentially Beneficial Alleles**

Two aspects of these experiments are worth discussing. First, the library was not infectious by itself. This is not an unexpected result given the dynamic properties of RNA virus mutant swarms presented in the Introduction. Quasispecies theory predicts the existence of an error threshold beyond which the fitness class distribution disappears into a new state, the error catastrophe, in which the frequency of mutants is not determined by

their fitness (Bull et al., 2005; Domingo et al., 2012). Increasing mutation rate pushes the viral population across the error threshold. A viral population replicating in the error catastrophe range is doomed to extinction in a process dubbed as lethal mutagenesis in the quasispecies literature and as mutational meltdown in the classic population genetics literature (Bull et al., 2007). Here, by generating all possible single-nucleotide substitution mutants (plus some with higher number of mutations) we have artificially created a population that is already into error catastrophe: the diversity composition in the library is far larger than observed for the WT quasispecies. To rescue the library from error catastrophe and establish a symptomatic infection, it was sufficient to mix it with as low as 5% of WT TEV.

Second, we have detected a limited number of potentially viable genotypes (23) (Table 2), five of which have fitness values estimated to be larger than the fitness of WT, whereas others have about the same fitness (*e.g.*, I32T or G34C) and some clearly much lower fitness (*e.g.*, G35V or W436C). Finding beneficial mutations at a frequency high enough as to be detected by the Illumina technique is not surprising. What is more surprising is the finding of largely deleterious 6K2 alleles (*e.g.*, A39E, G34S/A39S, and G34D) at noticeable frequencies. Five possible, nonmutually exclusive explanations, can be brought forward: (1) These deleterious mutations are linked to a beneficial one elsewhere in the genome that is being positively selected and are hitchhiking despite their negative effect. (2) The deleterious effect of these mutations may reverse due to epistatic interactions with mutations elsewhere in the genome that arose and reversed the deleterious effect well before the deleterious mutation was fixed, and the beneficial combination of both mutations then rose together to fixation (Cowperthwaite et al., 2006). (3) A nonhomogeneous spatial distribution of alleles in the plant implies that deleterious

alleles may persist longer in local subpopulations as long as they are not directly competing with better alleles present in their close neighborhood (Aguirre and Manrubia, 2008). (4) If multiplicity of infection (MOI) is high, then cells can be coinfecting by different variants and deleterious alleles may be easily complemented by shared common goods. Regarding this last explanation, experimental measures of MOI for TEV in tobacco rendered values  $< 1.5$  per cell (Tomas et al., 2014), although MOI has been estimated to be slightly higher (5-6 per cell) for soil-borne wheat mosaic potyvirus (Miyashita and Kishino, 2010), in any case, MOI should be low enough to reduce the likelihood of complementation to explain the observed high frequency of some deleterious alleles. (5) Finally, a less interesting possible explanation is that these mutations are artifacts produced during the preparation and sequencing of the Illumina libraries. Given the inherent high error rate per site of this sequencing technique, in the 0.3-1% range. In the worse scenario, assuming a 1% error rate, then all the mutants discussed in this study would be sequencing errors. However, this possibility makes little sense in the light that some of the mutants had appeared in independent experiments. Even accepting that some of the mutants analyzed in this study were the result of technical errors, by characterizing them at the structural and phenotypic level we have provided relevant information on the selective and functional constraints operating on potyvirus' 6K2.

#### **I 4.2. A Consideration on Epistasis and the Ruggedness of Adaptive Landscapes**

Another interesting evolutionary aspect is the observation that the effect of some mutations is contingent to the presence of other mutations in the protein. For instance, mutation I32S in the 6K2-WT background shows completely different structural and phenotypic effects than when it appears in combination with mutation G33D (Fig. 10).

The same situation is true for mutation D44E in WT background and in presence of mutation F46L (Fig. 10). In these two examples, fitness reversals from deleterious to beneficial effects result from epistatic interactions among mutations (Cowperthwaite et al., 2006). Epistasis determines the ruggedness of adaptive fitness landscapes and the accessibility of adaptive pathways (De Visser and Krug, 2014). Epistasis has been shown to be pervasive in the genome of RNA viruses (reviewed in Elena et al., 2010), including TEV (Lalić and Elena, 2012, 2015; Hillung et al., 2015), thus suggesting that TEV fitness landscape should be quite rugged in nature. Indeed, the ruggedness of TEV adaptive landscape has been shown to depend on the host species, being more rugged in a novel host (*A. thaliana*) than in the natural one, *N. tabacum* (Cervera et al., 2016a).

Recent empirical studies exploring the accessibility of adaptive peaks for TEV in *A. thaliana* at increasing mutational distances from the local optima have shown that the chances to return to this local peak decrease with mutational distance from it. At distances longer than one mutational step, viral populations tend to jump on the landscape and reach new distant peaks (Cervera et al., 2016b). When the contribution of adaptation, chance and contingency to evolution were evaluated, it turned out that adaptation, by large, was the most relevant factor, with contingency and chance event playing similar roles (Cervera et al., 2016b).

## **I 5. CONCLUSIONS**

Our results allow us to hypothesize a possible model of 6K2 molecular evolution in which different domains of the protein are subjected to different selective pressures: although the N- terminal domain is evolving under purifying selection, the transmembrane

domains are more evolvable but with most mutations either retaining the  $\alpha$ -helices or fusing them into a longer one, and the disordered C-terminal domain being also able to accommodate mutations. Mutations that are predicted in silico to have a major impact in 6K2 function result in asymptomatic or (most likely) no infections, whereas mutations predicted to have weaker functional effects result in infections which are similar to those induced by the WT virus. The in vivo estimated fitness effects of mutations are negatively associated with the ability to induce symptomatic infections and to the functional effects estimated in silico, which suggests that they strongly depend on the composition of the viral quasispecies. Observations also suggest a model in which a negative tradeoff may exist between within-host replicative fitness and severity of symptoms, with beneficial mutations being those associated with weaker symptoms and slower disease progression. This tantalizing possibility needs future experiments to further explore the evolution of this essential tiny viral protein.





## CHAPTER II

---

### Role of host genetic diversity for susceptibility-to-infection in the evolution of virulence of a plant virus

---

#### II ABSTRACT

Predicting viral emergence is difficult due to the stochastic nature of the underlying processes and the many factors that govern pathogen evolution. Environmental factors affecting the host, the pathogen and the interaction between both are key in emergence. In particular, infectious disease dynamics are affected by spatiotemporal heterogeneity in their environments. A broad knowledge of these factors will allow better estimating where and when viral emergence is more likely to occur. Here, we investigate how the population structure for susceptibility-to-infection genes of the plant *A. thaliana* shapes the evolution of TuMV. For doing so we have evolved TuMV lineages in two radically different host population structures: (1) a metapopulation subdivided into six demes (subpopulations); each one being composed of individuals from only one of six possible *A. thaliana* ecotypes and (2) a well-mixed population constituted by equal number of plants from the same six *A. thaliana* ecotypes. These two populations were evolved for twelve serial passages. At the end of the experimental evolution, we found faster adaptation of TuMV to each ecotype in the metapopulation than in the well-mixed heterogeneous host populations. However, viruses evolved in well-mixed populations were more pathogenic and infectious than viruses evolved in the metapopulation. Furthermore, the viruses evolved in the demes showed stronger signatures of local specialization than viruses evolved in the well-mixed populations. These results illustrate how the genetic diversity of hosts in an experimental ecosystem favors the evolution of virulence of a pathogen.

## **II 1. INTRODUCTION**

Since the term ‘emerging infectious disease’ was coined in the mid-1900s, its definition has evolved (Rosenthal et al., 2015). Woolhouse and Dye (2001) enunciated the most comprehensive definition of an emerging infectious disease as one *‘whose incidence is increasing following its first introduction into a new host population or whose incidence is increasing in an existing host population as a result of long-term changes in its underlying epidemiology’*. Engering et al. (2013) rephrased this definition and suggested that emergence events can be classified into three groups: (1) viruses showing up in a novel host, (2) mutant viral strains displaying novel traits in the same host, and (3) already known viral diseases spreading out in a new geographic area. Following these definitions, viral emergence is usually associated with cross-species transmission but it can actually occur with or without species jump (Di Giallonardo and Holmes, 2015).

The emergence and re-emergence of viruses cause serious threatening to public health (Morens and Fauci, 2013) and are responsible for large yield losses in crops that compromise food security (Vurro et al., 2010). Gaining knowledge on the principles that govern viral emergence would allow to predict when and where such events are more likely to happen. Achieving an accurate prediction of the emergence of viral diseases is a challenging task, because emergence is governed by multiple and diverse factors that remain poorly understood or completely unknown. Predictions will become even more difficult under the ongoing climate change that will favor the conditions for development and dispersal of the virus vectors (Garrett et al., 2006; Baylis, 2017).

The spectrum of disease severity can be attributed to heterogeneity in virus virulence or in host factors; the two are not necessarily independent explanations and they may actually complement and/or interact with each other. A problem faced by viruses is that host populations consist of individuals that had different degrees of susceptibility to infection (Schmid-Hempel and Koella, 1994; Pfennig, 2001). Therefore, adaptive changes improving viral fitness in one host may be selected against, or be neutral, in an alternative one. Genetic variability in susceptibility of hosts and infectiousness of viruses have been well studied in animals and plants (Schmid-Hempel and Koella, 1994; Altizer, 2006; Hughes and Boomsma, 2006; Brown and Tellier, 2011; Anttila et al., 2015; Parrat et al., 2016). Variability within host populations can arise from nonrandom spatial distributions of genotypes: social groups of animals are more closely related to each other than to other members of the population, and plant crops are generally cultivated as genetically homogeneous plots. These situations facilitate virus transmission among genetically similar host genotypes. Spatial structure and local migration predict evolution of less aggressive exploitation in horizontally transmitted parasites (reviewed in Parrat et al., 2016). For example, Boots and Meador (2007) observed that *Plodia interpunctella* granulovirus (PiGV) evolved in spatially structured lepidopteran host populations become less virulent than the one maintained in well-mixed host populations. More recently, Berngruber et al. (2015) showed that a latent  $\lambda$  bacteriophage won competitions against a virulent one in a spatially structured host populations but lost in well-mixed populations. In a natural context, it has been observed that virulent *Linum marginale* fungi were more frequent in highly resistant *Melampsora lini* plant populations whereas avirulent pathogens dominated susceptible populations (Thrall and Burdon, 2003; Thrall et al. 2012). Similar results were observed in laboratory evolution experiments with the pathosystem *A. thaliana*-TEV: a negative association between plant natural accessions

permissiveness to infection and TEV virulence was evolved (Hillung et al., 2014). Finally, host population structure also promotes coexistence of hosts and parasites by creating refugia due to limited dispersal of viruses; increased dispersal makes coexistence less stable (Brockhurst et al., 2006).

The role of host population heterogeneity has also received quite a lot of attention from theoreticians (Comins et al., 1992; Gandon et al., 1996; Boots and Sasaki, 1999, 2002; Gandon and Michalakis, 2002; Tellier and Brown, 2011), resulting in a number of interesting predictions that in many instances have not been properly tested experimentally. One of the most tantalizing predictions is that in the absence of host heterogeneity, parasites must evolve toward a host exploitation strategy that maximizes transmission with low virulence (Haraguchi and Sasaki, 2000; Regoes et al., 2000; Rodríguez and Torres, 2001; Ganusov et al., 2002; Gandon, 2004; Lively, 2010; Moreno-Gómez et al., 2013). However, in the context of emerging diseases, right after the spill-over of the new pathogen into the heterogeneous host population and prior to adaptive evolution to take place, Yates et al. (2006) showed that host heterogeneity has a small effect in the probability of establishing the disease. Very recently, Chabas et al. (2018) showed that evolutionary emergence is more likely to occur when the host population contains intermediate levels of resistant hosts, confirming this prediction using different phages and bacterial hosts with different alterations in the CRISPR/Cas immune systems that conferred the cells with different levels of resistance.

In this study, we use experimental evolution to explore the effect of host population structure for genes involved in susceptibility to infection in the evolution of infectiousness and virulence of a plant virus; in particular, we want to explore the extent

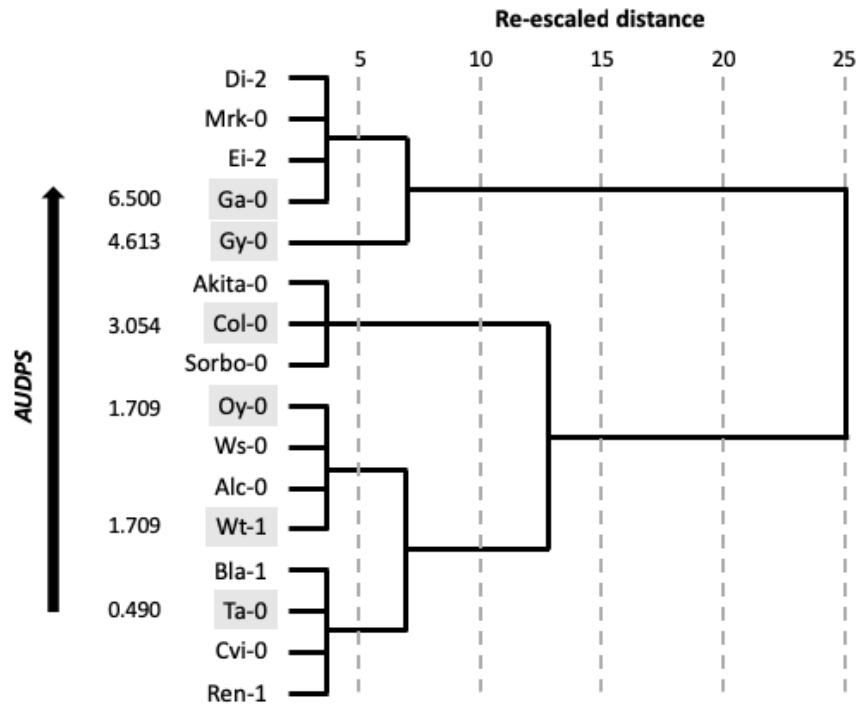
to which host heterogeneity determines the rate of evolution of these two fitness-related traits. According to the above theoretical predictions and experimental observations in other pathosystems, we expect the virus infectiousness and virulence to evolve to lower levels in genetically homogeneous host subpopulations, but at a faster rate, than in genetically diverse well-mixed host populations. The pathosystem we have studied is composed of TuMV as pathogen and six different natural accessions (hereafter ecotypes) of *A. thaliana* that differ in their susceptibility to infection as hosts.

## **II 2. METHODS**

### **II 2.1. Selection of *A. thaliana* Ecotypes for Evolution Experiments**

Before we could begin the evolution experiment, we sought to identify a set of *A. thaliana* ecotypes representative of the phenotypic variability in response to TuMV infection (Rubio et al., 2019) observed for a larger collection of ecotypes. Sixteen *A. thaliana* ecotypes were evaluated. Based on the similarity of their phenotypic responses to infection with TuMV (disease intensity over time measured as the area under the disease progression steps (*AUDPS*) curve; see Section II 2.4 for a definition and explanation) during 14 days post-inoculation (dpi), ecotypes were clustered as shown in Fig. 11 (UPGMA). Six accessions were chosen as representative of the five clusters shown in Fig. 11: Col-0, Ga-0, Gy-0, Oy-0, Ta-0, and Wt-1. A progressive *k*-clustering algorithm confirmed that adding additional clusters did not result in significant improvement in model predictability (five vs six clusters partial-*F* test:  $F_{1,9} = 2.841$ ,  $P = 0.058$ ). Ta-0 showed the least intense disease progression ( $AUDPS = 0.490$ ) while Ga-0 the most

intense one ( $AUDPS = 6.500$ ); Oy-0 and Wt-1 showed similar progression ( $AUDPS = 1.709$ ).



**Fig 11.** UPGMA clustering of *A. thaliana* ecotypes according to their response to TuMV infection. The six ecotypes selected for the evolution experiment are highlighted in gray. As a measure of virulence, the  $AUDPS$  of the ancestral TuMV isolate on each selected ecotype is indicated in the left scale.

Furthermore, the six ecotypes reached growth stage 3.5 in the Boyes' scale (Boyes 2001) at the same time after germination ( $\pm 1$  day), and at that moment they were inoculated. This synchronization ensures that they all were at the same phenological state when inoculated.

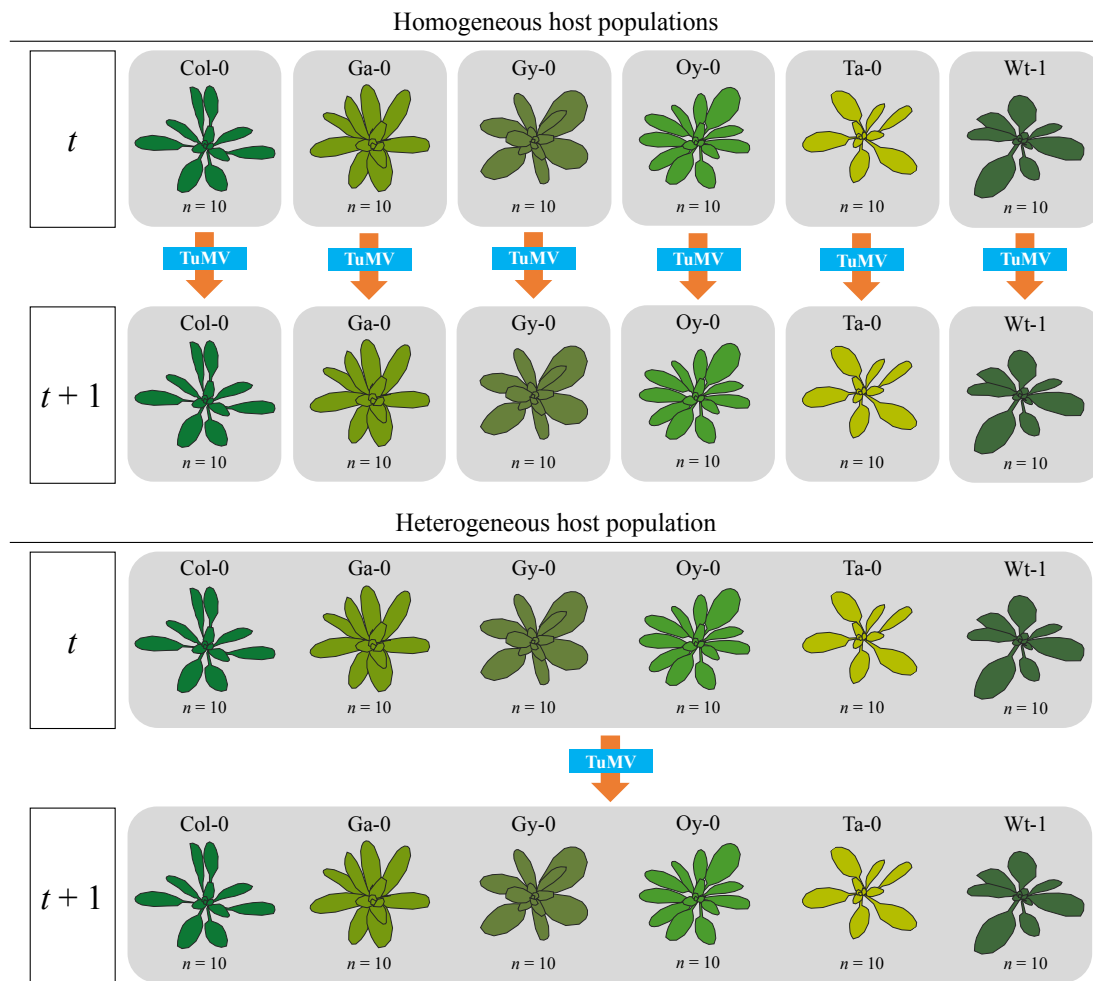
In all experiments performed in this study, plants were maintained in a BSL-2 growing chamber at 8 h light:16 h dark cycles and temperature variation of 24 °C day:20 °C night.

## **II 2.2. TuMV Original Isolate and Inocula**

Infectious saps were obtained from TuMV-infected *N. benthamiana* plants inoculated with an infectious plasmid containing TuMV genome cDNA (GenBank accession AF530055.2) under the control of the cauliflower mosaic virus 35S promoter. This TuMV sequence variant corresponds to YC5 isolate from calla lily (*Zantedeschia* sp.) (Chen et al., 2003). The same stock of plasmid was used to inoculate two batches of five *N. benthamiana* plants each, with a year of difference. After plants showed visible symptoms of infection, two independent infectious saps (viral stocks) were obtained by grinding infected tissues from *N. benthamiana* plants in a mortar with ten volumes of grinding buffer (50 mM KH<sub>2</sub>PO<sub>4</sub> pH 7.0, 3% PEG6000).

## **II 2.3. Experimental Evolution**

TuMV lineages were evolved during twelve consecutive serial passages as schematized in Fig. 12 in two treatments that differ in the composition of the host population. The first treatment consists of a host metapopulation structured in six genetically homogeneous demes. TuMV was evolved in this genetically homogeneous host demes. Passages were made by harvesting the symptomatic plants at 14 dpi, preparing infectious sap as described above and inoculating the virus to the next demes (ten plants each) of the same ecotype. Three leaves from 21-day old plants were rub-inoculated each with 5 µl of infectious sap and 10 per cent Carborundum (100 mg/ml). One TuMV lineage was evolved on each one of the six ecotypes chosen (Fig. 12). In this treatment, each TuMV lineage in the metapopulation only experienced a particular host genotype along their evolution. In this treatment, we expect evolution to be dominated by rapid local adaptation.



**Fig. 12.** Schematic representation of a passage during the evolution experiments. A total of twelve such passages were done, in each of two replications conducted at different dates. See IISection 2.3 for a detailed description.

The second treatment consists of a host population without subpopulation structure composed of ten well-mixed plants from each of the six ecotypes. TuMV was thus evolved in a host population with maximal genetic heterogeneity. In this case, the passages were made harvesting all the symptomatic plants of all the ecotypes, pooling the tissues, preparing sap and transmitting it again to ten plants from each of the ecotypes (Fig. 12). In this treatment, the evolving population of symptomatic TuMV strains has an equal opportunity of infecting each plant ecotype at each passage. More susceptible



ecotypes will contribute more to the next viral generation because more plants are infected and support more generations of replication, while more restrictive ones will contribute in a minor amount to the viral population. In this treatment, we expect a slower adaptation due to fluctuating selection and the evolution of generalist viruses.

This evolution experiment was replicated in two fully independent blocks, hereafter, referred as P and R, respectively. Each block was initiated with a different viral stock (as described in Section II 2.2), different light sources (P with fluorescent tubes at PAR 100-150  $\mu\text{mol}/\text{m}^2/\text{s}$  and R with LED tubes at PAR 90-100  $\mu\text{mol}/\text{m}^2/\text{s}$ ), inoculations were done by two different researchers, and started with more than year of difference in time (P started 12/02/2016 and R started 01/24/2018). In addition, the substrate mix used in passage eight of block P was not the *A. thaliana* standard one in some of the pots but the one used for growing tomato plants, as a consequence *A. thaliana* plants in these pots grew slightly smaller.

#### **II 2.4. Evaluation of infectiousness and virulence**

Upon infection, plants were observed every day and the number of plants showing symptoms was recorded. Infectiousness,  $i$ , was evaluated as the frequency of plants showing symptoms at 14 dpi out of the ten plants inoculated with a standardized amount of infectious sap. Using the frequency time-series within the duration of a passage, the *AUDPS* (Simko and Piepho, 2012) was evaluated as a proxy to virulence. *AUDPS* represents the intensity at which symptoms appear in a population of inoculated plants, and in our case, it is bounded between zero (no plant shows symptoms 14 dpi) and ten (all plants show symptoms at 1 dpi). In the case of the heterogenous host population treatment,  $i$  and *AUDPS* were evaluated in each of the ecotypes.

The  $i$  data were probit-transformed as

$$f = \text{probit}(i) = \sqrt{2} \text{erf}^{-1}(2i - 1) \quad (\text{Eq. 2})$$

hence the new variable  $f$  is now Gaussian distributed with mean zero and variance one. For visualizing infectiousness data in a meaningful scale,  $i$  will be presented in figures while the  $f$  transformed data will be used in the statistical analyses described in II Section 2.5.

At each passage, infection status of each plant was assessed by the presence of symptoms rather than by molecular detection methods. In our extensive experience with the TuMV/*A. thaliana* pathosystem, there is an almost one-to-one match between infection and the development of symptoms. Symptoms started with leaf curling and vein bleaching (~5 to 6 dpi) that quickly developed to diverse grades of leaf chlorosis and/or necrosis (~10 to 12 dpi). Plants also suffered a developmental arrest, with deformed new leaves, abortion of flowering button and abnormal growth of the caulinar apex.

## **II 2.5. Data Analyses**

*AUDP* and  $f$  data were fitted to a fully factorial multivariate analysis of covariance model (MANCOVA), in which experimental block ( $B$ ; P and R), population structure ( $D$ ; homogeneous or heterogeneous), and plant ecotype ( $E$ ) were treated as orthogonal factors and passage ( $t$ ) as a covariable. Main effects and all the interactions between factors and the covariable were incorporated into the model. The full model equation thus reads

$$\begin{aligned} \mathbf{X}_{ijkl} \sim & \mu + t_i + B_j + (B \times t)_{ij} + D_k + (D \times t)_{ik} + E_l + (E \times t)_{il} + (B \times D)_{jk} + (B \times D \times t)_{ijk} + (B \times E)_{jl} + \\ & (B \times E \times t)_{ijl} + (D \times E)_{kl} + (D \times E \times t)_{ikl} + (B \times D \times E)_{jkl} + (B \times D \times E \times t)_{ijkl} + \boldsymbol{\varepsilon}_{ijkl} \end{aligned} \quad (\text{Eq. 3})$$

where  $\mathbf{X}_{ijkl} = (AUDPS_{ijkl}, f_{ijkl})^T$  is the vector of phenotypic traits observed at time  $i$ , experimental block  $j$ , population structure  $k$ , and ecotype  $l$ ,  $\mu$  represents the vector of phenotypic grand mean values and  $\boldsymbol{\varepsilon}_{ijkl}$  stands for the vector of errors. In addition, univariate ANCOVA analyses were performed for  $AUDPS$  and  $f$  using the same model equation.

The magnitude of effects was evaluated using the  $\eta_p^2$  statistic (proportion of total variability in the traits attributable to each factor in the model). Conventionally, values  $\eta_p^2 < 0.05$  are considered as small,  $0.05 \leq \eta_p^2 < 0.15$  as medium and  $\eta_p^2 \geq 0.15$  as large effects.

Unless otherwise mentioned, all statistical analyses described in this work were performed using SPSS version 25 software (IBM, Armonk, NY).

## II 2.6. Evaluation of rates of phenotypic evolution

$AUDPS$  and  $i$  data obtained for each lineage were fitted to the following first-order autoregressive integrated moving-average, ARIMA(1, 0, 0), model:

$$Y_t - \rho_1 Y_{t-1} = Y_0 + \beta^Y t + \varepsilon_t \quad (\text{Eq. 4})$$

where  $Y_k$  represents the variable being analyzed at passage  $k$ ,  $\rho_1$  measures the degree of self-similarity in the time-series data (correlation between values at passages  $t$  and  $t - 1$ ),  $\varepsilon_t$  represents the sampling error at passage  $t$ , and  $\beta^Y$  represents the linear dependency of variable  $Y$  with passage number, that is, the rate of phenotypic evolution.

To explore the effect of factors  $D$  and  $E$  in the rates of *AUDPS* and  $i$  evolution, the following generalized linear model (GLM) was fitted to the  $\beta$  values:

$$\beta_{ijk}^Y \sim \beta^Y + D_i + E_j + (D \times E)_{ij} + \varepsilon_{ijk} \quad (\text{Eq. 5})$$

where superscript  $Y$  again refers to the trait being analyzed,  $\beta^Y$  the grand mean value for the rate of evolution of trait  $Y$  and all other terms are defined in II Section 2.5. A Normal distribution and identity link function were chosen based on the minimal *BIC* value among competing models. Notice that the error term  $\varepsilon_{ijk}$  is obtained from the differences in the estimates from both experimental blocks.

## II 2.7. Infection matrices

To analyze the specificity of adaptation of each evolved TuMV lineage, we performed a full cross-infection experiment in which all the fourteen evolved lineages were inoculated into ten plants of all six ecotypes. In the case of the lineages evolved in the well-mixed host population, the virus isolated at passage 11 was inoculated into ten additional plants of each of the six ecotypes to separate it into sub-samples. Infection matrices were analyzed using tools borrowed from the field of community ecology to explore whether

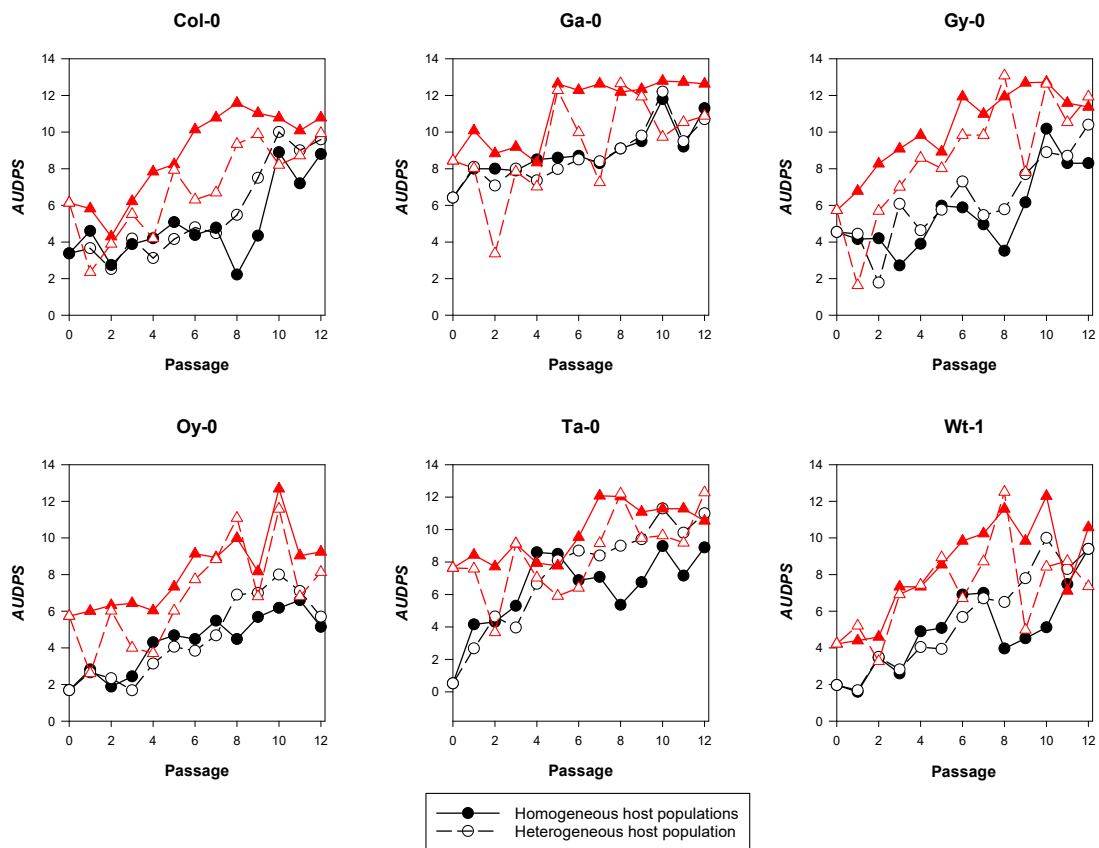
they show random associations between viral lineages and host genotypes, one-to-one associations, nestedness indicative of a gene-for-gene type of interaction, or modularity (Weitz et al. 2013). The statistical properties of the resulting infection matrices were evaluated using the bipartite version 2.11 package (Dormann et al., 2008) in R version 3.3.2 (R Core Team 2016). Three different summary statistics were evaluated: nestedness (Bascompte et al., 2003), modularity (Newman, 2006), and overall specialization  $d'$  index (Blüthgen et al., 2006).  $d'$  is based in Kullback-Leibler relative entropy, that measures variation within networks and quantifies the degree of specialization of elements within the interaction network. Statistical significance of nestedness and modularity was evaluated using Bascompte et al. (2003) null model.

## **II 3. RESULTS**

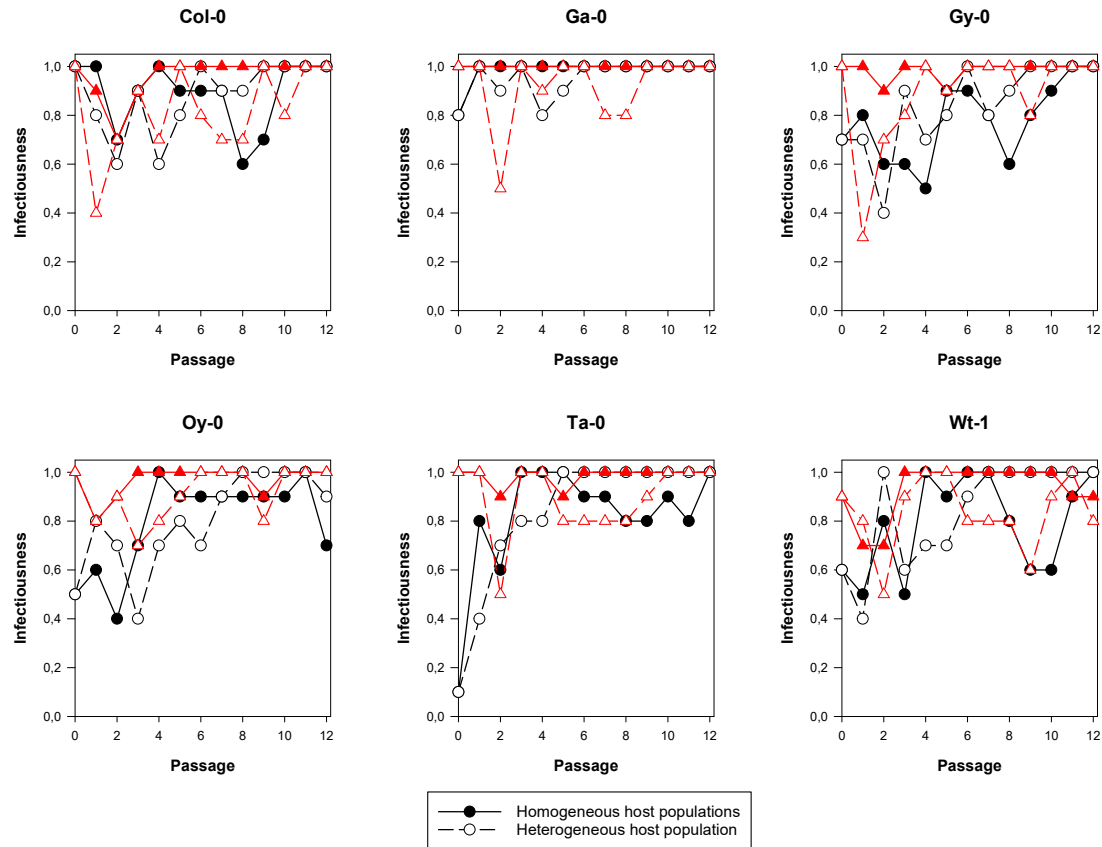
### **II 3.1. Evolutionary Dynamics of Virulence and Infectiousness**

Figure 13 illustrates the evolutionary dynamics of virulence, measured as  $AUDPS$ , observed on each of the host ecotypes and for both population structures. Likewise, Fig. 14 shows the evolution of infectiousness,  $i$ , along the passages. In both figures black symbols and lines correspond to the results of experimental block P while red symbols and lines correspond to experimental block R. Both variables are significantly correlated (Pearson's partial correlation coefficient controlling for  $B$ ,  $D$ ,  $E$ , and  $t$ :  $r_p = 0.640$ , 306 d.f.,  $P < 0.001$ ) and hence multivariate methods were used to analyze the data while increasing the power of the tests. Both datasets were fitted to the MANCOVA model equation described in Section II 2.5. In all cases, an overall trend to increase virulence and infectivity along time can be observed in the time-series data (Figs. 13 and 14). This

trend is statistically supported by a significant net effect of the passage (covariable  $t$ ) on both phenotypic traits (Table 3;  $P < 0.001$ ).



**Fig. 13.** Evolution of virulence ( $AUDPS$ ). Black circles and lines represent the results from experiment P and red triangles and lines from experiment R. Solid symbols represent evolution in the corresponding deme of host metapopulation while open symbols represent evolution in the well-mixed host populations.



**Fig. 14.** Evolution of infectiousness ( $i$ ). Black circles and lines represent the results from experiment P and red triangles and lines from experiment R. Solid symbols represent evolution in demes within a metapopulation while open symbols represent evolution in the well-mixed host populations. Notice that infectiousness data were probit-transformed for statistical analyses.

Next, given the differences in starting inocula, light conditions and experimenter responsible for performing each experimental block, we expect *a priori* to find significant block effects, either net or in combination with other factors in the model. Table 3 shows that experimental block has a net ( $B$ ) effect on the phenotypic vector ( $P < 0.001$ ). This effect changes along evolution passages ( $B \times t$ ) ( $P < 0.001$ ) and depends on the particular ecotype ( $B \times E$ ) ( $P = 0.006$ ). The observed net block effect is of large magnitude ( $\eta_P^2 =$

0.647), while the effect of its interaction with passage number ( $\eta_P^2 = 0.089$ ) and ecotype ( $\eta_P^2 = 0.046$ ) can be considered of small-medium size.

**Table 3.** MANCOVA analysis of the symptoms development (*AUDPS*) and infectiousness (*f*) data (Figs. 13 and 14). See Section II 2.5 for a description of the model equation and parameters.  $1 - \beta$  is the power of the corresponding test.

Source of variation <sup>1</sup>	Wilk's $\Lambda$	<i>F</i>	d.f.	<i>p</i>	$\eta_P^2$	$1 - \beta$
<i>Intersection</i>	0.244	406.426	2, 263	< 0.001	0.756	1
<i>t</i>	0.353	241.176	2, 263	< 0.001	0.647	1
<i>B</i>	0.787	35.600	2, 263	< 0.001	0.213	1
<i>B</i> × <i>t</i>	0.911	12.805	2, 263	< 0.001	0.089	0.997
<i>D</i>	0.943	7.905	2, 263	< 0.001	0.057	0.952
<i>D</i> × <i>t</i>	0.972	3.816	2, 263	0.023	0.028	0.690
<i>E</i>	0.713	9.678	10, 526	< 0.001	0.155	1
<i>E</i> × <i>t</i>	0.947	1.445	10, 526	0.157	0.027	0.731
<i>B</i> × <i>D</i>	0.993	0.969	2, 263	0.381	0.007	0.218
<i>B</i> × <i>D</i> × <i>t</i>	0.985	1.979	2, 263	0.140	0.015	0.407
<i>B</i> × <i>E</i>	0.911	2.524	10, 526	0.006	0.046	0.954
<i>B</i> × <i>E</i> × <i>t</i>	0.941	1.631	10, 526	0.094	0.030	0.794
<i>D</i> × <i>E</i>	0.991	0.225	10, 526	0.994	0.004	0.130
<i>D</i> × <i>E</i> × <i>t</i>	0.990	0.275	10, 526	0.986	0.005	0.152
<i>B</i> × <i>D</i> × <i>E</i>	0.985	0.402	10, 526	0.946	0.008	0.212
<i>B</i> × <i>D</i> × <i>E</i> × <i>t</i>	0.983	0.464	10, 526	0.913	0.016	0.316

<sup>1</sup>Model factors: passage  $\equiv t \in (0, \dots, 12)$ , block  $\equiv B \in (\mathcal{P}, \mathcal{R})$ , population structure  $\equiv D \in$  (homogeneous, heterogeneous), host ecotype  $\equiv E \in (\text{Col-0}, \text{Ga-0}, \text{Gy-0}, \text{Oy-0}, \text{Ta-0}, \text{Wt-1})$ .



Most interestingly, the population structure in which the TuMV lineages had evolved has a highly significant net ( $D$ ) effect on the phenotypic vector (Table 3;  $P < 0.001$ ), though it is of small-medium size in the multivariate analysis ( $\eta_p^2 = 0.057$ ), and of similar effect for both  $AUDPS$  and  $f$  (univariate ANCOVAs shown in Supplementary Table S2;  $\eta_p^2 = 0.042$  and  $\eta_p^2 = 0.051$ , respectively). As it can be seen in Figs. 13 and 14, this effect comes from a pattern (consistent across both experiments) that  $AUDPS$  and  $i$  values are usually larger for the lineages evolving in the well-mixed population during the early passages of evolution than for those evolved in the metapopulation (solid symbols are above open ones). Indeed, on average  $AUDPS$  was 7.42% larger for TuMV lineages evolved in well-mixed population than in metapopulation. Likewise,  $f$  was 12.25% higher for TuMV lineages evolved in the well-mixed plant population than in the genetically homogeneous demes. Furthermore, the effect of population structure changed along passages ( $D \times t$ ) (Table 3;  $P = 0.023$ ), though this effect is of small size in the multivariate analysis ( $\eta_p^2 = 0.023$ ) and also in the univariate ones (Supplementary Table S2;  $\eta_p^2 = 0.016$  and  $\eta_p^2 = 0.028$ , respectively).

The host ecotype in which lineages evolved has a highly significant effect on the magnitude of the phenotypic vector (Table 3;  $P < 0.001$ ), the size of this effect being large ( $\eta_p^2 = 0.155$ ). Indeed, the univariate analyses show that the effect is much larger for  $AUDPS$  than for  $f$  (Supplementary Table S2;  $\eta_p^2 = 0.233$  and  $\eta_p^2 = 0.085$ , respectively). The TuMV evolved in Ga-0 was the most virulent ( $AUDPS = 9.492$ ), followed by viruses evolved in Ta-0 and Gy-0, while the less virulent infection corresponds to a homogenous group formed by TuMV lineages evolved in Col-0, Wt-1, and Oy-0 ( $AUDPS$  in the range: 5.938-6.525) (sequential Bonferroni's *post hoc* test,  $P \leq 0.017$ ). This ranking of ecotypes according to virulence slightly differs from the ranking observed for the ancestral TuMV

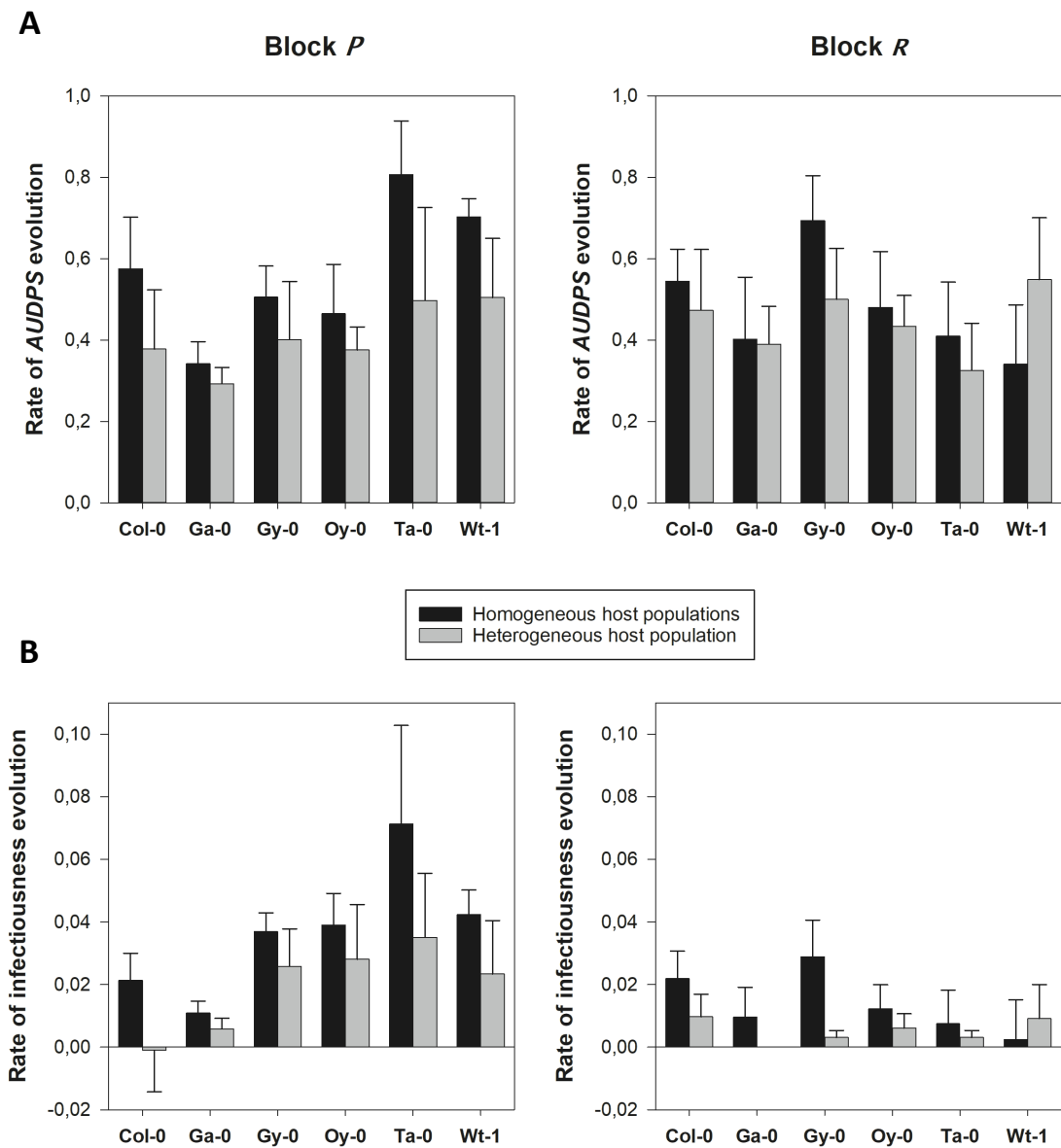
isolate (Fig. 11). Likewise, the virus evolved in Ga-0 was also the most infectious ( $f = 1.264$ ), followed by lineages evolved in Col-0 and Ta-0. The less infectious viruses were those evolved in Wt-1, Oy-0, and Gy-0 ( $f$  in the range: 0.905 - 0.972).

Therefore, we conclude in this first section that the presence of maximal host genetic diversity for genes involved in susceptibility to infection selects for more virulent and infectious viruses. Differences are mostly due to the dynamics of the initial passages. For both phenotypic traits studied, this effect depends on the presence of particular ecotypes.

### **II 3.2. Rates of Phenotypic Evolution**

In the previous section, we found a net effect of passage on the virulence-related traits. More interestingly, this effect depended on the population structure ( $D$ ) and was consistent across both experimental blocks (non-significant  $B \times D \times t$  effect; Table 3), which allows us to treat the estimates of rates of evolution from each block as replicates in a GLM analysis. Next, we sought to get a better understanding of the effect of population structure for susceptibility to infection on the rates of phenotypic evolution. To do so, we have estimated evolution rates as described in Section II 2.6. Summary statistics for the ARIMA(1, 0, 0) model fitting are shown in Supplementary Table S3. Figure 15 compares the estimated rates of evolution for  $AUDPS$  and  $i$  for both population structures analyzed. Rates of evolution were fitted to the GLM described in Section II 2.6. The results from these analyses are shown in Table 4. First, let's focus in the rate of  $AUDPS$  evolution. In all cases except for the lineages isolated from Wt-1, the rate of evolution for lineages evolved in the corresponding deme of the metapopulation is larger than for the corresponding lineages evolved in the well-mixed host population (Fig. 15A).

This trend is statistically significant (Table 4;  $P = 0.015$ ) and the size of effect must be considered as large ( $\eta_p^2 = 0.219$ ).



**Fig 15.** Mean rates of phenotypic evolution for the two traits studied, *AUDPS* (A) and *i* (B). Rates of evolution were estimated from the ARIMA(1, 0, 0) model described in Section II 2.6. Error bars represent  $\pm 1$  SEM.

Second, a similar result has been observed for the rates of I evolution (Fig. 15B): rates of infectiousness evolution are faster in each deme of the metapopulation than in the well-

mixed population. Again, differences among population structures are significant (Table 4;  $P = 0.026$ ) and of large effect  $\eta_p^2 = 0.187$ ).

**Table 4.** GLM analyses of the rates of evolution of *AUDPS* and *i* data (Fig. 15). See II Section 2.6 for a description of the model equation and parameters.  $1 - \beta$  is the power of the corresponding test.

Source of variation <sup>a</sup>	Rate of evolution					
	$(\beta^Y)$	$\chi^2$	d.f.	<i>P</i>	$\eta_p^2$	$1 - \beta$
Intersection ( $\beta^Y$ )	<i>AUDPS</i>	80.518	1	<0.001	0.965	1
	<i>i</i>	25.879	1	<0.001	0.660	0.993
<i>D</i>	<i>AUDPS</i>	5.936	1	0.015	0.219	0.393
	<i>i</i>	4.979	1	0.026	0.187	0.334
<i>E</i>	<i>AUDPS</i>	8.835	5	0.116	0.308	0.259
	<i>i</i>	6.176	5	0.289	0.227	0.179
<i>D</i> × <i>E</i>	<i>AUDPS</i>	3.378	5	0.642	0.131	0.110
	<i>i</i>	1.071	5	0.975	0.044	0.067

<sup>a</sup>Model factors: population structure  $\equiv D \in$  (metapopulation, well-mixed population), host ecotype  $\equiv E \in$  (Col-0, Ga-0, Gy-0, Oy-0, Ta-0, Wt-1).

For both traits, no significant differences in rates of evolution exist between ecotypes (*E*) nor for the interaction between ecotypes and population structure (*D*×*E*) (Table 4),

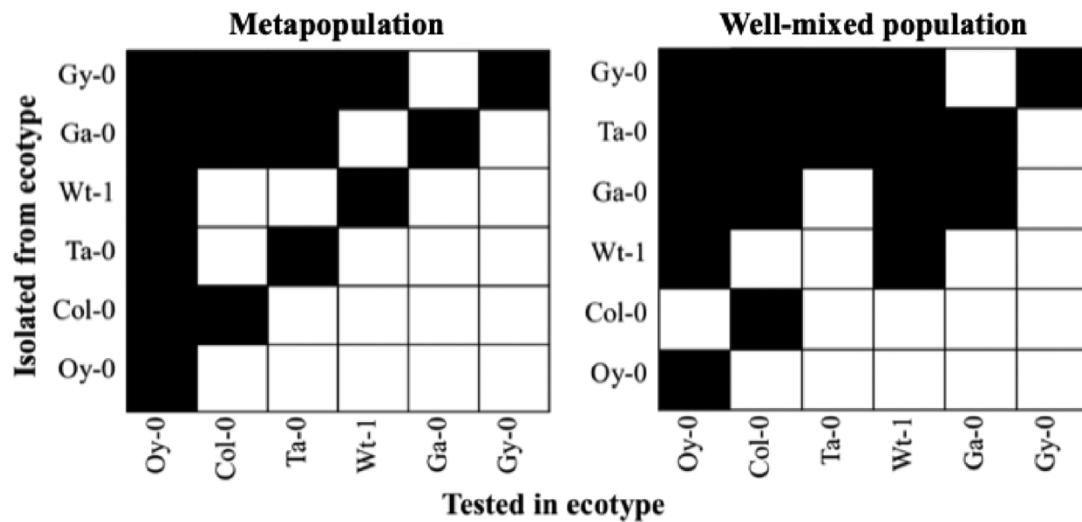
with the power of the tests being too low ( $1 - \beta \leq 0.259$ ) as to fully discard we are not accepting a false null hypothesis of no effects. Low statistical power must clearly be due to the small sample size used (only two experimental blocks).

As a conclusion for this second section, evolution of virulence and infectiousness took place at a faster pace in genetically homogeneous demes of a metapopulations than in a well-mixed genetically heterogeneous population.

### **II 3.3. Comparison of Infection Matrices**

Finally, we evaluated the degree of specificity of adaptation to the six ecotypes for all the evolved lineages. To build up infection matrices, TuMV lineages evolved in each of the ecotypes, or isolated from each ecotype in the case of the two blocks evolved in the well-mixed host population, were inoculated into each one of the six ecotypes and *AUDPS* was estimated as described above. Figure 16 shows a binary representation of the infection matrices estimated (averaging across experiments) for the two host population structures. Black squares represent host-virus combinations in which virulence was equal or greater than for the value observed for the viral lineage in its corresponding local host (host of isolation in the case of the well-mixed host population). White squares represent less virulent combinations. The first row in the matrices correspond to the most generalist TuMV lineage, in this case those evolved in Gy-0, whereas the last row represents the most specialist TuMV lineage, here those evolved in Oy-0. This observation is consistent for both matrices. Likewise, both matrices are also consistent regarding the rank order in susceptibility to infection of the different ecotypes: Oy-0 is the most susceptible ecotype, being infected by almost all viral lineages with similar virulence (exception being the virus isolated from Col-0 in the well-mixed host population) and Gy-0 being the most

resistant ecotype, infected with high virulence only by the lineages evolved in or isolated from Gy-0. In other words, this suggests that more permissive ecotypes select for less virulent viruses while more restrictive ecotypes select for more virulent ones.



**Fig 16.** Infection matrices obtained from the virulence data (mean *AUDPS* values of the two experiments). Black cells represent cases in which virulence was equal or greater than the value estimated for the corresponding viral lineage on its local host ecotype. In the case of the well-mixed population, it corresponds to the value estimated on the ecotype from which the virus was isolated in the last evolution passage. As described in Section II 3.2, in each matrix evolved viruses are ordered from the most virulent (upper row) to the less virulent (bottom row) and host ecotypes from the most sensitive (most left column) to the most resistant (rightest column).

To further test this hypothesis, we evaluated the nestedness and modularity of the two matrices. The infection matrix estimated for the demes of the metapopulations shows a significant nestedness (Fig. 16, left;  $P = 0.035$ ). This suggests that virus evolution in a single host genotype selects for a gene-for-gene interaction mechanism. This model of host-virus interaction is fully compatible with the above hypothesis. However, the matrix

estimated for the well-mixed population did not show significant nestedness (Fig. 16, right;  $P = 0.149$ ), suggesting that gene-for-gene interactions have not been selected under this ecological situation. Indeed, odds ratios indicate that the left matrix in Fig. 16 is 3.16 per cent more nested than the right one.

Both matrices show significant modularity ( $P = 0.022$  for the metapopulation and  $P = 0.012$  for the well-mixed population), though the modularity in the well-mixed host population matrix is slightly larger (odds ratio: 0.24%). A module is a group of viruses and hosts for which the viruses in the set are more likely to be virulent in these hosts than to any other host outside the group, and that hosts in the group are more likely to be infected with similar virulence by viruses from within the group. Such modules suggest common selective constraints imposed by the hosts and similar evolutionary solutions found by the viruses.

Finally, we sought to quantify the degree of specialization in the host ecotype—viral lineage, or partner diversity. To this end, we applied Blüthgen et al. (2006)  $d'$  index. For the infection matrix estimated for viruses evolved in the metapopulation (Fig. 16 left),  $d' = 0.00207$ , while for the matrix estimated for viruses evolved in the well-mixed population (Fig. 16 right),  $d' = 0.00061$ . Therefore, 3.39 times greater specialization evolved when the virus was facing a single host genotype during the evolution experiment.

## **II 4. DISCUSSION**

### **II 4.1 Host Population Structure and the Evolution of Specialist and Generalist Viral Strategies**

Most plant viruses are generalists capable of infecting more than one host species and thus generalism versus specialism should be properly defined on the basis of variance in infectiousness across different hosts (Leggett et al., 2013). In this sense, a generalist virus will be characterized by a low variance in infectiousness over different host genotype and/or species; in contrast a specialist virus will have a high variance. It is logical to expect that virus evolution should proceed faster in homogenous than in heterogeneous host environments because viruses with a narrower host range (specialist) have higher probabilities of fixing beneficial alleles, taking less time to do so (Gavrilets and Gibson, 2002; Whitlock, 2002, 2003; Whitlock and Gomulkiewicz, 2005; Papaïx et al. 2013). Consequently, viral species or genotypes with broader host ranges (generalists) must have slower rates of evolutionary response (Bennett et al., 1992; Fry, 1996; Whitlock, 1996; Kassen and Bell, 1998; Kassen, 2002). If evolution occurs in too many hosts, then the total selection for any particular host-specific trait would not be as effective, and viruses specializing into a single host would then evolve faster and outcompete their generalist counterparts. As a result of this faster evolutionary rate, specialist viruses may persist longer in time in a constant host landscape. In this study, we have directly challenged this hypothesis using experimental evolution. Lineages of TuMV, a prototypical RNA virus of the picorna-like superfamily *sensu* Koonin et al. (2008), were evolved either in six alternative single host demes constituting a metapopulation or in a well-mixed genetically heterogeneous host environment composed by equal numbers of the same six host genotypes. Giving support to the above hypothesis, we observed that rates of phenotypic



evolution were significantly faster for lineages evolved in the individual demes of the metapopulation than in the well-mixed population.

Pfennig (2001) made the distinction between polymorphism and polyphenism as causes of pathogens generalism. On one hand, polymorphism means that different strains of pathogens evolve specialized virulence strategies in different host genotypes. On the other hand, polyphenism means that pathogens facultatively express alternative virulence strategies depending on the host phenotypes. In case of viruses with compacted genomes and multifunctional proteins, polymorphisms seem a more plausible explanation, though we cannot rule out polyphenism. In this sense, highly polymorphic viral populations would behave as generalists owed to the diversity of specialists they contain.

## **II 4.2 Host Population Structure, Selection of Recognition Mechanisms and Evolution of Virulence**

We have also characterized the evolution of two-virulence-related traits, *AUDPS* and infectiousness. Before engaging ourselves in further discussion, it is worth mentioning that our definition of virulence, *AUDPS*, is to some extent different from the classical one used by evolutionary ecologists, namely the reduction in host fitness due to infection, which is the one generally used in most theoretical treatments of the evolution of virulence in infectious diseases. However, our definition is more akin with definitions used by plant pathologists: development of symptoms (*AUDPS*) or susceptibility to infection (infectiousness). Given that TuMV is a castrating pathogen, and hence infection has a strong impact in the plant fitness, *AUDPS* turns out to be a good predictor of the reduction in plant fitness.

According to theoretical predictions (Haraguchi and Sasaki, 2000; Regoes et al., 2000; Gandon, 2004; Lively, 2010; Moreno-Gómez et al., 2013), one should expect to observe viruses to evolve toward mild infections in genetically homogeneous host populations, as a more efficient exploitation strategy that maximizes transmission. In contrast, host heterogeneity is postulated to select for increased virulence (Ganusov et al., 2002). Indeed, in the absence of trade-offs between the virulence in the different host ecotypes, it is expected that virulence increases without bounds (Regoes et al., 2000). Our results also provide support to this hypothesis. TuMV lineages evolved in demes from a metapopulation are, on average, less virulent on their local hosts than those lineages evolved in the well-mixed host population. However, a note of caution must be added here: the aforementioned models assume a trade-off between virulence and transmission rates because more virulent viruses will reduce their host's lifespan or production of viable progeny. In our experiments, this trade-off has been broken since transmission rate is controlled by us and, therefore, virulence can increase without paying a cost. Therefore, our results are consistent with the expectation that mild infections are selected in genetically homogeneous host populations, though the precise mechanism does not rely in the transmission-virulence trade-off and needs to be explored in future work.

Genetic variability in host susceptibility to infection and the evolution of infectiousness and virulence of parasites has been well documented in animals and plants [see Schmid-Hempel and Koella (1994) and Parrat et al. (2016) for reviews]. Resistance is not always achieved against all parasite genotypes but results from the specific interaction between the genotypes of the host and the pathogen in agreement with the so-called gene-for-gene relationship. This mechanism proposes that for each locus determining resistance in the plant host, there is a corresponding locus for virulence in

the virus. This model predicts that infection matrices must show significant nestedness: the most susceptible host carrying few resistance genes will be successfully infected by every virus, while the host genotypes carrying large number of resistance loci will be infected by very few viral genotypes carrying more virulence loci (Weitz et al., 2013). To test whether the gene-for-gene relationship has evolved in our experiment, we evaluated the structure of the infection matrices. In agreement with the gene-for-gene model, we found that the infection matrix for the TuMV lineages evolved in homogeneous host demes was significantly nested, with viruses evolved in the most restrictive host Gy-0 deme being also the most generalist ones able of infecting all alternative ecotypes equally well. In contrast, viruses evolved in the most permissive host Oy-0 deme were the most specialized ones, infecting alternative hosts with less efficiency. Interestingly, the lineages evolved in the well-mixed host population did not generate a nested matrix, suggesting that fluctuating selection avoided the fixation of beneficial alleles in all required virulence loci. It is also noteworthy mentioning that, regardless whether or not nested, both infection matrices were significantly modular, suggesting that the mechanism underlying TuMV-*A. thaliana* interaction is far more complex than expected under a simple gene-for-gene model. Two possible mechanisms will contribute to create modularity: first, plant ecotypes sharing alleles in given sets of defense-response genes will likely impose similar selective constraints to the virus. Second, negative pleiotropic effect of mutations in small and compacted RNA genomes limits the number of alternative evolutionary pathways (Belshaw et al., 2008).

## **II 4.3 Are Our Findings in Agreement with Those From Other Previous Evolution Experiments?**

We would like to compare our findings with those from studies with small RNA and large DNA viruses. First, Cuevas et al. (2003) simulated in vitro the migration of vesicular stomatitis virus (VSV) between three alternative cell types, and observed that migration rate had a negative impact on the rate of fitness improvement on each cell type. With no migration, viruses evolved as specialists with high fitness in their local host, whilst at 50 per cent migration rate per generation VSV evolved as a lower fitness generalist metapopulation. This result is apparently at odds with our findings and with theoretical expectations. However, it is worth noticing that Cuevas et al. (2003) did not measure virulence but fitness in competition against a common reference strain. Since fitness and virulence are not perfectly correlated traits in cell cultures for VSV (Furió et al., 2012), this may account for the discrepancy.

Second, Hillung et al. (2014) evolved TEV in five different ecotypes of *A. thaliana* in a set up similar to our deme host metapopulations treatment. In full agreement with our results, the infection matrix of the evolved TEV genotypes on the five ecotypes was significantly nested and thus compatible with a gene-for-gene relationship.

Third, Boots and Meador (2007) evolved lineage of the large DNA betabaculovirus PiGV in lines of its host lepidopteran at varying viscosity of the food media in which larvae grew. Increasing viscosity decreased the mobility of the larvae and thus created stronger spatial structure. Viruses evolved in spatially structured host populations evolved to be less virulent than those maintained in well-mixed host populations, matching our findings.

Fourth, Kubinak et al. (2014) and Cornwall et al. (2018) observed that serial passages of the mouse-specific Friend virus (FV) complex in mouse populations that differed in diversity in the major histocompatibility complex (MHC) resulted in different fitness and virulence. Both traits rapidly increased when FV was evolved only in inbred mouse strains of the same genotype. However, both viral fitness and virulence were significantly reduced when FV was evolved in mouse populations containing different MHC genotypes.

Finally, Berngruber et al. (2015) showed that a latent  $\lambda$  phage won competitions against a virulent one in a spatially structured bacterial host populations but lost in well-mixed ones. This result agrees with our observation that evolution in isolated homogeneous host populations selected for less virulent TuMV while evolution in well-mixed heterogeneous population selected for more virulent virus.

#### **II 4.4 Are Our Results Compatible with Field Observations?**

There is a number of field studies performed with plant viruses whose results are fully compatible with those from our experimental approach. We are just mentioning a few. Malpica et al. (2006) analyzed the prevalence of five plant viruses on twenty-one wild plant species and found that specialization was the most successful viral strategy, with highly selective viruses (specialist) being the most prevalent ones. In a recent follow up study, McLeish et al. (2017) have applied ecological networks theory to explore the association between plant host diversity and virus spatial distributions. Two important observations were made: (1) an association between host abundance and viral prevalence and (2) most prevalent viruses behave as facultative generalists, meaning that they have

the widest host range breadth but can narrow it to the most permissive host without paying a fitness cost.

Rodelo-Urrego et al. (2013) explored the association between the population structure of *Capsicum annuum glabriusculum* and the prevalence of pepper golden mosaic and pepper huasteco yellow vein begomoviruses, founding that landscape heterogeneity affected the epidemiology and genetic structure of the begomoviruses. When host population diversity was removed by human intervention, the prevalence of the viruses increased. Higher levels of anthropization and loss of plant biodiversity (i.e. similar to our genetically homogenous host populations) resulted in an increase in the genetic diversity of the begomoviruses (Rodelo-Urrego et al., 2015).

## **II 5. CONCLUDING REMARKS**

Understanding of ecological factors, such as the spatial distribution of host genotypes, is critical for deciphering and predicting the evolutionary and epidemiological dynamics of infectious diseases. The role of spatial host structure has attracted quite a lot of attention from theoreticians that made a number of interesting testable predictions. Unfortunately, many of these still remain untested. Here, we have tried to provide experimental evidences for a plant-virus pathosystem to test the hypothesis that host population homogeneity would promote fast local adaptation and low virulence of the virus, whereas the presence of host genetic heterogeneity for susceptibility to infection will slow down the rate of viral evolution and favor more virulent viruses. Supporting these predictions, we found faster TuMV adaptation to homogeneous than to heterogeneous *A. thaliana*

experimental populations. However, viruses evolved in heterogeneous host populations were more pathogenic and infectious than viruses evolved in the homogeneous population. Furthermore, the viruses evolved in homogeneous populations showed stronger signatures of local specialization than viruses evolved in heterogeneous populations. These results illustrate how the genetic diversity of hosts in an experimental ecosystem favors the evolution of virulence of a pathogen and may help agronomists to handle crops in ways that will minimize the rise and spread of virulent viral strains.

## CHAPTER III

---

### From foes to friends: Viral infections expand the limits of host phenotypic plasticity

---

#### III ABSTRACT

Phenotypic plasticity enables organisms to survive in the face of unpredictable environmental stress. Intimately related to the notion of phenotypic plasticity is the concept of the reaction norm that places phenotypic plasticity in the context of a genotype-specific response to environmental gradients. Whether reaction norms themselves evolve and which factors might affect their shape has been the object of intense debates among evolutionary biologists along the years. Since their discovery, viruses have been considered as pathogens. However, new viromic techniques and a shift in conceptual paradigms are showing that viruses are mostly non-pathogenic ubiquitous entities. Recent studies have shown how viral infections can even be beneficial for their hosts. This may happen especially in the context of stressed hosts, where the virus infection can induce beneficial changes in the host's physiological homeostasis, hence changing the shape of the reaction norm. Despite the fact that underlying physiological mechanisms and evolutionary dynamics are still not well understood, such beneficial interactions are being discovered in a growing number of plant-virus systems. Here, we aim to review these diverse studies and place them into the context of phenotypic plasticity and the evolution of reaction norms. This is an emerging field that is posing many questions that still need to be properly answered. The answers would clearly interest virologists, plant pathologists and evolutionary biologists and likely they will suggest possible future biotechnological applications, including the development of crops with higher survival rates and yield under adverse environmental situations.



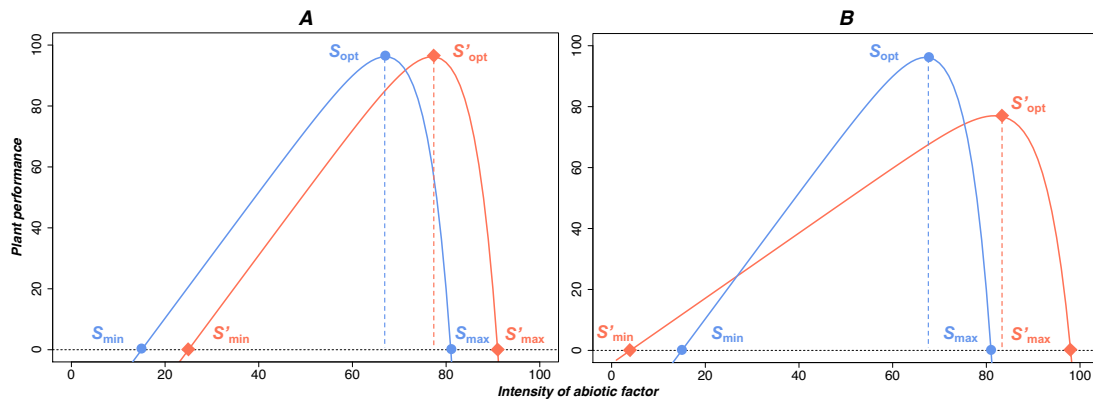
### **III 1. INTRODUCTION**

#### **III 1.1 Phenotypic Plasticity, Reaction Norms and Niche Breadth**

Plant phenotypes result from the highly coordinated expression of their genes under a set of environmental conditions. Understanding the interaction between these environmental factors and the genotype in determining the phenotype, *i.e.*, the so-called genotype-to-phenotype developmental map, has been the focus of intense research almost since the birth of Genetics as a scientific discipline (Ahnert, 2017; Lewontin, 1974). Despite huge efforts, we are still far from having a real understanding of these maps and of the factors that may influence their properties. In nature, plants are facing the challenge of maximizing fitness under sometimes unpredictably fluctuating and heterogeneous environmental conditions. Natural selection has solved this problem in several different ways, including environmental robustness (or canalization) and phenotypic plasticity. Environmental robustness involves a series of genetically-encoded mechanisms that reduce environmental influences on the expression of phenotypic traits (*e.g.*, de Visser et al., 2003). By contrast, phenotypic plasticity is the ability of one genotype to produce more than one phenotype when exposed to different environments (Bradshaw, 1965). Phenotypic plasticity and robustness are thus opposing mechanisms: a highly robust developmental program would result in a single phenotype and, henceforth, will have no phenotypic plasticity. Adaptive phenotypic plasticity is presumed to evolve in response to contrasting selection pressures that arise when organisms confront environmental heterogeneity (Bradshaw, 1965).

The notion of phenotypic plasticity is interwoven with the concept of the reaction norm. A reaction norm describes the pattern of phenotypic expression of a single genotype across a range of environments, thus reaction norms place phenotypic plasticity in the context of genotype-specific responses to varying environments. In other words, reaction norms describe how plastic a phenotypic trait is (*e.g.*, fitness, viability or fecundity) to changes in environmental components (Fig. 17). Reaction norms are used in evolutionary ecology to investigate environmental effects on a variety of phenotypes. A reaction norm can be defined by three parameters (Fig. 17): its minimum ( $S_{min}$ ), optimum ( $S_{opt}$ ) and maximum ( $S_{max}$ ) values.  $S_{min}$  and  $S_{max}$  represent the values of the environmental factor below and above which the organism is no longer viable, respectively. The difference between these two extreme values is known as the niche breadth.  $S_{opt}$  corresponds to the value of the environmental factor at which the performance of the genotype is optimal.

The fact that no perfectly plastic organism exists suggests that constraints to the evolution of phenotypic plasticity must exist (Auld et al., 2010; DeWitt et al., 1998; Murren et al., 2015). DeWitt et al. (1998) suggested two types of restrictions: (*i*) costs that will lead to reduced fitness when a trait is produced via plasticity rather than constitutively and (*ii*) limits due to the inability to produce the optimal trait value. Obvious costs would be associated with the need of carrying extra genetic machinery to detect environmental perturbations and respond by adjusting the phenotype (*e.g.*, physiological homeostasis), or the lack of genetic variation for phenotypic plasticity. Limits may arise, for instance, due to a lag-time between the detection of the environmental cue and the phenotypic response that can reduce fitness.



**Fig 17.** Blue lines represent a hypothetical reaction norm for a given abiotic factor (e.g., temperature).  $S_{min}$  and  $S_{max}$  represent the lower and upper critical values of the abiotic factor beyond which the genotype cannot survive.  $S_{opt}$  represent the value of the abiotic factor at which performance of the plant is optimal and  $S_{max} - S_{min}$  represents the niche breadth. Panels (A) and (B) represent two, out of many, possible effects of viral infection on the reaction norm (red lines). **(A)** In this example, infection switches the reaction norm towards higher values of the abiotic factor.  $S'_{min}$ ,  $S'_{max}$  and  $S'_{opt}$  are now all higher although the niche breadth remains the same  $S_{max} - S_{min} = S'_{max} - S'_{min}$ . **(B)** In this example infection affects the shape of the reaction norm,  $S'_{opt} > S_{opt}$  is also switched towards a higher value but  $S'_{min}$  and  $S'_{max}$  are switch to even more extreme values, thus widening the niche breadth  $S_{max} - S_{min} < S'_{max} - S'_{min}$ .

Host genotypes suffer more or less from infections depending on the environmental settings (Vale et al., 2008a; Wolinska and King, 2009). Indeed, evidence has been reported for bacteria (Knies et al., 2006; Listmann et al., 2016), plants (Gorovits et al., 2019; Hily et al., 2016; Laine, 2007) and animals (Lazzaro and Little, 2009; Vale et al., 2008b) supporting the idea that the outcome of infection strongly depends on environmental factors such as nutrients or temperature. Fig. 17 illustrates possible effects of viral infection in the reaction norm of a given host genotype. In the example illustrated

by Fig. 17A, infection produces a shift in the entire reaction norm, moving the triplet ( $S_{min}$ ,  $S_{opt}$ ,  $S_{max}$ ) to a new set of higher values ( $S'_{min}$ ,  $S'_{opt}$ ,  $S'_{max}$ ) of the environmental abiotic factor under consideration. However, the niche breadth remains the same. Fig. 17B illustrates a somehow different situation in which not only the triplet ( $S_{min}$ ,  $S_{opt}$ ,  $S_{max}$ ) changes to ( $S'_{min}$ ,  $S'_{opt}$ ,  $S'_{max}$ ), but also the shape of the reaction norm. In this particular example,  $S_{min}$  and  $S_{max}$  move to more extreme values, widening the niche breadth, and  $S_{opt}$  moves to a higher value. These are just two examples of many possible effects of infection in the reaction norm of a host genotype, but they have been chosen to illustrate a tantalizing hypothesis: viral infection may change the reaction norm of host genotypes making them more resistant to environmental stresses.

This review is organized as follow. First, we will review mounting evidence that plant viruses are not always pathogenic but, especially in natural situations, peacefully coexist with their natural reservoir hosts. Then, keeping in mind readers not familiar with plants, we will provide a short overview of the mechanisms that plants have to deal with abiotic environmental stresses that will provide a better understanding of the rest of the article. Next, to provide support to the above hypothesis, we review experimental infection data gathered along decades for different plant-virus pathosystems under permissive and stressful conditions (Table 5). We will continue discussing some possible mechanisms by which viruses may increase plant tolerance to abiotic environmental stresses. Finally, we will draw some general conclusions, putting the emphasis into an evolutionary/ecological perspective.

**Table 5.** Summary of the documented plant-virus interactions that lead to changes in the reaction norm for tolerance to different abiotic stresses.

Stress	Host	Virus	Reference	
Water	Drought	Rice	BMV	Xu et al. (2008)
		Beet	CMV	
		Pepper		
		Goosefoot		
		Watermelon		
		Cucumber		
		Squash		
		Tomato		
		Wild tomato		
		Tobacco		
		<i>N. benthamiana</i>		
		<i>N. benthamiana</i>	BMV	
			TMV	
			TRV	
		Tobacco	TMV	
<i>A. thaliana</i>	CMV	Westwood et al. (2013)		
<i>A. thaliana</i>	PPV	Aguilar et al. (2017)		
	PVX			
	PPV+PVX			
<i>N. benthamiana</i>	PPV			
	PVX			

		PPV+PVX		
		<i>N. benthamiana</i>	YTMMV	Dastogeer et al. (2018)
Flood	Wheat		BYDV	Andrews and Paliwal (1986)
	Barley			
Temperature	High	Green bean	TMV	Yarwood (1957)
			TRSV	
			TSWV	
			AMV	
		<i>Curvularia</i>	CThTV	Márquez et al. (2007)
		<i>protuberata</i> infecting		
		panic grass		
		Tomato	TYLCV	Gorovits et al. (2019)
Low	Wheat		BYDV	Fitzgera and Stone (1967)
	Wheat		BYDV	Andrews and Paliwal (1983, 1986)
	Beet		CMV	Xu et al. (2008)
	<i>A. thaliana</i>		TRV	Fernández-Calvino et al. (2014)
Excess of light	<i>Chlorella</i> sp.		CVM-1	Seaton et al. (1996)
radiation	<i>Synechococcus</i> sp.		S-PM2	Mann (2003)
Oxidative	<i>N. benthamiana</i>		PVX	Shabala et al. (2011a, 2011b)

### III 1.2 Viruses as Non-Pathogenic Symbionts

Viruses are the most abundant biological entities on Earth, exceeding cells by a factor of 10 to 100 (Suttle, 2007). They are also a major force in delineating landscapes (*e.g.*,

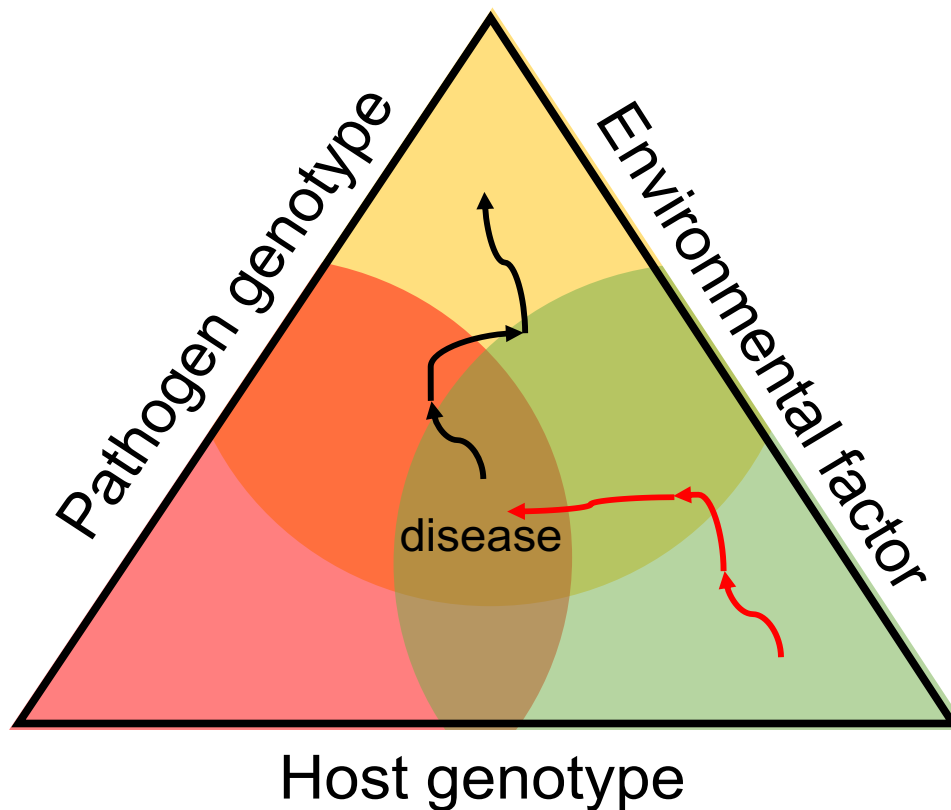
deposition of skeletons of dead foraminifera and diatoms) and providing organic matter for heterotrophic organisms, thus determining the biota's composition (Koonin, 2016; Koonin and Dolja, 2013). Viruses are usually seen as selfish genetic elements highly dependent on their hosts, other selfish elements and the environment (Iranzo et al., 2016). However, with the rise of new sequencing technologies, we are seeing that most plants in nature remain symptomless while infected with one or more viruses (Prendeville et al., 2012; Roossinck, 2012; Roossinck et al., 2010). These host-virus asymptomatic interactions fall on a spectrum between antagonism (pathogenesis) and mutualism as viruses tend to switch between these lifestyle choices (Pagán et al., 2010; Roossinck, 2010, 2011, 2015a, 2015b; Roossinck and Bazán, 2017). Even though most of the current research is still, and logically, driven by disease phenotypes, we are starting to realize the importance viruses play in evolution of Life in general. Sometimes they can promote host survival and transmission in situations of abiotic stress (Gorovits et al., 2019; Hily et al., 2016), which can also be viewed as a way of promoting their own survival (Roossinck, 2010, 2011, 2015a, 2015b). These observations put viruses into a different evolutionary perspective: mutualistic relationships in which the survival of the virus and of the host are intimately associated. The longer the host is alive and more or less able to support viral replication, the more benefit for the virus.

Plant disease has been classically described using the so-called disease triangle (Agrios, 2005) (Fig. 18). Each dimension in the triangle represents the three factors which drive the outcome of the interaction: environmental conditions, virus genotype and plant genotype. As show in the Fig. 18, under the specific combination of these three factors, the outcome of the interaction is the development of disease. Outside the disease area, the interaction between plant and virus may not result in disease but in other symbiotic

relationships. Evolution both of viruses and plant genomes can also change the relationship. For instance, an antagonistic interaction in a given environment can evolve towards a mutualistic one simply by changing the environmental factors (exemplified by the black arrows in Fig. 18), resulting in conditional mutualism (Hily et al., 2016). Likewise, genetic changes in either the host or the virus (or both) may change the relationship from commensalism to parasitism (exemplified by the red arrows in Fig. 18). Indeed, we can consider the disease triangle as an alternative representation of a reaction norm, at one side of the norm being parasitism (antagonism) and at the other side finding mutualism (Roossinck, 2011; Roossinck and Bazán, 2017). Viruses may lie anywhere along this norm. They must use the resources of their hosts for replication, transcription and packaging, as many other symbionts do, but the nature of the relationship does not have to be always antagonistic. In principle, there is no reason why they cannot be commensals or even mutualists, if the benefits of such benign associations outweigh possible costs. Results of plant virus metagenomic studies have detected large numbers of asymptomatic infections in thousands of wild plants (Roossinck, 2012; Roossinck et al., 2010; Thapa et al., 2015; Wren et al., 2006), suggesting that viruses may be non-pathogenic members of microbial communities and that pathology appears as a consequence of the change in host's composition, transmission patterns and environmental settings (Bass et al., 2019; Roossinck and Bazán, 2017). The replacement of natural plant biodiversity by large-scale crop monocultures generally leads to the emergence of epidemics and more virulent viruses (Lefeuvre et al., 2019; Paudel and Sanfaçon, 2018). But this antagonistic effect is also important because the emerging viruses may control and stabilize novel ecosystems. On the one hand, they may help to maintain genetically diverse ecosystems and avoid overpopulation and homogenization by a certain plant species that may indeed help ecosystems to be more resilient to



environmental changes. On the other hand, viruses may help in preventing the invasion of ecosystems by foreign plant species by infecting them and being virulent enough to stop the invader from overrunning the ecosystem (Lefeuvre et al., 2019).



**Fig. 18.** The disease triangle illustrates the outcome of the interaction between a virus and its host plant in a given environmental condition. Parasitism, resulting in disease, only appears under certain combinations of the three factors, here indicated by the overlap between the three colors. Environmental changes and genetic modifications in both players may result in a different symbiotic interaction without resulting in disease development. Black arrows represent a possible evolutionary trajectory due to environmental or genetic changes that transform an antagonistic interaction into a commensal one. Red arrows indicate a possible evolutionary trajectory from a mutualistic relationship into a parasitic one. Obviously, the triangle is a simplification

of the real situation as it ignores the role of vectors. A tetrahedron in which the additional dimension would represent the effect of vectors would be a more appropriate representation.

Plant viruses cannot be understood without adding their vectors into the picture. This will add a new dimension to the disease triangle, transforming it into a tetrahedron. Indeed, growing evidences support the notion that viruses actively modify the response of plants against herbivores to maximize their own transmission. Infection may result in the alteration of plant nutritional quality, palatability or defense responses to herbivores, thereby facilitating their transmission to uninfected hosts (Dáder et al., 2017; Gibbs, 1980; Mauck et al., 2019). As illustrative examples of this virus-vector mutualistic interaction, we can firstly mention the case of tomato spotted wilt virus (TSWV) infections that transform infected pepper (*Capsicum annuum*) plants into a better host for its vector thrips (Belliure et al., 2004). When the plant is eaten by virus-free thrips, the thrips juveniles have a lower survival and developmental rate, while this is reversed when the plant is damaged by virus-infected thrips, therefore, making the virus-infected host a better home for new thrips offspring (Belliure et al., 2004). The mechanism behind this phenomenon involves the production of salicylic acid (SA) in TSWV-infected plants which counteracts the jasmonic acid (JA) response that works as a thrips repellent (Li et al., 2017; Roossinck, 2015b). A second interesting study by Mauck et al. (2010) used cucumber mosaic virus (CMV) to characterize interactions of its host squash (*Cucurbita pepo*) with the insect community in a semi-natural field setting. CMV induced an increase of plant volatile emissions that enhance aphid attraction to infected plants but it decreased the quality of the plant host, so the aphids disperse more rapidly when they acquire the virus. Another interesting finding came up: CMV-infected plants were significantly more

resistant to *Anasa tristis* bug (a non-vector insect) damage than non-infected plants. So, it appears the plants infected with CMV may escape additional biotic stressors, despite the obvious cost of virus-reduced fruit production and disrupted seed development (Mauck et al. 2010). From the virus evolutionary perspective, this is a beneficial situation because as long as the host avoids death it will persist longer in the environment and serve as inoculum for new CMV infections. A contrasting example of manipulation of vector's behavior has been very recently reported by Safari et al. (2019) for pepper cryptic virus 1 (PCV-1). PCV-1 infects jalapeño pepper (*C. annuum*) plants in a persistent manner and is vertically transmitted with 100% efficiency. In stark contrast with the above results for CMV, plants infected with PCV-1 are less attractive for the aphid vector *Myzus persicae* compared with PCV-1-free plants. Hence, PCV-1 protects plants from being superinfected by acute viruses that the aphids may be harboring. This is a clear example of mutualism: plants benefit for being protected from secondary acute infections and the virus minimizes within-host interspecific competition and ensures its own vertical transmission.

A lot of research focuses on studying the pathogenic effects that viruses have on their cultivated plant hosts, but nowadays this narrow view is changing. Now we are starting to see that the actual benefit or harm that viruses have on their host plants depend on direct and/or indirect interactions between the virus coinfecting the same individual host, the plant genotype and the environment. These complex multiplayer interactions may change the viral composition, leading to fitness costs or gains for the plant. One illustrative example can be found in the viruses infecting oat (*Avena sativa*), where infection rates and competition among viruses are affected by the balance of concentrations of nitrogen and phosphorous in the soil (Lacroix et al, 2014). If the oat

plants are grown on a phosphorous and nitrogen rich soil, cereal yellow dwarf virus (CYDV) -RPV prevalence is low. In contrast, if the concentrations of phosphorous are high and of nitrogen are low, barley yellow dwarf virus (BYDV) -PAV outcompetes CYDV-RPV in coinfections. In the opposite conditions, the antagonistic interaction between the two viruses disappears (Lacroix et al., 2014).

### **III 1.3 Plant Responses to Abiotic Environmental Stresses**

Plants are sessile organisms permanently exposed to their environment and so they are prone to be affected by adverse conditions, both biotic or abiotic. The abiotic stress can be defined as environmental conditions that reduce growth and yield below optimum levels (Cramer et al., 2011).

Abiotic stressors are a major factor in yield loss, causing a decrease in crop productivity by 50 - 80% depending on the crop species and geographical location (Shinozaki et al., 2015). The effect is not only noticeable in quantitative but also in a qualitative way: environmental stress factors perturb the amount of proteins and lipids of the harvest (Wang and Frei, 2011). This situation is only going to get worse under the conditions predicted for the ongoing climate change: *(i)* average temperatures will increase, especially affecting higher latitudes and cropland areas; *(ii)* global mean precipitation and runoff will also increase in higher latitudes and the equatorial region, but potential reductions in available water resources may happen in the middle latitudes; and *(iii)* relatively localized extreme events will increase in frequency and intensity (Bruinsma, 2003). Following the short-term climate conditions, the abiotic stress suffered by plants will be consequently more common and severe.

The nature of the abiotic stress depends on the conditions of the environment, so this can cause a diverse array of stresses: water-related stresses, extreme temperatures, UV-C radiation, deficiency or excess of nutrients, and chemical situations as abnormal quantities of metals, ozone or salts. Usually abiotic stressors happen in combination with other types of perturbations so it is expected that in nature two or more abiotic stresses might appear combined (Atkinson and Urwin, 2012). For the plant, the combination of different abiotic stresses is not equivalent to the response it would have for each stress individually: the plant responds in a unique way to the combination of stresses (Mittler, 2006). Pandey et al. (2015) reviewed the shared and unique responses to combinations of stresses. The interaction depends on different factors such as the severity of each stress, the age of the plant and the tolerance or susceptibility of the plant to the individual stress. Interaction between stresses can result in antagonisms or synergies, but the overall response is apparently governed by the more severe stress. In a transcriptomic and metabolic analysis *A. thaliana* response to a combination of heat, drought and infection with TuMV it was found that that gene expression under multifactorial stress is not predictable from single stress treatments (Prasch and Sonnewald, 2013).

Acclimation to adverse conditions is the strategy that plants use to deal with abiotic stress. The process starts with the plant receiving an incoming environmental signal related to stress, which triggers changes in the gene expression and leads to a more appropriate physiological status in the plant. Despite acclimation to stress being a relatively fast process (it can be accomplished in hours, as happens with heat stress), other conditions (such as cold and drought stress) may require days or even weeks (Gusta, 2012). Acclimation to a variety of stresses can be achieved by some species; the process can be less successful in many cultivated plants that can struggle to face environmental

changes since they may lack mechanisms that were lost during selection in order to achieve a higher crop yield (Basu et al., 2016). The acclimation strategy occurs in other sessile organism, such as corals, and it is likely that their reaction norms are also affected by the presence of parasites.

Plant are versatile organism that have a wide range of mechanisms to adapt to their environment and the interactions they have with other organisms may depend on their circumstances. For example, in alpine plant species competition was the dominant interaction between plants but in environments where the abiotic stress was high positive interactions predominated: biomass, growth and reproduction were higher when other plants were nearby (Callaway et al., 2002). So, plant responses to their environment depend on multiple facts and, even between plants of the same species belonging to different cultivars, an opposite stress-responsive strategy and a different molecular response may take place (Aprile et al., 2013).

### **III 2.1 Water Stress**

Plants can face an environment where the soil and/or the air moisture are outside the niche breadth of the plant. Whenever plants suffer a water deficit (drought) or excess (flood) an adverse impact on plant growth and yield results. If the situation lasts for longer periods it may even compromise the plant survival (Carr, 2017).

#### **III 2.1.1 Drought**

Xu et al. (2008) studied the role of different viruses in the induction of drought tolerance. They used four RNA viruses: brome mosaic virus (BMV), CMV, tobacco mosaic virus (TMV), and tobacco rattle virus (TRV). It was found that all these viruses induced some

sort of drought tolerance in their hosts. Xu et al. (2008) investigate this effect in a wide range of hosts and viruses (Table 5): rice (*Oryza sativa*) (infected with BMV), beet (*Beta vulgaris*), pepper, watermelon (*Citrullus lanatus*), cucumber (*Cucumis sativus*), squash, wild tomato (*Solanum habrochaites*), tomato (*Solanum lycopersicum*), and goosefoot (*Chenopodium amaranthicolor*) (all infected with CMV); and *N. benthamiana* plants infected with BMV, CMV, TMV or TRV. In addition, tobacco (*Nicotiana tabacum*) plants were infected with TMV. In all the interactions investigated, a delayed onset and severity of drought symptoms (wilted shoot tips or plant collapse) was observed. Infection with CMV delayed the appearance of drought symptoms by two and five days in comparison with the time it took after water withdrawal for mock-inoculated plants. CMV was able to inducing tolerance in a broad range of hosts, and the degree of the benefit was species-dependent. The pathogenesis of the virus does not seem to be a key factor in tolerance, as the infection at the time of water deprivation induced severe systemic symptoms in some hosts but mild or non-systemic ones in beet, goosefoot and wild tomato. Different RNA viruses in the same host also showed a variable degree of delay in the appearance of symptoms: depending on the virus species of the four viruses that infected *N. benthamiana* the plant showed a delay of two to five days in showing drought stress symptoms. In the case of the rice, the difference between infected plants and mock-inoculated ones was in survival: after a drought period, plants were re-watered for two weeks and while none of the non-infected plants recovered, all of the BMV-infected ones survived. At face value, these results suggest that the reaction norm for plant response to water has displaced the  $S_{min}$  towards lower values and hence increased tolerance to drought.

The metabolic changes under drought conditions were studied in the above-ground tissue of healthy and BMV-infected rice, and CMV- and BMV-infected beet plants. In BMV or CMV infections, the level of metabolites that changed were significantly reduced compared to the non-infected plants. The virus-infected plants showed an increase in the levels of SA, a hormone that gets activated in response to pathogens but is also involved in abiotic stress resistance in plants. An increase in the levels of osmoprotectants (*e.g.*, trehalose, putrescine and proline) and antioxidants (*e.g.*, anthocyanins, tocopherols and ascorbic acid) was associated with drought tolerance.

Westwood et al. (2013) used the CMV-*A. thaliana* pathosystem to study the mechanisms of virus-induced drought tolerance. They found that CMV infected plants had a higher water content than mock-inoculated ones and they presented no drought-related symptoms. However, when the plants were infected with a CMV mutant lacking the *2b* cistron (CMV $\Delta$ 2b), the infected plants were wilted and showed the same low tolerance as mock-inoculated ones. The 2b viral protein is involved in local and systemic virus movement, induction of symptoms and suppression of host RNA silencing. It seems that this viral protein is responsible for providing drought tolerance, at least in *A. thaliana* plants. Westwood et al. (2013) corroborated this function by creating two transgenic lines of arabidopsis expressing 2b. The two transgenic lines expressed 2b at different levels: the line with lower 2b levels showed milder viral-like symptoms. Despite that, the drought tolerance in both transgenic lines was similar: two-fold higher than in wildtype plants.

Aguilar et al. (2017) studied the effects of levels of virulence in drought tolerance. They used two experimental hosts (*A. thaliana* and *N. benthamiana*) and two viruses: plum pox virus (PPV) and potato virus X (PVX). They found that the tolerance induced



in *N. benthamiana* correlates with the virulence of the virus. Drought-tolerance was enhanced in plants under the synergistic coinfection of PPV and PVX. The tolerance was also enhanced in plants infected with PPV expressing PVX virulence protein P25. These differences in drought tolerance were correlated with differences in transpiration, since more virus-induced tolerance also caused a reduced rate of transpiration. The effects were similar in *Arabidopsis*, where virus infections led to higher levels of water content, with these levels being higher in plants infected with the engineered PPV-P25. However, the increase in tolerance was countered by the higher virulence effects on fecundity. Infection reduced seed production in well-watered plants, yet the situation changed in drought conditions in which infected plants produce a higher amount of seeds than non-infected ones. However, the PPV-P25 and the PPV-PVX infection caused such a large decrease in the progeny production that the induced drought tolerance did not compensate the reduced host reproduction. This indicates that a trade-off exists between virulence and the degree of drought tolerance in infected plants.

As the concurrence of parasites may be common, it is important to study the effects of multiple infections. Dastogeer et al. (2018) used *N. benthamiana* as a host to compare and study the combinatory effects of an endophyte fungus and a virus infection. They used the fungus *Cladosporium cladosporioides* and the yellow tailflower mild mottle virus (YTMMV). For separated infections, both the virus and the fungus induced drought tolerance but their coinfection had no additive effect. Comparing both infections, it was shown that both the virus and the fungus trigger similar changes in the host: increase in the osmolytes accumulation, antioxidant enzyme activity and expression of genes that are also induced by drought.

### **III 2.1.2 Flood**

An excessive amount of water can be very harmful to land plants. Submergence or soil waterlogging can compromise plant survival as plants may be subjected to hypoxia (Loreti et al., 2016). Andrews and Paliwal (1986) explored the role of BYDV infection in winter cereal tolerance to flooding. Wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*) were infested with viruliferous aphids and it was observed that the infection slightly reduced the aerial weight of barley and wheat both in flooded and non-flooded conditions. Nevertheless, the total number of unfertile and fertile tillers increased in BYDV-infected wheat, and this increase was higher after flooding. These positive effects were not found in barley, which suffered detriment in all circumstances. This result suggests that the reaction norm for plant response to water has displaced  $S_{max}$  towards higher values and hence increased tolerance to flooding.

Combining the results of the experiments we have just shown in subsections III 2.1.1 and III 2.1.2, we can conclude that viral infections can change the shape of reaction norm for plant response to water, widening the niche breadth  $S_{max} - S_{min}$  (Fig. 17B). However, the information available does not allow us to make any inference about the  $S_{opt}$  values.

### **III 2.2 Extreme Temperatures**

Each plant genotype has a specific optimum temperature within its thermal niche breadth. Thermal reaction norms typically have a common general shape in which performance increases with temperature, reaches an optimum at some intermediate value and then quickly declines if temperature keeps increasing (Fitter and Hay, 2002) (Fig. 17). Temperature is a principal regulator of plant growth and development, values very close to or outside the thermal niche breadth are a primary factor affecting the rate of plant

development. One of the most sensitive stages to temperature extremes is pollination (Hatfield and Prueger, 2015), which would affect plant fecundity and yield. Reduced plant fertility and yield obviously would impose a strong burden to virus transmission, especially for those viruses able to be vertically transmitted through seeds. Viruses also have a different range of temperatures where they are able to replicate in their host plants. Therefore, it is logical to expect that whenever a virus evolves in a plant under a certain temperature regime, it adapts its  $S_{opt}$  to that temperature (Jones, 2016). Furthermore, if viral infection induces a shift in the plant's thermal niche breadth, it will survive adverse environments and therefore the virus will be able to continue replicating and transmitting, thus having higher chances to adapt to even more extreme scenarios.

### **III 2.2.1 High Temperatures**

Yarwood (1957) studied the thermostability of cells in healthy green bean (*Phaseolus vulgaris*) leaves and compared it with the thermostability of the leaves infected by several viruses. As expected for a shift in the thermal reaction norm, he found that the cells surrounding the virus-induced necrotic spots had higher tolerance to heat.

Tomato yellow leaf curl virus (TYLCV) induces a paradoxical form of thermal tolerance in tomato plants (Gorovits et al., 2019). One of the consequences of TYLCV infection is the shutdown of the transcription and translation levels of heat-inducible genes (*e.g.*, the HSP90 chaperone complex) and the inhibition of the ubiquitin 26S proteasome degradation, resulting in a reduced cell death due to the thermal stress. By doing so, the virus keeps cells alive for longer, gaining extra time for its own replication. As a side effect, infected plants are tolerant to increased temperatures. Hence, TYLCV infection has shifted the  $S_{max}$  value of tomato thermal reaction norm upwards.

Usually plants accommodate diverse communities of symbiotic microbes (Bass et al., 2019) that play a role in plant development and in its interaction with the environment. Márquez et al. (2007) provided a particularly interesting example of thermotolerance induced in a plant not by a virus but by a consortium formed by the fungal virus *Curvularia thermal tolerance virus* (CThTV) and its host the fungal endophyte *Curvularia protuberata*. The tropical panic grass (*Dichanthelium lanuginosum*) lives in geothermal soils in Yellowstone National Park in the USA. When plants are colonized by the fungus, they show enhanced growth and greater tolerance to increased temperature and drought. In this case, neither the plant or the fungus can grow separately above 38 °C, but they can tolerate and survive up to 65 °C when a symbiotic association is established between them. The peculiarity of this three-ways relationship is that heat tolerance is only conferred when the fungus is infected with CThTV. Virus-free fungal isolates were not able to confer heat tolerance, but if the same fungal isolate was infected *de novo* with CThTV, they were able to confer heat tolerance to the plants at the same level the CThTV-infected fungus wildtype isolates did. Again, this study exemplifies the role of viral infection in shifting up the  $S_{max}$  value of the thermal reaction norm of the host fungus and plant.

### **III 2.2.2 Low Temperatures**

Fernández-Calvino et al. (2014) studied the gene expression and the metabolite profile of arabidopsis infected with TRV. They found that TRV-infected plants had a transcriptional response that broadly overlaps with the one reported in stressed plants. In particular, they found that putrescine levels were almost two-fold higher in infected plants than in non-infected ones and, as putrescine is essential for arabidopsis tolerance to low temperatures,

they studied the effect of the TRV infection in cold tolerance. Under a low temperature regime, the rate of plants that survived and produced siliques (an index of fertility and plant fitness) was higher among infected plants (67%) than in the mock-inoculated ones (48%). In order to confirm the role of the putrescine in TRV-induced tolerance, they repeated the experiment using an arabidopsis mutant defective in putrescine biosynthesis (*adc1-2*) that is less tolerant to low temperatures. This time, the survival of the mock-inoculated plants was 45%, while the survival of the TRV-infected plants reduced to a 55%, lower than in wildtype TRV-infected plants but still higher than in wildtype non-infected plants. This shows that in the TRV-*A. thaliana* pathosystem, putrescine plays a major role in cold tolerance but other mechanisms are also responsible for the tolerance phenotype.

In their study about tolerance to drought, Xu et al. (2008) also observed an increase in the putrescine levels in virus-infected plants. Then, Xu et al. (2008) explored virus-induced tolerance to cold. They submitted mock-inoculated and CMV-infected beet plants to low temperature conditions. After exposing plants to this cold environment for three days, all the non-infected plants died while the infected ones tolerated the frost stress and survived.

Andrews and Paliwal (1986) studied the interactions between BYDV and wheat and found a somewhat more intriguing relationship between tolerance to cold and virus infection. Plants of the cultivar Fredrick and Norstar were infected by aphid inoculation and then submitted to different conditions: pre-hardened (the plant were prepared for cold environments) or non-pre-hardened with short or long disease development periods. At odds with our hypothesis of viruses pushing down the  $S_{min}$  of the thermal niche breadth,

their results indicated a decline in the plant tolerance to cold when infected with BYDV. Nevertheless, non-pre-hardened Fredrick plants showed an interesting interaction with BYDV. Infected plants did not show ice tolerance and cold tolerance during the first two months, but after four months of exposure to low-temperatures the tolerance was significantly higher in the BYDV-infected plants than in the non-infected ones: the survival rate of the infected plants was 85% while only 30% of the non-infected plants survived. The authors measured the dry weight of all plants, finding that the infected ones had greater mass than the non-infected ones. They hypothesize that a higher dry weight would provide a greater energy reserve for the plant that can be used for anaerobic respiration during ice encasement. This conditional induced tolerance to low temperature was previously observed by the same authors in field experiments (Andrews and Paliwal, 1983): BYDV infection lead to different survival rates depending on the period of infection development and the host cereal species [black oat (*Avena strigosa*), barley, rye (*Secale cereale*), or wheat] or even the cultivar of wheat (Fredrick or Kharkov). Again, in general BYDV-infected plants had a reduced survival except in the case of wheat cv Fredrick where the survival in ice encasement was increased when plants were subjected to 17 days of BDYV infection followed by 33 days of cold hardening. Infected plants had, after seven days of ice encasement, a survival rate of 75% meanwhile only the 30% of healthy plants survived. Consistently, Fitzgera and Stoner (1967) studied the effects of BDYV in the survival of three winter wheat cultivars in South Dakota field conditions: infection caused an increase in plant survival in winter if they were infected when they were only seven days old; this survival was more pronounced if plants were infected after one month of fall growth. Together, these experiments with the BYDV-cereal pathosystems show that the shift down in  $S_{min}$  occurs under long periods of infection and not for all cereal species and cultivars.

As a conclusion from the experiments discussed in subsections III 2.2.1 and III 2.2.2, viral infections can change the shape of plant's thermal reaction norms, widening the niche breadth  $S_{max} - S_{min}$  (Fig. 17B). Again, as it was the case for water stress, the available information does not allow us to make any inference about whether the  $S_{opt}$  value has also changed or remains the same as for non-infected plants.

### III 2.3 Excess of Light Irradiation

So far, there is very little literature describing virus-induced tolerance to light stress in plants. However, there are examples of viruses infecting cyanobacteria and green algae that increase tolerance to this type of stress. These organisms are related phylogenetically to plants and share an evolutive history (Bhattacharya and Medlin, 1998; Hagemann et al., 2016).

Unlike vascular plants, aquatic green algae have an undeveloped mechanism of non-photochemical quenching (NPQ). This allows the dissipation of light excess trapped by the pigment pool of photosystem II. Seaton et al. (1996) described the NPQ changes induced in the phytoplankton *Chlorella* strain Pbi when it was infected with the *Chlorella* virus Marburg-1 (CVM-1). *Chlorella* cells showed an increased NPQ and a xanthophyll cycle activity when infected, leading to reduction in the amount of absorbed light energy from photosystem II. This would help the plant in situations of light stress where an excess of light energy would be detrimental for the organism. The CVM-1-induced increase of NPQ happened even at very low light intensity, reminiscent of the NPQ enhancement observed in vascular plants when infected with a microbial pathogen (e.g., Aucique-Pérez et al., 2017; Prokopová et al., 2010). This mechanism has led to the

hypothesis that both systems have an “acquired systemic response to excess excitation energy and to pathogens”, resulting in an enhanced NPQ induced by external stimuli of an abiotic or biotic nature (Finazzi and Minagawa, 2014). In the context of our reaction norm framework, this study suggests that CVM-1 moves the  $S_{max}$  value of *Chlorella*'s light intensity response reaction norm upwards.

Mann (2003) described the genomic sequence of cyanophage S-PM2. S-PM2 contain the *psbA* and *psbD* genes that encode two proteins D1 and D2, respectively, that are key components of photosystem II. They form a dimer at its core that binds pigments and co-factors necessary to perform photosynthesis. Light excess can cause a degradation of D1 and D2, leading to the need for replacement with new polypeptides. If the degradation is higher than the replacement, repair is not possible and a process of photo-inhibition occurs. Under such situations, an infected cell can take advantage of the virus-encoded D1 and D2 proteins to keep the photosystem II functional, allowing maintenance of the cell photosynthetic activity and virus replication. Therefore, under stressful light conditions S-PM2-infected cyanobacteria cells gain a clear advantage compared to non-infected cells while the virus promotes its own replication, a clear example of mutualism. These observations are also consistent with a virus-induced change in the shape of the light intensity response reaction norm.

### **III 2.4 Oxidative Stress**

Oxidative stress causes an increase in ROS production in chloroplasts that leads to higher levels of  $Ca^{++}$ , which can be harmful for plant cells if it persists for long periods of time. Tolerance of *N. benthamiana* to oxidative stress was studied by Shabala et al. (2010, 2011a, 2011b). *N. benthamiana* plants were more tolerant to oxidative stress when



infected with PVX. Stress was induced by exposing plants to intense UV-C light radiation, which caused a significant impairment of the electron transport chain and a progressive decline in the functionality of the plant's photosystem. PVX-infected plants tolerate this stress better than non-infected ones, as control plants showed high levels of  $\text{Ca}^{++}$  while the infected ones had intrinsically lower levels. These low  $\text{Ca}^{++}$  levels are achieved because PVX-infected plants have a more active  $\text{Ca}^{++}$  removal from the cytoplasm due to upregulation of two  $\text{Ca}^{++}$  efflux systems: the plasma membrane  $\text{Ca}^{++}/\text{H}^{+}$  exchanger and the endomembrane  $\text{P}_{2\text{A}}$  and  $\text{P}_{2\text{B}}$   $\text{Ca}^{++}$ -ATPases. Hence, PVX-infected *N. benthamiana* plants show a shift upwards in their tolerance to oxidative stress that translates into higher  $S_{\text{max}}$  values in the corresponding reaction norm.

### **III 3. MECHANISMS OF VIRUS-INDUCED TOLERANCE**

Little is known about molecular and physiological processes behind virus-induced tolerance to abiotic stresses (Ramegowda and Senthil-Kumar, 2015; Rejeb et al., 2014). The only consensus point is that mechanisms may differ widely among pathosystem. As it was stated above, the same host plant can express different responses to different viruses, and the same virus can induce different responses in different hosts. There is a real possibility that even cultivars of the same plant species may show differences. Likewise, different strains of the same virus may trigger different responses in the host genotype.

To tolerate stressful abiotic environments, plants have to be physiologically and phenotypically plastic, being capable of acquiring a metabolic status that maximizes the fitness of the individuals under harsh circumstances. If the stress is too intense, or the

plant does not have the metabolic capability to respond to it, the most likely outcome would be the death of the plant. Here we have worked around the hypothesis that viral infection may provide a possible mechanism to minimize the chances of such adverse situation by changing the shape of the reaction norm and broadening niche breadth. For a plant, hosting a virus represents an intense biotic stress; virus infection can disrupt the physiological homeostasis of the plant by altering hormone-mediated signaling pathways and their crosstalk and inducing new pathways. This new physiological status would be negative in the current permissive environment, but could be positive if infection has the side effect of triggering the overexpression of signaling pathways that result in enhanced tolerance to potential negative environmental changes. Furthermore, as the virus further evolves to adapt to the host in the new harsh environment, the level of tolerance to abiotic stress may even improve. Such situations may result in the evolution of a new mutualistic association.

### **III 3.1 Plant Phenotypic and Physiological Changes in Response to Infection**

Plants rely on crosstalk between complex signaling networks to respond to biotic and abiotic stress (Zhu, 2016). Each stress has a specific mechanism of response, which is tuned according to the stress intensity. However, responses to abiotic and biotic stresses are not fully independent, but instead signaling pathways are inter-wired and metabolic responses overlap. This leads to situations where, depending upon the combination of stresses, interactions will be negative or positive for the plant. In general, plant exposure to abiotic stressors results in increased susceptibility to pathogens and herbivores (Suzuki et al., 2014). On the other hand, biotic stress can induce a response in the plant that is also effective in tolerance of certain abiotic conditions.

As already mentioned, virus infections induce physiological changes in the plant, for instance enhanced accumulation of carbohydrates, particularly of starch, which can be used as an energy reservoir for the plant in order to survive low-temperature environments (Diener 1963). Infection can also increase the level of certain metabolites that are beneficial for the plant, especially osmoprotectants and antioxidants that can mitigate abiotic stresses (Dastogeer et al., 2018; Fernández-Calvino et al., 2014; Xu et al., 2008). Infected plants may also suffer higher levels of senescence, which may be beneficial under certain stress conditions as this process allows the plant to relocate resources and reduce the loss of water (Munné-Bosch and Alegre, 2004).

Pathogenic infections usually cause a reduction in the size of the plants (Hull, 2002), which immediately reduces water requirements and improves their tolerance to drought (Xu et al., 2008). A switch in the size and morphology also may be beneficial under adverse temperature conditions (Patel and Franklin, 2009).

In addition to the aerial parts of the plant roots also are modified under abiotic stress. Westwood et al. (2014) found that a transgenic plant expressing a CMV virulence factor 2b had an increased drought tolerance and, despite not being able to find changes in the root architecture, they concluded that the root physiology was altered to uptake and retain water more efficiently.

Plants change their hormonal status in order to coordinate antiviral responses. As mentioned, changes in the levels of different hormones can prime plants into a more favorable status to face adverse abiotic conditions. Drawing here a largely oversimplified picture of hormonal signaling crosstalk, we can focus on the two most essential players:

SA, whose accumulation dramatically increases in response to viral infection (Alazem and Lin, 2015), and abscisic acid (ABA) that regulates plant development and responses to abiotic stresses (Khan et al., 2015). During viral infection a regulatory crosstalk between SA and ABA is established. ABA is known for its implications in the regulation of abiotic stress but recently it has emerged as a positive or negative regulator of plant defenses, depending on the particular pathosystem (Denancé et al., 2013). ABA regulates different stages of the plant cycle through massive transcriptional reprogramming, impacting plant development and its tolerance to diverse stresses. The upregulation of ABA levels has been correlated with an increased tolerance to drought as it leads to stomata occlusion, preventing water losses. However, it has also been shown that high levels of ABA during long periods result in a vegetative growth retardation, photooxidative damage and seed abortion (Sreenivasulu et al., 2012). Therefore, as SA levels increase in response to viral infection, ABA levels can be reduced, inducing a tolerance status. This tolerance can happen even under prolonged adverse conditions, as high SA levels would keep the ABA homeostasis and help the plant to survive longer. Interestingly, ABA is a positive regulator of siRNA machinery that results in enhanced antiviral activity mediated by the AGO complexes (Alazem and Lim, 2015). Therefore, by suppressing ABA expression, the SA-mediated response could result in an increased susceptibility to viruses. Unfortunately, the tradeoff between SA and ABA expressions in plant susceptibility to viral infection have not received the well-deserved attention from an evolutionary point of view (*i.e.*, would evolving viral mutant swarms modify the plant's balance between SA and ABA?).

### III 3.2 Plants Short-Term Adaptation to Stress by Epigenetic Changes

Exposure to stress leads to selection of beneficial traits in plants but this process may take a long time depending on the evolutionary response of such traits to selection and the intensity and duration of the stress. Plants can still acquire a heritable resistance status in a single generation through a process called acclimation (Boyko and Kovalchuk, 2011). One of the most pervasive stresses faced by plants are pathogens. A way for the plants to defend themselves is by gene-for-gene-based resistance, where there is direct or indirect recognition of pathogen effectors (avirulence factors) by plant resistance gene products that results in local lesions and blockage of the systemic spread of the virus (Chisholm et al., 2006; van der Biezen and Jones, 1998). Plants that do not encode for resistance genes necessary to oppose to a particular parasite, for example the TMV resistance *N* gene absent in *N. benthamiana* SR1 plants, respond to pathogen's infection by a systemic signal. This systemic signal leads to an increase in the frequency of somatic homologous recombination (Yao and Kovalchuk, 2011). Boyko and Kovalchuk (2011) found that the progeny of *N. benthamiana* SR1 plants infected with TMV showed higher levels resistance to TMV. These progeny plants showed more polymorphisms at loci that have a 60% similarity to the Leu-rich repeat region of the *N* gene. Along with these higher levels of somatic homologous recombination, compared to non-infected plants, the progeny of TMV-infected plants showed increased resistance to methyl methane sulfonate, increase in callose deposition, higher expression of the *PATHOGENESIS-RELATED 1* gene, and a global change in methylation patterns in euchromatic areas. Hypermethylation could be a response to prevent genome reshuffling in the progeny (Boyko and Kovalchuk, 2011). These finding points towards an adaptive inheritance, epigenetic in nature (Kathiria et al., 2010), that can contribute to plant's phenotypic plasticity (Zhang et al., 2012). Since plants produce new vegetative tissue throughout

most of their lives, but do not produce the germline until late in development, all of these stress-related changes could be passed on as a form of epigenetic memory to the progeny. Accumulating evidences suggest that plants possess dual forms of inheritance based on genetic and epigenetic mechanisms (Arnholdt-Schmitt, 2004; Madlung, 2004; Takeda and Paszkowski, 2006), which allow them to follow long- and short-term strategies of adaptation, respectively. Another mechanism of epigenetic inheritance is based on non-symmetrical genome methylation maintained by siRNA, which is a highly flexible epigenetic marker that is quickly lost in the absence of the triggering siRNA. This mechanism enables continuous reprogramming of gene expression across the plant genome. This mechanism could lead to the selection of virus-susceptible individuals if the virus alters methylation patterns in stress-responsive genes (Diezma-Navas et al., 2019).

Hence, phenotypic plasticity can be bolstered by epigenetic changes (Zhang et al., 2012). It has been suggested that prolonged exposure to stress throughout generations could turn an epigenetic modification into a stable genetic modification, leading to a fixation of a novel trait (Dickins and Rahman, 2012).

### **III 3.3 Viral Population Dynamics and Virus Adaptation to Abiotic Stresses**

Environmental stresses can directly affect virus fitness and epidemiological parameters. For example, recent work by Bergès et al. (2018) showed the negative effect of drought stress in CaMV transmission rate and viral load in arabidopsis. Under well-watered conditions CaMV transmission rate, viral load and virulence increased, resulting in a faster spread of the virus, compared to drought conditions. In drought, transmission rate and virulence were negatively correlated traits, but without concomitant changes in viral

load. These observations lead to the conclusion that the optimal virulence of CaMV was strongly dependent on environmental permissiveness. How do viruses cope with such environmental stresses?

As rapidly evolving organisms, viruses can quickly adapt to adverse environments due to their high mutation rates, large population sizes and short generation times. The combination of these three factors facilitate rapid generation of genetic diversity that results in a highly complex and dynamic population structure known as quasispecies or mutant swarms. Error-prone RNA-dependent RNA-polymerases make more mistakes than DNA polymerases as they also lack proof-reading mechanisms, which is why RNA viruses have higher mutation rates than DNA viruses (Sanjuán et al., 2010). Not surprisingly, RNA viruses also show faster rates of evolution than DNA viruses (Duffy et al., 2008; Sanjuán, 2012). Whether such high mutation rates have been selected because they provide a fitness advantage in an always fluctuating world, or they are a side-effect of a parasitic lifestyle in which selection favors fast replication at the cost of high mutation is not yet obvious (Elena and Sanjuán, 2005; Belshaw et al., 2008). High mutation rates also create the conditions for selection to favor mechanisms of mutational robustness (Elena et al., 2006; Belshaw et al., 2008). Robustness is a population-level emerging property by which the phenotype (symptomatology) stays constant in the face of genomic mutations (Elena, 2012). Indeed, quasispecies theory predicts such population-level robustness, the so-called “quasispecies” effect, by which a wide spectrum of mutant genotypes exists together as an equilibrium population where the most fit genotype makes up the majority of sequences (Dolan et al., 2018; Perales and Domingo, 2019).

Sometimes, stressful abiotic conditions can last for long periods of time, thus allowing selection to favor the evolution of viral genotypes more tolerant to the stress, either by some intrinsic mechanism or by altering their cellular environment. However, the efficiency of selection may be hampered if mutations come at a fitness cost for another trait like growth or reproduction. It has been observed (Dessau et al., 2012; McBride et al., 2008; Wieczynski et al., 2018) that the RNA bacteriophage  $\phi 6$  is able to induce stress-tolerant mutants that withstand extremely high temperatures but show significantly reduced reproductive ability compared to wildtype  $\phi 6$ . This means that environmental fluctuations can lead to opposite selective forces, favoring stress-tolerant viral phenotypes under harsh environmental conditions but reproductively superior phenotypes during benign periods. Therefore, evolving the ability to tolerate environmental stress depends on the amount and frequency of stress experienced by the viral population and the way opposed selective forces, that result in phenotypic trade-offs, are distributed over time (Wilke et al., 2006). Wieczynski et al. (2017) showed that stress tolerance plays an important role in shorter periods of environmental perturbations, while the reproductive ability shows a greater importance in the face of larger intervals of environmental stressors.

Plant tolerance to viral infection is an interesting mechanism of host-virus interaction on the commensal side of the symbiosis continuum (Little et al., 2010). Tolerance is the accumulation of viruses within the host without any apparent negative effect on host fitness (Shukla et al., 2017). In general, tolerance comes with a reallocation of host resources. For instance, it has been shown that the tolerant interaction between CMV and arabidopsis results from a deviation of plant resources from vegetative growth to reproduction (Shukla et al., 2017). Tolerance may generate a beneficial ball park for



subsequent evolution of a mutualistic interaction as far as the host may receive some benefit, like protection of the plant against super-infection by other viruses, and the virus can benefit from longer periods of time where it is able to replicate and transmit horizontally, or even evolve the capacity for vertical transmission. Mechanisms of tolerance are as diverse as the amount of possible virus-plant pathosystems that may exist. A common emerging picture is that tolerance may result from the proper balance between plant defense responses and virus counter-defense responses mechanisms (Paudel and Sanfaçon, 2018).

These mechanisms of viral survival and adaptation to new environments could also result in benefits to the host. As we have discussed above, by adapting to environment-mediated stresses, viruses could trigger mechanisms in the host that would promote their own survival and give the virus more opportunities to replicate and spread to new hosts. Viral population dynamics change during periods of stress and can also change the response of the host. But it is yet to be studied in detail how abiotic stress affects the relationship between plant-virus pathosystems and the fate of diversity within the viral mutant swarm.

### **III 3.4 The Environment as a Key Factor in Host-Virus Interaction**

As we have seen, the environment plays a key role in the outcome of the interaction between a host and its parasites (Fig. 18). Browder (1985) showed that the rate of infection of barley cultivars with rust fungus was dependent on the season. In winter the rate was low, while in the summer it was significantly higher. Surprisingly at that time, this observation was reversed if oats were infected with a different rust fungus strain. This pathogen genotype-by-environment interaction shown in disease outcome is clearly

associated with an effect of the rust fungus in the oat thermal reaction norm. The degree of this effect on the host can be more or less intense depending on the context of the infection (Wolinska and King, 2009).

In a very illuminating study, Hily et al. (2016) tested the hypothesis that plant-virus interactions vary between the continuum from antagonism to conditional mutualism depending on environmental conditions. They designed an experiment with four arabidopsis CMV-infected and mock-inoculated ecotypes under different temperature and light conditions where they measured CMV virulence and plant tolerance to abiotic stress. They discovered that the interplay between CMV virulence and arabidopsis tolerance fits the expectation from the life-history theory, with resource reallocation from growth to reproduction. Life-history theory is all about resources allocation in organisms and the emerging tradeoffs between different fitness traits, and how such tradeoffs might be modified by environmental conditions. In the case of antagonism, the plant will divert resources to reproduction, thereby reducing growth needs. With highly virulent CMV strains, the pre-reproductive period of the plant will be shortened to produce progeny before resources would be depleted, while for less virulent CMV strains plants will extend the pre-reproductive period to compensate for the parasite damage. Hily et al. (2016) also pointed out another interesting feature of infected arabidopsis: overcompensation. Overcompensation is the situation when the infected plant has a higher fitness than the non-infected ones. This has been mostly studied in relation to herbivory in arabidopsis, where the plant reallocates resources for reproduction. Environment-dependent overcompensation can be viewed as a case of conditional mutualism or pleiotropic parasitism. This has been demonstrated in many plant-virus systems where the plant was attacked by herbivores (see above: Gibbs, 1980) or in water stress. Conditional mutualism

between the virus and the plant can happen under certain environmental conditions where the cost of parasitism changes and can be beneficial for the survival of the host. In general, longer host survival will also benefit the pathogen, as more generations of viral replication may take place and more chances for transmission may exist.

To fully understand the role of the environment in plant-virus interactions we need to take into account temperature changes, seasonal changes, geographical, chemical or physical barriers, and most importantly the availability and distribution of vectors (Engering et al., 2013). One example of the essential role of vectors comes from the observation that elevated temperatures lead to favorable conditions for some vectors that thrive at higher temperatures, thus altering the composition of the parasite community that will transmit among plants. But as we have been discussing, the virome may broaden the niche breadth of plants too. By doing so, viruses may get better access to potential vectors. For example, van Munster et al. (2017) reported that cauliflower mosaic virus (CaMV) and TuMV infected plants grown under drought were better transmitted by vectors. Comparing the results of transmission experiments in drought and permissive conditions, CaMV had a 34% increase in vector-transmission rate while the effect on TuMV was 100%. No differences existed in viral accumulation between drought-stressed plants and well-watered plants, so the explanation for the observed differences in transmission rates was other than simply virus accumulation. In the other hand, parasites can also produce an effect in the environment through the impact on host growth. Considering the close relationship between the host-vector and pathogen-vector, it would be appropriate to include the vector as an extra dimension in the disease triangle (Fig. 18) (Jones and Barbetti, 2012): vector feeding can alter plant growth thus affecting virus

replication; pathogen replication can affect the attractiveness of plants to their vectors (Mauck et al., 2018; Roossinck, 2015a, 2015b; Safari et al., 2019).

There are two factors that are having a big role in plant spatiotemporal dynamics and environmental changes that will have a tremendous impact in plant virus epidemics (Jones and Barbetti, 2012). First is the ongoing global warming is dramatically changing environmental conditions that will lead to changes in the plants, from the genome to the ecosystem levels (Garret et al., 2006). This will clearly affect plant disease resistance and the ability of pathogens to infect and propagate successfully. Viral epidemics are increasing in higher and lower latitudes while the opposite is being observed in the subtropical arid regions. Second is human management of expanding agroecosystems. Monocultures reduce the habitat diversity and increase crop density, which leads to loss of genetic diversity. These alterations induce a change in the pathogen's infection risk and the eco-evolutionary dynamics of viruses. Plant community diversity has a great importance in virus prevalence that can be predicted from anthropogenic factors and ecological conditions (Fraile et al., 2017). The genetic diversity of tolerance and susceptibility genes in plant populations can also shape the evolution of the pathogen. González et al. (2019) have recently shown with the *A. thaliana* - TuMV pathosystem that increasing genetically heterogeneous plant populations negatively affects the rate of virus evolution, at the cost of increased virulence (González et al., 2019).

### III 4. CONCLUDING REMARKS

In this review we have shown examples of viral infections that contribute to increase the tolerance of infected plants to diverse abiotic stresses. This phenomenon can be contextualized in the theory of phenotypic plasticity and the resulting reaction norms. As presented, phenotypic plasticity would be a universally beneficial trait. Therefore, why are not all plant phenotypes maximally plastic? Obviously, the answer once again results from the balance between benefits, costs and limits. The benefits are obvious and have been extensively discussed along this review. Costs and limits of phenotypic plasticity are thought to have important ecological and evolutionary consequences, and we cannot wrap up this review without mentioning them. DeWitt et al. (1998) identified four main costs: (i) the *maintenance cost* associated with the expression and production of sensory and regulatory mechanisms even in absence of stimuli; (ii) the *information acquisition cost* necessary for interacting with the environment and identification of correct environmental cues before triggering plasticity mechanisms; (iii) the *developmental stability cost* inherent to the one-to-many genotype-to-phenotype developmental maps that may result in reduced fitness under stabilizing selection; and (iv) *genetic costs* such as the linkage between plasticity genes and other undesired genes, negative pleiotropic effects of plasticity genes in other traits or epistatic interactions between plasticity genes and the expression of other genes. Likewise, DeWitt et al. (1998) proposed four limits to the benefit of phenotypic plasticity: (i) an *information reliability limit* consequence of the production of maladapted phenotypes if the environmental cues are not processed correctly; (ii) a *lag-time limit* that emerges from the unavoidable delay between detecting the environmental stress and producing the corresponding new phenotype; (iii) development from starting phenotype to the final one may involve intermediate stages

with low fitness, thus creating a *developmental range limit*; and (iv) the so-called *epiphenotype limit* that means epigenetically encoded traits may be less effective than if the same trait is integrated during early plant development. As we have illustrated in this review, virus infection may alleviate some of these costs and relax the limits either because (i) the activation of antiviral responses beneficial in the very short-term will switch on responses to abiotic stress, (ii) the physiological host status is modified by the virus for its benefit and this new status fits better than the previous one in a stressful environment or (iii) the accumulation of viral particles may alter the physicochemical properties of cells making them more resistant to stresses such as desiccation. Regardless of the molecular details, the conclusion is that by alleviating costs and relaxing limits of phenotypic plasticity, reaction norms change their shape, generally resulting in a wider niche breadth.

The efficiency of selection on host and parasite genotypes depends on the amount of genetic variance, which is known to be environment-dependent (Falconer and Mackay, 1995). Consequently, environmental variation may influence the intensity of coevolution, potentially creating coevolutionary cold and hot spots in different environments. The above reviewed experiments suggest that abiotic stressful situations for plants may favor a win-win situation in which plant and virus may coevolve from an antagonistic relationship towards a commensal or even mutualistic situation (Vale et al., 2008b).

We have collected diverse evidence demonstrating that even pathogenic viruses may become beneficial for their hosts under adverse abiotic circumstances. Environment modulates the interaction between the plant host and the virus. In abiotically stressful environments, viruses can provide their plant host with increased tolerance. So far,

induced tolerance to drought, flood, high and low temperatures, high light, and oxidative stress were described. The mechanisms that rule the induced tolerance are convoluted and seem to be distinct in each host-virus-environment situation. Environment plays a key role in the outcome of the virus-host interaction; whether the plant is damaged or benefited by a virus infection may depend on the environment surrounding the interaction. If the interaction is beneficial, both parts are favored under the new adverse environment: infected hosts see an increase in their fitness in the new environment; and viruses are able to propagate and find new hosts. Since their discovery at the end of the XIX<sup>th</sup> century, plant viruses have been studied as pathogens. Brought on by the symptoms caused by some viral infections, the focus of the studies was to search for symptoms, caused by pathogenic viruses. Modern techniques have shown us that we are surrounded by viruses, most of them being asymptomatic. A great deal is still unknown about the asymptomatic viruses but it seems clear that most of them do not cause any apparent harm to their host and are essential components of the now called phytobiome (Bass et al., 2019).

The research of virus-induced tolerance to abiotic stresses is promising, especially in the perspective of warmer and drier climate conditions. There are still a lot of unknown aspects and unanswered questions in the evolution of the virus-host pathosystem under abiotic stress and its ecological consequences. A lot of additional information concerning the mechanisms that rule the induced tolerance needs to be accumulated. Furthermore, wild ecosystems need to be deeper analyzed to better understand the biology of plant viruses and their hosts and vectors in natural systems. Studying the response of virus infected organisms under all sorts of environments can help us to (i) understand the dynamics of heterogeneous complex systems, (ii) predict and modify the outcome of

virus-host interactions and (iii) discover and improve approaches to deal with adverse conditions.





## CHAPTER IV

---

### **Plant virus evolution under strong drought conditions results in a transition from parasitism to mutualism**

---

#### **IV ABSTRACT**

Environmental conditions are an important factor driving pathogens evolution. Here we explore the effects of drought stress in plant virus evolution. We evolved turnip mosaic potyvirus in well-watered and drought conditions in *A. thaliana* accessions that differ in their response to virus infection. Virus adaptation occurred in all accessions independently of watering status. Drought-evolved viruses conferred a significantly higher tolerance to drought to infected plants. By contrast, non-significant increases in tolerance were observed in plants infected with viruses evolved under standard watering. The magnitude of this effect was dependent on the plant accessions. Differences in tolerance were correlated to alterations in the expression of host genes, some involved in regulation of the circadian clock, as well as in deep changes in the balance of phytohormones regulating defense and growth signaling pathways. Our results show that viruses can promote host survival in situations of abiotic stress, being the magnitude of such benefit a selectable trait.

#### **IV 1. INTRODUCTION**

Viruses are the most abundant biological entities, having enormous diversity and ubiquitous distribution (Páez-Espino et al., 2016). Traditionally, they have been studied in the context of disease but nowadays numerous beneficial viruses are being identified in a diverse range of host species (Roossinck, 2011). Wild plant populations are

frequently asymptotically infected with viruses that in some cases produce diseases in cultivated plants (Roossinck, 2012). This happens because host-virus interactions fall on a spectrum between pathogenesis and mutualism and during their lifecycle viruses might switch between these two lifestyles (Roossinck, 2015; Roossinck and Bazán, 2017). This transition may happen depending on the environment and the genetics of hosts and viruses (Prendeville et al., 2012; González et al., 2020).

To face frequent environmental abiotic perturbations, plants have evolved acclimatation and tolerance mechanisms. Plant responses triggered by some stressors interacts with the response caused by others, *e.g.*, drought and cold (Shinozaki et al., 2003). This also happens between abiotic and biotic stresses (Atkinson and Urwin, 2012; Atkinson et al., 2013), meaning that under certain environmental perturbations (*i.e.*, water availability, extreme temperatures, excess of light irradiation, or oxidative stress) even pathogenic viruses can be beneficial for their host, since infection induces changes in the plant physiological homeostasis that may enhance its survival (González et al., 2020). Drought is one of the main plant stressors that, depending on its intensity and duration, causes major fitness reductions or even death. Xu et al. (2008) showed that plants infected with certain viruses can improve their tolerance to drought. It has been shown that the combination of drought and infection with TuMV affects different signaling networks in *A. thaliana* plants (Prasch and Sonnewald, 2013). Additional examples have been recently summarized in (González et al., 2020). Environmental perturbations can also affect pathogens evolution as changes in the environment can influence the specificity of selection (Wolinska and King, 2009).

Here we have studied how severe drought conditions influence TuMV evolution and, more interestingly, modified the way in which the virus and the host interact at different physiological levels. Firstly, we describe phenotypic and genotypic changes occurring in TuMV along evolution. Secondly, we characterize the differences in the host's transcriptome and hormonal profiles resulting from the infection with viruses evolved under permissive or drought conditions. Thirdly, we evaluate the selective pressures on the virus, specifically testing if evolution under drought conditions selected for viruses that provide an increased tolerance to this adverse environment that is plant genotype-specific or general.

## **IV 2. MATERIALS AND METHODS**

### **IV 2.1 Plant Material.**

Four accessions of *A. thaliana* were used as hosts. These accessions showed different response to potyvirus infection (Hillung et al., 2012), being classified into two groups according to their response to the potyviral infection: (i) G1, inducing a severe infection, up-regulating defense genes and shutting down the production of cell wall components (St-0 and Ler-0) and (ii) G2, with milder symptoms and lower virus accumulation, up-regulating genes involved in abiotic stress and cell wall construction (Wt-1 and Oy-0).

The selected accessions were exposed to standard watering and drought conditions. Standard conditions consisted of watering every two days until the plants were harvested at 14 days post inoculation (dpi). Drought conditions consisted of water withdrawal from 7 dpi until 14 dpi (time at which the plant tissue was harvested).

The evolution experiment was performed in a BSL-2 greenhouse at 24 °C with 16 h light:8 h dark photoperiod. The rest of the experiments were done in a growing chamber at 24 °C with 16 h light:8 h dark photoperiod, 45% relative humidity and 125  $\mu\text{mol m}^{-2}\text{s}^{-1}$  of light intensity (1:3 mixture of 450 nm blue and 670 nm purple LEDs).

#### **IV 2.2 Experimental Virus Evolution.**

The infections were initiated using homogenized TuMV-infected tissue preserved at  $-80$  °C. This virus stock was created from infected tissue of *N. benthamiana* plants previously inoculated with an infectious clone derived from TuMV isolate YC5 (GenBank accession AF530055.2) from *Zantedeschia* sp. (Chen et al., 2003).

The stock was used to inoculate four *A. thaliana* accessions. The inoculum used consisted of 100 mg of homogeneous N<sub>2</sub>-frozen infected tissue mixed with 1 mL of phosphate buffer and 10% Carborundum (100 mg/mL). For each accession 10 plants were inoculated and kept in standard conditions (well-watered) until infected plants were harvested at 14 dpi and another 10 under drought conditions (no watering from 7 dpi until the harvest at 14 dpi). Only the symptomatic infected plants were collected, making a pool of infected tissue from each condition and accession, using it as inoculum to start a five-passage evolution. For each accession and condition, three lineages were established (Fig 19A).

#### **IV 2.3 AUDPS.**

Upon inoculation, plants were inspected daily for visual symptoms. The infectivity data during the 14 dpi were used to calculate the *AUDPS* as described by Simko and Piepho (2012). This formula transforms data from disease progression, allowing us to express

the virulence and dynamics of the disease into a single figure. The *AUDPS* ranges between zero and the total number of observation time points along the experiment; larger *AUDPS* values mean that the virus infects a higher number of plants more quickly. *AUDPS* values were computed using the *agricolae* R package version 1.3-2 with R version 3.6.1 in RStudio version 1.2.1335.

Depending on the particular experiment being analyzed, *AUDPS* data were fitted to two fully factorial generalized linear model (GLM). In the first type of experiments, plant accession (*A*) and environmental conditions (*C*) were treated as orthogonal factors and evolutionary passage (*t*) as a covariable. The full model equation reads

$$AUDPS_{ijk}(t) \sim \alpha + A_i + C_j + P + (A \times C)_{ij} + (A \times t)_i + (C \times t)_j + (A \times C \times t)_{ij} + \varepsilon_{ijk} \quad (\text{Eq. 6})$$

where  $\alpha$  stands for the intercept, and  $\varepsilon_{ijk}$  represents the Gaussian error associated with each individual *k* plant measured at passage *t*.

In the second type of experiments, plant accession (*A*), environmental conditions being tested (*C*), and environmental conditions where the virus evolved (*E*) were treated as orthogonal factors. The full model equation now reads

$$AUDPS_{ijkl} \sim \alpha + A_i + C_j + E_k + (A \times C)_{ij} + (A \times E)_{ik} + (C \times E)_{jk} + (A \times C \times E)_{ijk} + \varepsilon_{ijkl} \quad (\text{Eq. 7})$$

where  $\alpha$  and  $\varepsilon_{ijkl}$  had the same meaning than in the Eq. 6. In both cases, a Gaussian distribution and identity link function were chosen based on the minimal *BIC* value among competing models. Hereafter, all GLM fitting were done with SPSS version 26 software (IBM).

#### **IV 2.4 *In silico* Evaluation of Functional Effects Associated with Observed Mutations in VPg.**

The functional effects of the mutations in the VPg protein were studied *in silico* using the Screening for Nonacceptable Polymorphisms (SNAP2) web server ([rostlab.org/services/snap2web/](http://rostlab.org/services/snap2web/); last accessed May 20, 2020). SNAP2 machine learning tools provide a score for all possible variants at each residue of the protein (Hecht et al., 2015). This score indicates if there is any effect of the variant in the protein function, regardless if the effect is positive or negative. The score value ranges between  $-100$  (no effect) and  $100$  (maximal effect).

#### **IV 2.5 Next Generation Sequencing.**

RNA was extracted from infected plant tissue using Plant RNA Isolation Mini Kit (Agilent). The quality of the RNAs used to prepare RNA-seq libraries was checked with the Qubit RNA BR Assay Kit (ThermoFisher). SMAT libraries, Illumina sequencing (paired end, 150 bp), and quality-check of the mRNA-seq libraries were done by Novogene Europe (UK). Seventeen bases from the 5' end and twelve from the 3' of the reads were trimmed with cutadapt v2.10 (Martin, 2011). Trimmed sequences were mapped with HiSat2 v2.1.0 (Kim et al., 2019) to the ENSEMBL release 47 of the *Arabidopsis* TAIR10 genome assembly. For viral genome SNP calling, trimmed reads were mapped with HiSat2 to the TuMV isolate YC5 with a modified minimum score

parameter ( $L$ , 0.0,  $-0.8$ ) to allow more mismatches. Resulting SAM files were BAM-converted, sorted, indexed and analyzed with SAMtools v1.10 (Li et al., 2009). SNP calling was performed using bcftools v1.6 by first using the mpileup subroutine. Read counting in features was done with htseq-count, using The *Arabidopsis* Reference Transcript Dataset (AtRTD2) (Zhang et al., 2017) as input annotation file. Differential expression analysis was done with DESeq2 v1.24.0 (Love et al., 2014), considering only genes having a total of at least 10 reads for each pairwise comparison. Characterization of DEGs was done with plant GOSlim implemented in the Cytoscape plugin Bingo (Maere et al., 2005) and MapMan (Thimm et al., 2004). Functional profiling was done using gProfiler (Raudvere et al., 2019).

#### **IV 2.6 RNA Isolation and cDNA Synthesis.**

Total RNA was extracted from plant tissues using Total Quick RNA Cells and Tissues Kit (Talent SRL), following the protocol established by the manufacturer. Further, DNase treatment (TURBO DNA-free Kit, Ambion) was performed to remove genomic DNA. RNA quantification was performed by spectrophotometric analysis and its integrity was checked by denaturing agarose gel electrophoresis. The absence of genomic DNA from the RNA samples was additionally tested by the null PCR amplification of the universal rDNA primer pair ITS1/ITS4. Then cDNA was synthesized from 2  $\mu$ g of total RNA, using SuperScript III H-Reverse Transcriptase (Invitrogen) and 100 pmol of random hexamers (Pharmacia Biotech) according to suppliers' instructions.

#### **IV 2.7 Quantitative RT-PCR Analysis.**

Quantitative real time PCR (RT-qPCR) was performed on a Thermal Cycler CFX96™ Real-Time System (BIO-RAD) using Power SYBR® Green PCR Master Mix (Applied



Biosystems), 11  $\mu\text{L}$  reactions contained 4.9  $\mu\text{L}$  of 1:6 diluted cDNA samples (8.5 ng of cDNA), 0.3  $\mu\text{L}$  (300 nM) of each primer (forward and reverse) and 5.5  $\mu\text{L}$  of SYBR<sup>®</sup> Green PCR Master Mix. PCR conditions were as follows: two initial steps of 50 °C for 2 min and 95 °C for 2 min, followed by 40 cycles of 95 °C for 30 s and 60 °C for 30 s. Afterwards, the dissociation protocol was performed to identify possible unspecific products. Three biological replicates per treatment were analyzed by RT-qPCR. For each transcript, the threshold cycle ( $C_T$ ) was determined using Bio-Rad CFX Manager 3.1 software. Primers used in the RT-qPCRs are described in Supplementary File S2 at [doi.org/10.1093/gbe/evz069](https://doi.org/10.1093/gbe/evz069).

Viral load was estimated by RT-qPCR using primers that amplify the *CP*. Viral load data, that were fitted to a fully factorial GLM with the same factors and structure than Eq. 7.

#### **IV 2.8 Infection Matrices.**

A full cross-infection experiment was performed where all the 22 evolved lineages were inoculated into ten plants of all four accessions. Infection matrices were analyzed using tools borrowed from the field of community ecology (Weitz et al., 2013). The statistical properties of the resulting infection matrices were evaluated using the bipartite R package version 2.11 (Dormann et al., 2008) in R version 3.6.1 (R Core Team 2016) in RStudio version 1.2.1335. Three different summary statistics were evaluated:  $T$ -nestedness (Bascompte et al., 2003),  $Q$ -modularity (Newman., 2006), and overall specialization  $d'$  index (Blüthgen et al., 2006).  $d'$  is based in Kullback-Leibler relative entropy, that measures variation within networks and quantifies the degree of specialization of

elements within the interaction network. Statistical significance of  $T$  and  $Q$  was evaluated using Bascompte et al. null model (Bascompte et al., 2003).

#### **IV 2.9 Survival Analysis.**

Lineages evolved under drought or standard conditions in a certain accession where inoculated in 24 plants of the same accession. Twenty-four plants were mock-inoculated as control. Seven dpi, a severe drought was simulated in the plants by withdrawing water for 14 days. After this period of drought plants were watered again during seven days and their survival was evaluated. This experiment was done twice for each one of the four accessions.

Survival frequency ( $S$ ) data were fitted to a factorial GLM in which natural accession ( $A$ ) and type of virus inoculum ( $V$ ) were treated as orthogonal random factors. A Binomial distribution and logit link function were chosen based on the minimal  $BIC$  value among competing models. The full model equation reads

$$S_{ijk} \sim \Sigma + A_i + V_j + (A \times V)_{ij} + \varepsilon_{ijk}, \quad (\text{Eq. 8})$$

where  $\Sigma$  corresponds to the model intercept, and  $\varepsilon_{ijk}$  represents the Gaussian error associated with each individual  $k$  plant.

#### **IV 2.10 Hormone Quantification.**

Hormone extraction and analysis were carried out as described in (Durgbanshi et al., 2005) with few modifications. Briefly, plant tissue was extracted in ultrapure water in a MillMix20 (Domel) after spiking with 10 ng of [ $^2\text{H}_2$ ]-IAA and 50 ng of the following

compounds: [<sup>2</sup>H<sub>6</sub>]-ABA, [<sup>13</sup>C]-SA, [<sup>2</sup>H<sub>3</sub>]-PA and dihydrojasmonic acid. Following centrifugation, supernatants were recovered and pH adjusted to 3.0. The water extract was partitioned against diethyl ether and the organic layer recovered and evaporated under vacuum. The residue was resuspended in a 10:90 CH<sub>3</sub>OH:H<sub>2</sub>O solution by gentle sonication. After filtering, the resulting solution was directly injected into an Acquity SDS ultra-performance LC system (Waters Corp.). Chromatographic separations were carried out on a reversed-phase C18 column (50×2.1 mm, 1.8 μm particle size, Macherey-Nagel GmbH) using a CH<sub>3</sub>OH:H<sub>2</sub>O (both supplemented with 0.1% acetic acid) gradient. Hormones were quantified with a TQS triple quadrupole mass spectrometer (Micromass). Multivariate analysis was performed using the package FactoMineR (Lê et al., 2008) in R version 3.6.1 (R Core Team 2016) in RStudio version 1.2.1335.

Hormones concentration were fitted to a fully factorial GLM with the same factors and structure than Eq. 7.

### **IV 3. RESULTS**

TuMV was evolved in four different natural accessions of *A. thaliana* that vary in their responses to infection with potyviruses (Lalić et al., 2010; Hillung et al., 2012). These accessions classified into two groups according to their phenotypic and transcriptomic responses (Hillung et al., 2012): accessions in Group 1 (G1) *Ler-0* and *St-0* showed severe symptoms and strong induction of defense genes; accessions in Group 2 (G2) *Oy-0* and *Wt-1* showed milder symptoms and over expression of genes involved in abiotic stress.

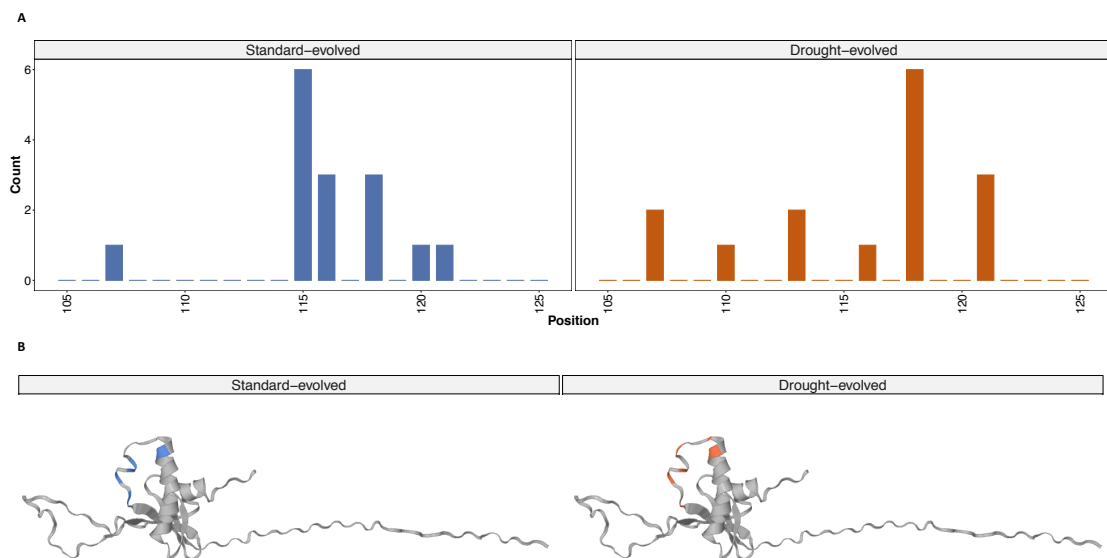
An *A. thaliana*-naïve TuMV isolate was evolved in each of the accessions for five passages in standard watering or drought conditions.

#### **IV 3.1 Phenotypic and Genotypic Changes in TuMV Lineages Evolved Under Standard and Drought Conditions.**

At the end of the evolution experiment (Fig. 19A) we obtained 12 lineages evolved in standard and 10 in drought conditions (two of the Wt-1 lineages evolved under drought conditions became extinct). All resulting lineages had experienced significant increases in their capacity to infect: infect more plants and show symptoms faster. Disease progression has been summarized using the area under the disease progress stairs (*AUDPS*) (Simko and Piepho, 2012), a value that integrates both infectivity and the speed of inducing symptoms (Fig. 19B). For the viral lineages evolved in G1 accessions, the increase in *AUDPS* was significantly larger when plants were grown in standard than in drought conditions ( $P < 0.001$ ; Fig 19B). Viruses evolved in G2 accessions also showed a significant increase in *AUDPS* relative to the ancestral virus, but this increase was larger for the lineages evolved in plants grown in drought conditions compared to the standard conditions ( $P < 0.001$ ; Fig. 19B). When facing abiotic stress, plants adjust their gene expression and metabolism to cope with to the stress (Krasensky and Jonak, 2012). These physiological changes may facilitate or jeopardize virus adaptation depending on the host genetics.



Next, seeking to characterize the spectrum of mutations in the evolved viral genomes, the nucleotide sequences of the ancestral and evolved viruses were obtained (Fig. 19C). Viruses evolved in standard conditions accumulated 32 mutations, five were fixed and 27 were polymorphisms. Nonsynonymous substitutions were the most common type, 27 out of 32 mutations. These mutations were not randomly distributed along the viral genome, but mainly concentrated in the *VPg* cistron (15 out of 32). Viruses evolved in drought conditions accumulated 26 mutations, only one was fixed and 25 remained polymorphisms. Again, most mutations were nonsynonymous (21) and preferentially were observed in the *VPg* cistron. Interestingly, all mutations observed in the *VPg* fall within a narrow domain encompassing amino acids 107 - 120 (Fig. 20).



**Fig 20.** Nonsynonymous mutations in VPg for standard- (left, blue color) and drought-evolved (right, orange color) viruses. (A) Distribution of the mutations. (B) Location of the mutations in the predicted tridimensional structure of VPg.

As an intrinsically disordered viral protein (Hébrard et al., 2009), VPg plays a role in virus-virus and virus-host protein-protein interaction networks (Jiang and Laliberté,

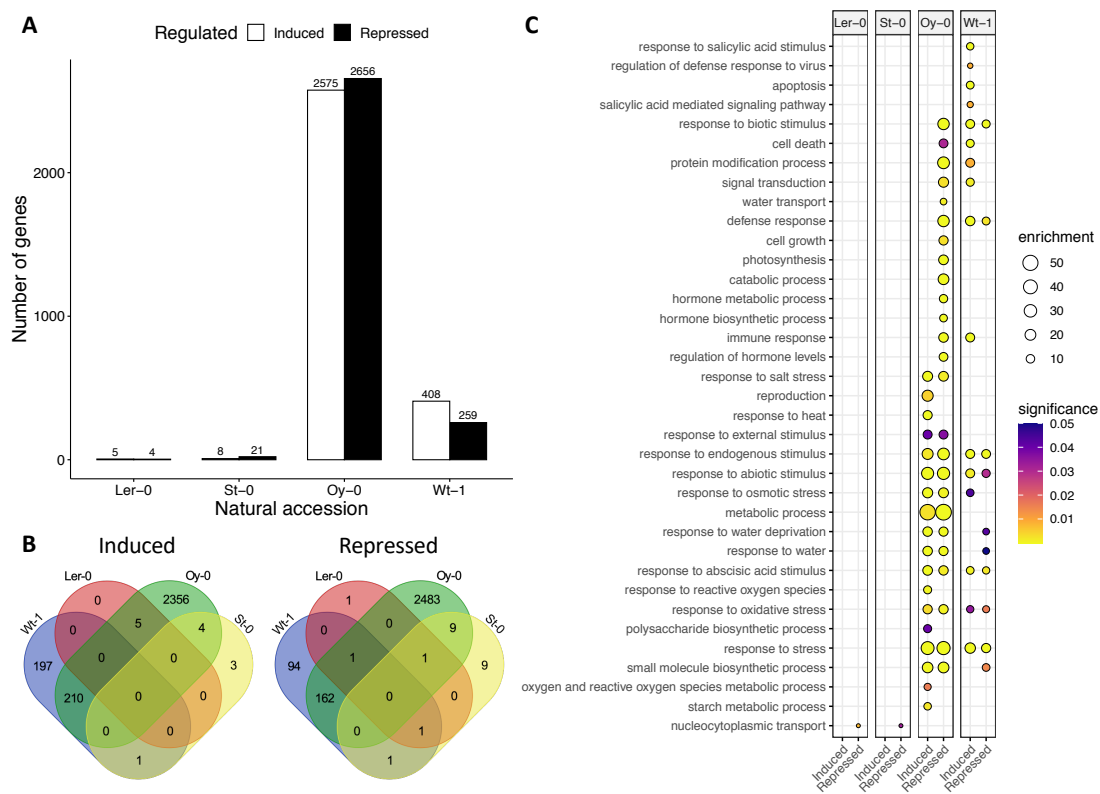
2011; Martínez et al., 2016). It is involved in virus movement, genome replication and suppression of host antiviral RNA silencing (Revers and García, 2015; Cheng and Wang, 2017). The functional effects of the mutations in the VPg protein were studied *in silico* using SNAP2 webserver (Hecht et al., 2015). Mutations fixed in standard-evolved viruses were predicted to have a significantly weaker effect (mean  $\pm 1$  SD =  $-0.400 \pm 22.831$ ) than those fixed in drought-evolved viruses ( $22.067 \pm 26.797$ ) (two-samples *t*-test,  $t_{28} = 6.109$ ,  $P = 0.020$ ), which are predicted to be more structurally and functionally disruptive.

#### **IV 3.3 Changes in Host's Transcriptomes when Facing Drought and Virus Infection.**

The whole-genome transcriptomic profiles of plants grown in drought conditions and infected with the drought-evolved viruses were compared with the transcriptomes of plants kept in standard conditions and infected with the standard-evolved viruses. Overall, the number of differentially expressed genes (DEGs) was significantly lower in the G1 than in the G2 accessions (Fig 21A;  $\chi^2 = 39.953$ , 3 d.f.,  $P < 0.001$ ).

The functional profiling of the DEGs (Dataset S1 at doi:10.1073/pnas.2020990118/-/DCSupplemental) shows a significant over-representation of genes involved in circadian rhythm in the under-expressed DEGs of the G1 accessions. All accessions share a few over- or under-expressed genes with other accessions (Fig. 21B), being some of those few shared genes also associated with circadian biological events. *Ler-0*, *Oy-0* and *St-0* show in common the under-expression of *PSEUDO-RESPONSE REGULATOR 5* (*PRR5*). The *PRR5* protein is a transcriptional repressor of the MYB-related transcription factors involved in circadian rhythm *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL 1* (*LHY1*). The repression of *PRR5* would lead to higher levels of *LHY1* expression, a gene that promotes expression of ABA-responsive genes

responsible for increased tolerance to drought (Adams et al., 2018). *FLAVIN-BINDING KELCH REPEAT F BOX 1 (FKF1)*, another gene involved in circadian rhythm, is under-expressed in *Ler-0*, *St-0* and *Wt-1*. FKF1 stabilizes *CONSTANS (CO)* expression and a reduction in *FKF1* expression will result in lower *CO* activity. A *CO*-like gene in rice has been shown to reduce drought resistance when overexpressed and to increase drought tolerance when knocked out (Liu et al., 2016). These observations align with evidences supporting circadian clock as a contributor to plants tolerance to abiotic stresses (Grundy et al., 2015). Circadian rhythms also play a role in infection, as it affects traits that could increase the fitness of both hosts and parasites (Westwood et al., 2019).



**Fig. 21.** Transcriptomic responses of different *A. thaliana* accessions to TuMV infection.

In each accession, the response of a pool of eight to ten plants infected with each one of



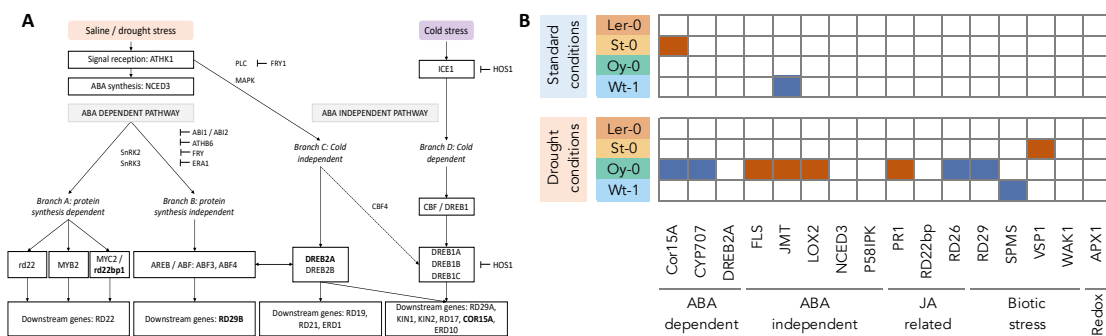
the corresponding drought-evolved TuMV lineages was compared to the response of plants infected with the standard-evolved viral lineages. (A) Number of DEGs obtained for each accession. Over-expressed genes are represented by white bars and under-expressed genes by black bars. (B) Over- (left) or under-expressed (right) DEGs shared between different accessions. (C) Gene ontology analysis for DEGs between drought-evolved viruses and standard-evolved ones for each one of the accessions (columns). Circle size represents the level of enrichment and color indicate adjusted  $P$  values.

Focusing in biological functions, no significant functional enrichment has been observed for G1 accessions, but for both accessions there is a reduction in DEGs involved in nucleocytoplasmic transport. It has been described that the disruption of genes involved in nuclear transport leads to an increase in drought tolerance (Yang et al., 2017). In the case of G2 accessions, the number of enriched and depleted biological categories were higher than in the G1 accessions. However, no obvious similarities in the pattern of enrichment between the two G2 accessions was observed. Interestingly, Wt-1 shows an enrichment in defense responses to virus infection, which may explain why two lineages were extinctic early on the evolution experiment.

To further evaluate how each accession responded to virus infection and drought, the expression of a set of key genes in stress regulation (Fig. 22A) were quantified in the combination of all environmental and virus evolution conditions. Comparison of the gene expression in plants infected with standard- and drought-evolved viruses showed that most of the differential expression happens in the drought environment. Even in these stressful conditions, the number of genes differentially expressed depends on the plant accession that the viruses were evolved in (Fig. 22B), with viruses evolved in Oy-0

showing the most extreme differences, specially down-expressing markers involved in ABA-independent responses (Fig. 22B).

These observations suggest that virus adaptation under drought conditions results in a differential change in the local hosts transcriptome. Previous work has shown how the degree of adaptation of a potyvirus differentially affects the transcriptome of infected plants (Agudelo-Romero et al., 2008). It also likely that drought- and standard-adapted viruses alter gene expression by manipulating certain methylation patterns in their host, as recently observed in TuMV lineages *naïve* and well adapted to *A. thaliana* (Corrêa et al., 2020), though this hypothesis remains to be tested here.



**Fig. 22.** (A) Schematic representation of *A. thaliana* network regulating the response to drought stress; in bold some of the genes whose expression was evaluated. (B) Comparison of the  $2^{-\Delta\Delta C_T}$  values of plants infected with standard- or drought-evolved viruses. Significant differences are marked in blue when the levels are significantly higher in plants infected with standard-evolved viruses and in orange when infected by drought-evolved viruses (pairwise *post hoc* Bonferroni tests in the GLM model described in Eq. 7; in all cases  $P \leq 0.040$ ). The accessions and the conditions where the sample was taken from are

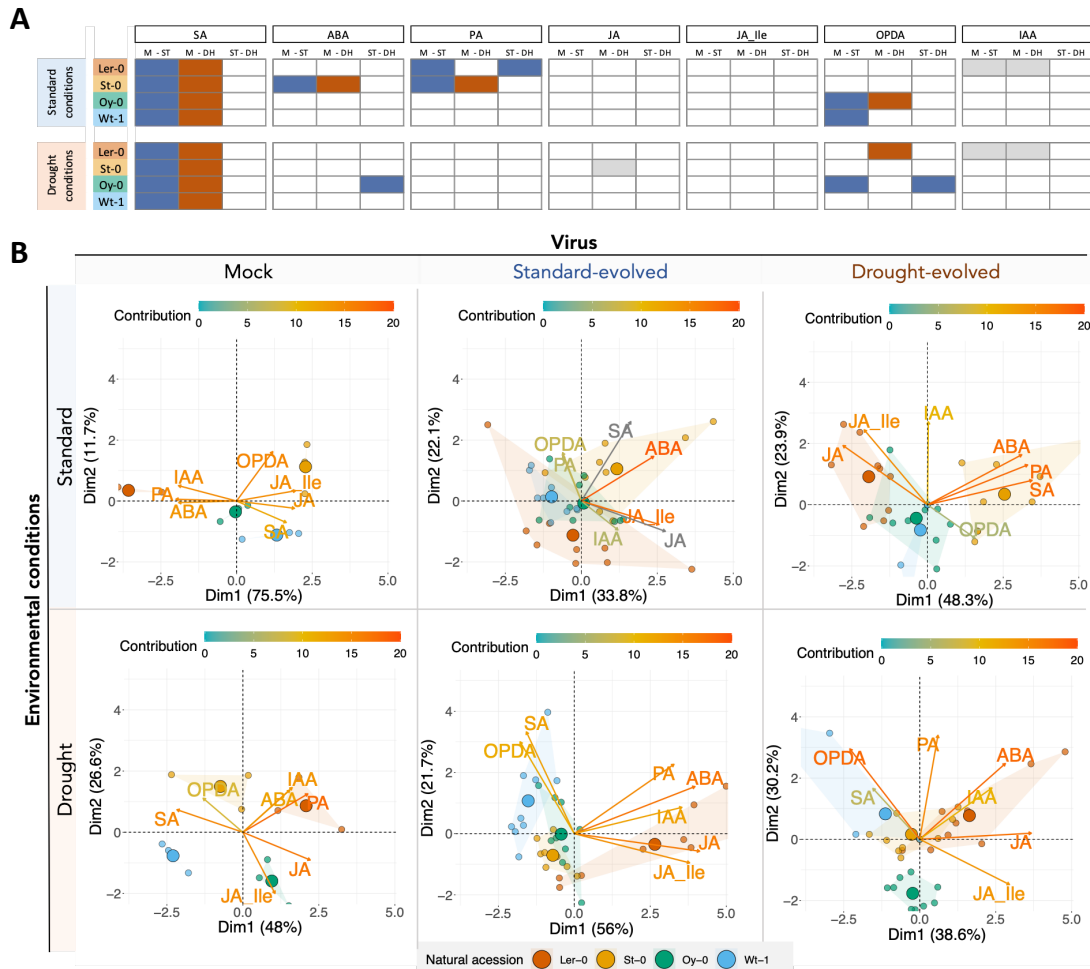
indicated in the left. Participation of the measured genes in particular responses to stress are indicated under the table.

#### **IV 3.4 Differences in Hormones Profiles.**

Plant response to stresses depends on the fine tuning among different phytohormones. We have studied the hormonal levels of plants in both environmental conditions, paying attention to differences among non-inoculated plants and plants infected with the drought- and standard-evolved viruses (Fig. 23). In both standard and drought conditions the levels of salicylic acid (SA) were significantly higher in infected plants (regardless in which conditions the virus was evolved) than in non-infected plants (Fig. 23A). This increase is expected as SA is a key component in defense signaling, inducing the expression of many defense-related genes (Delaney et al., 1994). However, SA not only plays a role in plant defense but also in plant growth regulation and responses to abiotic stresses (Miura and Tada, 2014). Xu et al. (2008) found high SA concentrations in plants infected with brome mosaic virus (BMV) and cucumber mosaic virus (CMV) although it could not be unambiguously associated with the improved drought tolerance provided by the infection (Prasch and Sonnewald, 2013). Aguilar et al. (2017) using SA-deficient transgenic lines observed that SA has a role in the tolerance provided by the virus infection. The observed increase in SA levels in infected plants was similar in plants infected with viruses evolved in standard or drought conditions (Fig. 23A). Therefore, the enhanced tolerance caused by drought-evolved viruses cannot be solely explained by SA levels, suggesting that other hormones could be also involved. Therefore, abscisic acid (ABA), phaseic acid (PA), jasmonic acid (JA), jasmonoyl isoleucine (JA-Ile), oxo-phytodienoic acid (OPDA), and indole-3 acetic acid (IAA; the main auxin) levels were also quantified. In standard growth conditions, the only significant difference in the hormonal levels between plants infected

with standard- and drought-evolved viruses was for the *Ler-0* accession, where the level of PA is significantly higher in plants infected with standard evolved-viruses (Fig. 23A). In drought conditions, the differences are significant for viruses evolved in *St-0*, with plants infected with the standard-evolved viruses having higher levels of ABA and PA than drought-evolved ones (Fig. 23A). Viruses evolved under drought in *Oy-0* showed a depletion in ABA and OPDA compared to viruses evolved in standard conditions (Fig. 23A).

A principal component analysis (Fig. 23B) shows that the hormonal profile of the four accessions changes in response to the infection status and the growing conditions. Interestingly, ABA shows an expression profile in non-infected plants grown in standard conditions (over the abscissa of the upper-left quadrant) that markedly differs from the pattern shown in non-infected plants grown in drought conditions, which actually is similar to the pattern shown by all infected plants, regardless the virus type (in all cases lying in the upper-right quadrant). SA clearly distinguishes between infected plants grown in standard (upper-right) and drought conditions (upper-left) (Fig. 23B). JA and JA-Ile also show a highly correlated pattern while the vectors lie close to the abscissa of the upper-right quadrant in non-infected plants grown in standard conditions, they both move to the upper-left quadrant in plants infected with drought-evolved viruses kept in standard conditions and move to the lower-right quadrant in all other situations (Fig. 23B).



**Fig. 23.** Quantification of stress-related hormones. (A) Comparison of the hormone profiles between non-infected plants (M), plants infected with standard- and drought-evolved viruses. Significant (pairwise *post hoc* Bonferroni tests the GLM described by Eq. 7; in all cases  $P \leq 0.039$ ) differences in the comparison are marked in color: blue when the levels are significantly higher in samples from plants infected with standard-evolved viruses, orange for plants infected with drought-evolved viruses and grey for non-infected plants. The accessions and the conditions where the sample was taken are indicated in the left. (B) Principal component analysis of the quantified hormones. In all cases, the first two components explain more than 55% of observed variability.

### IV 3.5 Changes in the Nature of Host-Virus Interactions.

We studied the survival of each accession to drought conditions when not infected or when infected with standard- or drought-evolved viruses (Fig. 24A). *Ler-0* showed almost no survival regardless of their infection status, with no significant differences in mean probability of survival between non-infected plants and plants infected with the standard-evolved viruses ( $P = 1.000$ ) or plants infected with the drought-evolved viral lineages ( $P = 1.000$ ). For the rest of the accessions, plants infected with the standard-evolved viruses had a higher mean survival probability to drought than non-infected plants, though the differences were not statistically significant (Fig. 24A;  $P \geq 0.202$  in all cases). In sharp contrast, the comparison of mean drought survival probabilities of plants infected with drought-evolved viruses, showed that drought tolerance was significantly higher than in non-infected plants (Fig. 24A;  $P \leq 0.023$  in all cases). Hence, we conclude that viruses can adapt to promote host tolerance to the environmental perturbations. This may lead to a transition into a mutualistic relationship between the virus and the host as both of them benefit from the infection: the virus is able to replicate and spread while the infected host acquires a physiological and/or morphological change that promotes its survival in the adverse environment (Hily et al., 2016).

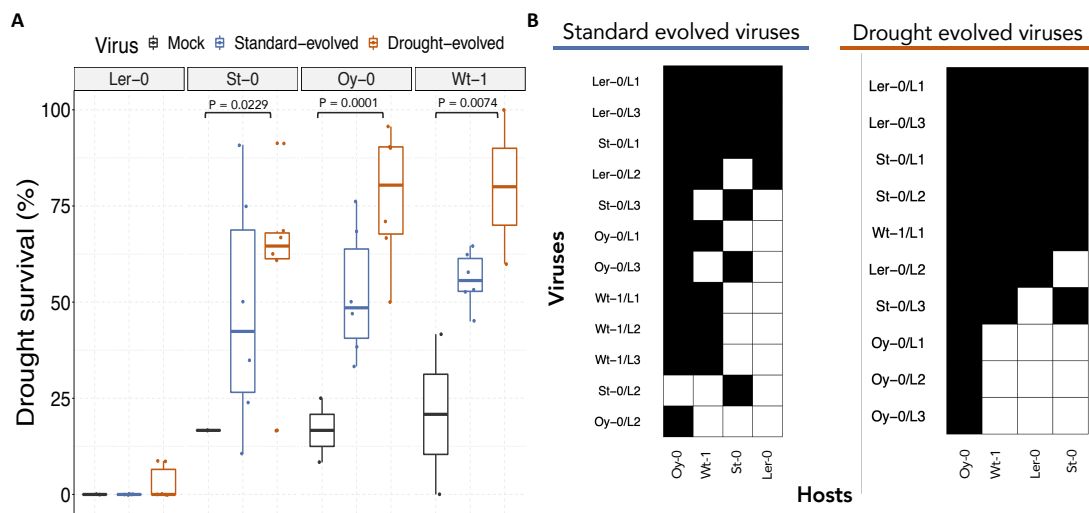
To analyze the specificity of adaptation of each evolved TuMV lineage, we inoculated the 22 evolved viruses into each one of the four accessions and their performance was evaluated using *AUDPS*. With this data, we built up two infection matrices (Fig. 24B), one for plants inoculated with standard-evolved and another with drought-evolved viruses. In each matrix, black squares represent host-virus combinations in which *AUDPS* was equal or greater than the value observed for the viral lineage in its corresponding local host. Therefore, the upper rows correspond to more generalist lineages while the lower ones correspond to more specialist ones. In general, viruses

evolved in accessions from G1 (*Ler-0* and *St-0*) are more generalist than viruses evolved in accessions from G2 (*Oy-0* and *Wt-1* lineages). Likewise, plants from G2 are more susceptible to infection than those from G1. This indicates that accessions which are more permissive to infection gave rise to less pathogenic viruses regardless the watering status, while the more restrictive accessions selected for viruses with greater pathogenicity. Similar results have been previously reported for potyvirus - *A. thaliana* pathosystems (Hillung et al., 2014; González et al., 2019).

Next, we sought to explore if environmental stress would affect the specificity of adaptation. To quantify the degree of specialization, we calculated the partner diversity  $d'$  index (Blüthgen et al., 2006). For the infection matrix estimated for standard-evolved viruses the value was 372-fold higher ( $d' = 0.037$ ) than in the matrix estimated for drought-evolved viruses ( $d' = 0.0001$ ). This large difference indicates that a severe environmental stress favored the evolution of more generalist viruses that provide higher tolerance to drought to a diverse array of host accessions.

Finally, we evaluated the nestedness and modularity of the two matrices. The infection matrix estimated for the standard-evolved viruses shows a significant  $T$ -nestedness (Bascompte et al., 2003) (Fig. 24B, left;  $T = 30.441$ ,  $P = 0.029$ ), while the matrix estimated for the drought-evolved viruses did not show significant nestedness (Fig. 24B, right;  $T = 18.506$ ,  $P = 0.053$ ). This suggests that virus evolution in standard conditions selects for a gene-for-gene interaction mechanism in which more susceptible hosts select for more specialized viruses while more resistant viruses select for more generalist viruses. However, under drought conditions this highly specific mechanism has been overcome, suggesting that it may limit the potential advantage provided by the

infection. We also studied the modularity of the infection matrices, as the presence of modules suggest that common selective constraints are imposed by different hosts (*e.g.*, G1 and G2) and similar evolutionary solutions are found by viruses. Both matrices show significant  $Q$ -modularity (Newman, 2006) ( $P = 0.019$  for standard-evolved viruses and  $P < 0.001$  for the drought-evolved ones), with the modularity observed in the matrix of the standard-evolved viruses being 1.735 times larger than in the matrix of the drought-evolved ones. This difference suggests that similarity in constraints among accessions represent a case of soft selection and contribute less to virus adaptation than the hard selection created by drought conditions.

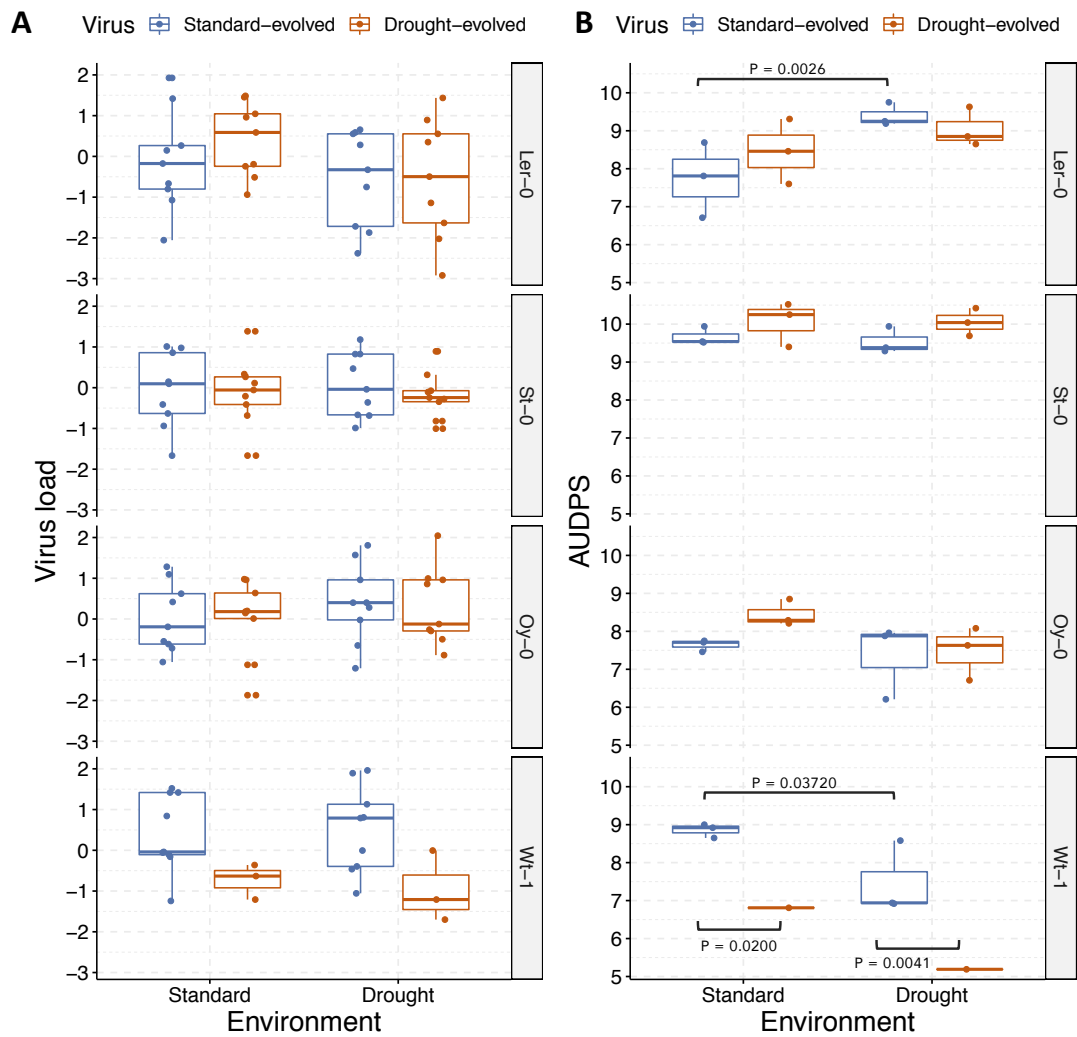


**Fig. 24.** Host-virus interactions. (A) Host survival in severe drought stress in different accessions. Comparison between non-inoculated plants (gray), plants inoculated with the standard- (blue) and with the drought-evolved (orange) viruses. Significant differences are marked with brackets and the  $P$  values are indicated (pairwise *post hoc* Bonferroni tests in the GLM described in Eq. 8). (B) Packed infection matrices in standard conditions for standard- and drought-evolved viruses. Viruses used as inocula are ordered (from the most generalist to the most specialists) in the rows and the



different hosts (from the most permissive to the less one) in the columns. Black squares represent virus-host combinations in which *AUDPS* was equal or greater than the value observed for the corresponding viral lineage in its corresponding local host.

Next, we inoculated all of the viral lineages in their corresponding local accessions in both standard and drought conditions. We found no changes in the viral load of both viruses in all accessions (Fig. 25A). Despite not being able to observe significant differences in the virus accumulation, viruses evolved in Wt-1 have a lower viral load that results in a significant reduction of *AUDPS* (Fig. 25B). The drought-evolved viruses performed worse in Wt-1 accessions than the standard-evolved viruses, an observation that contributes to better understand why two lineages of Wt-1 drought-evolved viruses ended up extinct early in the evolution experiment. Previously, Aguilar et al. observed that the drought-tolerance induced by PVX and PPV in *N. benthamiana* and *A. thaliana* (Aguilar et al., 2017) was caused by the over-expression of a virulence factor. In contrast, our results suggest that the enhanced host drought tolerance triggered by drought-evolved viruses does not depend on a higher virulence.



**Fig. 25.** Two fitness-related traits quantified in standard and drought conditions; blue color for standard- and orange for drought-evolved viruses. Significant differences are marked with brackets and the  $P$  values are indicated (pairwise *post hoc* Bonferroni tests in the GLM described in Eq. 7). (A) Z-scores of viral loads quantified as copies of CP RNA per ng of total RNA (the presence of this protein ensures that the whole virus genome was transcribed and no defective particles were quantified). (B) Disease progression measured as *AUDPS*.

#### **IV 4. DISCUSSION**

The environment where a virus evolves shapes the virus-host interaction. In general, plant viruses can adapt to extreme drought conditions but in certain host genotypes this process can be harder. In this refractory hosts, viral populations might be driven to extinction or will reach a lower fitness than they would in standard conditions, as observed in the two Wt-1 lineages that were extinct. Nevertheless, in other host genotypes plant viruses adapted and increased their fitness equally well regardless of the watering conditions. The environment also influences the mechanisms of selection in virus evolution: (i) the evolutionary solution reached by viruses evolved in drought conditions did not match a gene-for-gene interaction mechanism, (ii) drought-evolved viruses tend to be less specialized and (iii) the selective constraints imposed by the host are more diverse under drought conditions.

In our experimental conditions, the drought environment selected for TuMV lineages that confer enhanced and not accession-specific drought tolerance. Therefore, under conditions of drought stress, infected plants will better tolerate water deficit and, consequently, virus replication and transmission will be increased. The underlying mechanism that promotes drought tolerance seems to be specific for each accession. Hosts whose response to infection was similar also had similar responses to drought during the infection but, even within groups, each accession had a particular response. Differences in the host response are likely triggered by adaptative substitutions in the viral VPg protein. This highly multifunctional protein accumulated mutations in all lineages, but mutations found in drought-evolved lineages were predicted to be more functionally disruptive than the ones fixed in lineages evolved in standard conditions.

The fact that viruses evolved in drought promoted a higher rate of plant survival demonstrates how virus-host interactions are dependent on the environment and their natural history. It was observed before that viral infection can confer drought tolerance to their plant host but this is the first study (i) exploring how abiotic stresses shape the evolution of a host-virus interaction and (ii) showing that virus-induced tolerance is a selectable trait encoded in the viral genome. In our study we have observed that the response to virus infection in drought conditions is diverse within the same species, suggesting that the mechanisms used by viruses to induce drought tolerance are not universal and different mechanisms could be activated depending on the virus and the host genotypes. Xu et al. (2008) observed that drought tolerance improved with virus infection and found an increase in several osmoprotectants and antioxidants and that changes in the metabolite profiles were different depending on the pathosystem (Prasch and Sonnewald, 2013). As an example: trehalose, putrescine and SA levels were increased in virus-infected plants under water deficit conditions but proline, ascorbic acid and sucrose were increased only in BMV-infected rice while galactose, maltose and anthocyanins were only increased in CMV-infected beet. Aguilar et al. (2017) found that hormone levels and metabolite profiles also varied among plants under drought-conditions depending on the virus infecting them. Gorovits et al. (2019) found that tomato plants infected with tomato yellow leaf curl virus had tolerance to several abiotic stresses. This tolerance was found to be achieved by the viral repression of the ubiquitin 26S proteasome degradation and heat shock transcription factors. The variety in the mechanisms found in different pathosystems parallels the diversity we found within *A. thaliana* accessions.

Bergès et al. (2020) illustrated a high level of variability in the response to virus infection and drought within the same species. They studied the response of multiple *A. thaliana* accessions to CaMV infection in drought conditions. They found that under water-stress symptom appearance and rate of systemic spread accession-dependent. Interestingly they found that most of the studied accessions had a bigger survival rate during infection when they were cultivated in drought conditions compared to well-watered conditions. The beneficial virus-host interaction under drought conditions may also expand into other organisms that interact with the pathosystem, such as viral vectors. For example, in the pathosystem wheat - barley yellow dwarf virus it was shown that drought and virus infection enhance the performance of the aphid vector *Rhopalosiphum padi* (Davis et al., 2015).

As a final take-home message, here we have shown that under environmental perturbations, virus-host interactions can evolve from pathogenic to mutualistic in a relatively short evolutionary time, being the extent of such switch host genotype-dependent and via different changes underlying regulatory and signaling mechanisms.

## GENERAL DISCUSSION

---

This work used plant-virus pathosystem to study evolutionary processes through virus experimental evolution. To study them, viruses from the *Potyvirus* genus have been studied in different plant hosts: *Nicotiana* plants were used to generate virus stocks and to study TEV, while *C. quinoa* was used for virus quantification and *A. thaliana* for TuMV studies. Plants are a versatile, low-cost host for virus evolution studies. Some are well established model organisms, for example *A. thaliana*. This small plant allows us to work with a large number of individuals and has plenty resources available (methods, mutants seeds, genetic information, different genotypes, ...). About the different genotypes of *A. thaliana* a special discussion is necessary. In Chapter II, the term “ecotype” was used to refer to *A. thaliana* sampled from different geographic origins. In the following chapter the term used is “natural accessions”, the term the *A. thaliana* community has adopted. This change makes sense as the term ecotype implies that a line has a unique ecology and is adapted to specific environments, as opposed to differing only in genotype from other varieties. Furthermore, the short life cycle of *A. thaliana* allows multiple evolutive passages in a relative short amount of time. Thanks to this it was possible to perform multiple evolutive experiments during the duration of the predoctoral contract, being able to evaluate virus adaptation at different levels.

The potyvirus-plant is a useful and versatile pathosystem that allows us to explore and empirically test diverse evolutionary questions. Viruses are ideal entities for testing evolutionary hypothesis and perform experimental evolution as their evolutionary processes are in line with other living entities. Potyviruses were chosen for this study since (1) they are important crop pathogens, so there are well known viruses in which

new discoveries would be of interest and also easily translated to agricultural systems, (2) independently of any evident practical interest, studies performed in a plant-virus pathosystem are of high interest to increase our knowledge in virology and evolution, as the findings in this system could be extrapolated to others, (3) potyviruses are RNA viruses so they have high replication and mutation rates that allows us to perform evolutionary experiments in a short time and their small genome make it possible to easily study genomic changes, and (4) *Potyvirus* infect multicellular complex organism without rising any bioethical concern; so large number of individuals can be studied without any problem.

Measuring virus' degree of adaptation can be challenging as different fitness-related traits can be used. For example, if we would like to tell which countries are the best adapted ones we may classify them by their population (this will be more or less equivalent to virus load). So, by population (in 2020) the top five countries would be China, India, USA, Indonesia, and Pakistan (in that order). Some will say that this is not the most appropriate way of measuring a country adaptation, as large populations will not do much without economic power (this will be more or less equivalent to the infectious units that virus generate). By economic power the most adapted countries would be USA, China, Japan, Germany, and UK. But this classification may be tricky too, as by this parameters Spain is at the same level than Russia but in case of a clash between these two countries all Spaniards would be drinking vodka within two weeks. So, you could classify countries by military power (this will be more or less equivalent to viral symptomatology or disease progression) and the list will change to USA, Russia, China, India, and France. Even other less popular factors, as happiness, could be used for measuring adaptation. So, measuring the degree of adaptation is complicated and may depend on the perspective

used during the evaluation. In this work different approaches have been used: viral load (measuring the number of *CP* gene copies and therefore the number of complete genomes), infectious units (measuring the infection foci in a given viral dilutions) and the disease progression (calculating *AUDPS*). Luckily, in potyvirus there is a direct correlation between viral load, infectious units and disease progression. Nevertheless, this correlation may be not happening in all hosts or in all environments so it is important to take this into consideration when measuring virus adaptation. With this in mind, for example, both the viral load and the disease progression were evaluated when testing the adaptation under drought conditions.

Adaptation may be evaluated at different levels. In this work we evaluated three. First, the evolutive constrains of a viral protein were studied by performing a deep mutational scanning on the 6K2 gene of TEV. Second, TuMV degree of adaptation was evaluated accounting for the genetic structure of the host population, for susceptibility genes, the virus infects. Third, we tested TuMV adaptation to different environments: one where the host was in standard conditions and other where it was suffering an abiotic stress, such as drought. More factors affecting virus adaptation are yet to be explored. Virus evolution may depend in multiple factors that have been neglected in virus evolutionary biology or are yet to be discovered.





## GENERAL CONCLUSIONS

---

The research conducted for this PhD has led us to the following conclusions:

### Chapter I

I A. Different domains of the 6K2 protein are subjected to different selective pressures: N-terminal domain is evolving under purifying selection while the transmembrane domains are more evolvable, the C-terminal domain being able to accommodate mutations.

I B. *In silico* prediction of the impact of the mutations are negatively associated with their real effect: the major impact mutations correspond with no infective viruses meanwhile the weaker impact mutations result in WT-virus like infections.

I C. The identified beneficial mutations were associated with weaker symptoms and slower disease progression, suggesting a negative tradeoff between within-host replicative fitness and severity of symptoms.

### Chapter II

II A. Genetic diversity in susceptibility to infection of hosts in an experimental ecosystem influences the evolution of a pathogen, inducing changes in its adaptation to the host and its virulence towards it.

II B. TuMV adapts faster and shows stronger signatures of local specialization in homogeneous than in heterogeneous *A. thaliana* experimental populations.

II C. Viruses evolved in heterogeneous populations were more pathogenic and infectious than viruses evolved in the homogeneous population.

### **Chapter III**

III A. Most viruses have adapted to not harm their host and are essential components of the phytobiome.

III B. Pathogenic viruses can, under adverse abiotic environments, become beneficial for their hosts.

III C. If the host is in an abiotic stressful situations, plant and virus interactions may coevolve from an antagonistic relationship toward a non-antagonistic one (commensal or mutualistic).

### **Chapter IV**

IV A. Viruses evolved in drought conditions promote a higher plant survival under such stressful conditions.

IV B. Hormonal signaling of biotic and abiotic stresses are reprogrammed by the drought-evolved virus. The magnitude and direction of effects are ecotype-dependent.

IV C. This phenomenon suggests than viruses that were pathogenic for their host can evolve towards non-pathogenic relationships with them if the environment endangers the host and the virus can promote host survival.





## SUPPLEMENTARY MATERIAL

---

**Supplementary Table S1.** Pairs of primers used for generating the mutant 6K2 sequences.

### I32S

Primer F 5' GTCTGTGTTAAGTGGTGGTGGATGGATGC 3'

Primer R 5' GCATCCATCCACCACCACTTAACACAGAC 3'

### I32S-G33D

Primer F 5' GTCTGTGTTAAGTGATGGTGGATGGATGC 3'

Primer R 5' GCATCCATCCACCATCACTTAACACAGAC 3'

### G34D

Primer F 5' GTGTTAATTGGTGATGGATGGATGCTTG 3'

Primer R 5' CAAGCATCCATCCATCACCAATTAACAC 3'

### G34V

Primer F 5' GTGTTAATTGGTGTTGGATGGATGCTTG 3'

Primer R 5' CAAGCATCCATCCAACACCAATTAACAC 3'

### G34R/G35A

Primer F 5' GTTAATTGGTCGTGCATGGATGCTTGCAAC 3'

Primer R 5' GTTGCAAGCATCCATGCACGACCAATTAAC 3'

### G34S/A39S (primers to insert the G34S mutation)

Primer F 5' GTTAATTGGTAGTGGATGGATGCTTGC 3'

Primer R 5' GCAAGCATCCATCCACTACCAATTAAC 3'

### G34S/A39S (primers to insert the A39S mutation)

Primer F 5' GGATGGATGCTTTCAACGTACTIONTCAAG 3'

Primer R            5' CTTGAAGTACGTTGAAAGCATCCATCC 3'

#### G35V

Primer F            5' GTTAATTGGTGGTGTATGGATGCTTGC 3'

Primer R            5' GCAAGCATCCATACACCACCAATTAAC 3'

#### A39V

Primer F            5' GGATGGATGCTTGTAACGTACTIONTCAAG 3'

Primer R            5' CTTGAAGTACGTTACAAGCATCCATCC 3'

#### A39E

Primer F            5' GGATGGATGCTTGAAACGTACTIONTCAAG 3'

Primer R            5' CTTGAAGTACGTTTCAAGCATCCATCC 3'

#### D44E

Primer F            5' CGTACTTCAAGGAAAAGTTCAATGAAC 3'

Primer R            5' GTTCATTGAACTTTTCCTTGAAGTACG 3'

#### D44E/F46L

Primer F            5' CGTACTTCAAGGAAAAGTTAAATGAAC 3'

Primer R            5' GTTCATTTAACTTTTCCTTGAAGTACG 3'

---

**Supplementary Table S2.** Results of the univariate MANOVAs. Model equation fitted to each phenotypic variable is the same as for the multivariate analyses (section 2.7).

Source	Variable	SS	Df	MS	F	P	$\eta_p^2$	1 - $\beta$
<i>Intersection</i>	<i>AUDPS</i>	1777.003		1777.003	814.700	< 0.001	0.755	1
	<i>f</i>	44.188	1	44.188	302.233	< 0.001	0.534	1
<i>t</i>	<i>AUDPS</i>	978.630	1	978.630	448.671	< 0.001	0.630	1
	<i>f</i>	11.722	1	11.722	80.175	< 0.001	0.233	1
<i>B</i>	<i>AUDPS</i>	152.964	1	152.964	70.129	< 0.001	0.210	1
	<i>f</i>	5.698	1	5.698	38.970	< 0.001	0.129	1
<i>B</i> × <i>t</i>	<i>AUDPS</i>	0.448	1	0.448	0.206	0.651	0.001	0.074
	<i>f</i>	2.547	1	2.547	17.420	< 0.001	0.062	0.986
<i>D</i>	<i>AUDPS</i>	25.107	1	25.107	11.511	0.001	0.042	0.922
	<i>f</i>	2.083	1	2.083	14.244	< 0.001	0.051	0.964
<i>D</i> × <i>t</i>	<i>AUDPS</i>	9.070	1	9.070	4.158	0.042	0.016	0.529
	<i>f</i>	1.100	1	1.100	7.523	0.007	0.028	0.780
<i>E</i>	<i>AUDPS</i>	175.155	5	35.031	16.061	< 0.001	0.233	1
	<i>f</i>	3.608	5	0.722	4.935	< 0.001	0.085	0.982
<i>E</i> × <i>t</i>	<i>AUDPS</i>	14.695	5	2.939	1.347	0.245	0.025	0.474
	<i>f</i>	1.168	5	0.234	1.598	0.161	0.029	0.553
<i>B</i> × <i>D</i>	<i>AUDPS</i>	4.236	1	4.236	1.942	0.165	0.007	0.284
	<i>f</i>	0.128	1	0.128	0.874	0.351	0.003	0.154
<i>B</i> × <i>D</i> × <i>t</i>	<i>AUDPS</i>	8.599	1	8.599	3.942	0.048	0.015	0.507
	<i>f</i>	0.288	1	0.288	1.970	0.162	0.007	0.288
<i>B</i> × <i>E</i>	<i>AUDPS</i>	15.355	5	3.071	1.408	0.222	0.026	0.494
	<i>f</i>	2.863	5	0.573	3.916	0.002	0.069	0.943
<i>B</i> × <i>E</i> × <i>t</i>	<i>AUDPS</i>	16.678	5	3.336	1.529	0.181	0.028	0.532
	<i>f</i>	1.514	5	0.303	2.072	0.069	0.038	0.685
<i>D</i> × <i>E</i>	<i>AUDPS</i>	3.512	5	0.702	0.322	0.900	0.006	0.131
	<i>f</i>	0.179	5	0.036	0.245	0.942	0.005	0.110
<i>D</i> × <i>E</i> × <i>t</i>	<i>AUDPS</i>	5.351	5	1.070	0.491	0.783	0.009	0.183
	<i>f</i>	0.175	5	0.035	0.239	0.945	0.005	0.108
<i>B</i> × <i>D</i> × <i>E</i>	<i>AUDPS</i>	6.331	5	1.266	0.581	0.715	0.011	0.212
	<i>f</i>	0.349	5	0.070	0.478	0.793	0.009	0.179
<i>B</i> × <i>D</i> × <i>E</i> × <i>t</i>	<i>AUDPS</i>	8.256	5	1.651	0.757	0.582	0.014	0.271
	<i>f</i>	0.520	5	0.104	0.711	0.616	0.013	0.255
Error	<i>AUDPS</i>	575.830	264	2.181				
	<i>f</i>	38.598	264	0.146				



**Supplementary Table S3.** Results of the fitting of the ARIMA(1, 0, 0) model  $Y_t - \rho_1 Y_{t-1} = Y_0 + \beta t + \varepsilon_t$  to the *AUDPS* and *I* time series data. Error represent  $\pm 1$  SEM. See section 2.6 for definition of each model parameter. Statistical significance of each parameter is estimated using a *t*-test (10 d.f.).

Block	Ecotype	Population	$Y_0$	$t$	$P$	$\rho_1$	$t$	$P$	$\beta$	$t$	$P$
<i>AUDP</i>											
<i>S</i>											
<i>P</i>	Col-0	Metapopulation	2.179	2.363	0.040	0.443	1.559	0.150	0.576	4.543	0.001
			$\pm 0.922$		$\pm 0.284$		$\pm 0.127$				
		Well-mixed	2.733	2.663	0.024	0.193	0.582	0.574	0.378	2.596	0.027
			$\pm 1.026$		$\pm 0.332$		$\pm 0.146$				
	Ga-0	Metapopulation	6.656	17.445	<	-0.230	0.750	0.471	0.342	6.286	< 0.001
			$\pm 0.382$		0.001	$\pm 0.307$		$\pm 0.054$			
		Well-mixed	7.115	25.558	<	-0.461	1.596	0.142	0.293	7.317	< 0.001
			$\pm 0.278$		0.001	$\pm 0.289$		$\pm 0.040$			
	Gy-0	Metapopulation	3.189	6.000	<	-0.292	0.924	0.377	0.506	6.637	< 0.001
			$\pm 0.532$		0.001	$\pm 0.316$		$\pm 0.076$			
		Well-mixed	3.233	3.165	0.010	0.253	0.829	0.427	0.401	2.819	0.018
			$\pm 1.021$			$\pm 0.305$		$\pm 0.142$			
Oy-0	Metapopulation	1.652	2.030	0.070	0.408	1.089	0.302	0.465	3.838	0.003	
		$\pm 0.814$			$\pm 0.375$		$\pm 0.121$				
	Well-mixed	2.054	5.353	<	-0.111	0.269	0.793	0.376	6.639	< 0.001	
		$\pm 0.384$		0.001	$\pm 0.413$		$\pm 0.057$				
Ta-0	Metapopulation	2.326	2.472	0.033	0.350	1.129	0.285	0.807	6.123	< 0.001	
		$\pm 0.941$			$\pm 0.310$		$\pm 0.132$				
	Well-mixed	3.361	2.014	0.072	0.431	1.528	0.157	0.497	2.175	0.055	
		$\pm 1.669$			$\pm 0.282$		$\pm 0.229$				
Wt-1	Metapopulation	1.342	4.290	0.002	-0.307	1.013	0.335	0.703	15.706	< 0.001	
		$\pm 0.313$			$\pm 0.303$		$\pm 0.045$				
	Well-mixed	1.951	1.897	0.087	0.343	1.063	0.313	0.505	3.470	0.006	
		$\pm 1.029$			$\pm 0.322$		$\pm 0.145$				
<i>R</i>	Col-0	Metapopulation	3.539	6.402	<	-0.379	1.311	0.219	0.544	6.882	< 0.001
			$\pm 0.553$		0.001	$\pm 0.289$		$\pm 0.079$			
		Well-mixed	5.854	5.314	<	0.605	2.353	0.040	0.474	3.175	0.010
			$\pm 1.102$		0.001	$\pm 0.257$		$\pm 0.149$			

Ga-0	Metapopulation	6.806	6.375	<	-0.144	0.459	0.656	0.403	2.649	0.024
		±1.068		0.001	±0.315			±0.152		
	Well-mixed	8.811	13.467	<	0.100	0.306	0.766	0.390	4.208	0.002
		±0.654		0.001	±0.328			±0.093		
Gy-0	Metapopulation	4.472	5.809	<	-0.342	1.147	0.278	0.691±	6.300	< 0.001
		±0.770		0.001	±0.298			0.110		
	Well-mixed	7.015	8.226	<	0.366	0.989	0.346	0.500	3.982	0.003
		±0.853		0.001	±0.370			±0.126		
Oy-0	Metapopulation	3.966	4.140	0.002	-0.201	0.627	0.544	0.481	3.511	0.006
		±0.958			±0.320			±0.137		
	Well-mixed	5.499	10.489	<	-0.312	0.958	0.361	0.434	5.757	< 0.001
		±0.524		0.001	±0.326			±0.076		
Ta-0	Metapopulation	5.923	6.366	<	-0.134	0.413	0.688	0.410	3.087	0.012
		±0.930		0.001	±0.324			±0.133		
	Well-mixed	7.716	9.437	<	0.388	1.214	0.252	0.326	2.819	0.018
		±0.818		0.001	±0.320			±0.116		
Wt-1	Metapopulation	5.170	5.094	<	-0.176	0.540	0.601	0.342	2.352	0.041
		±1.015		0.001	±0.326			±0.145		
	Well-mixed	4.986	4.617	0.001	0.136	0.424	0.680	0.549	3.604	0.005
		±			±0.320			±0.152		
		1.080								

*I*

<i>P</i>	Col-0	Metapopulation	0.754	12.359	<	-0.191	0.616	0.552	0.012	2.453	0.034
			±0.061		0.001	±0.310			±0.009		
		Well-mixed	0.902	9.583	<	0.215	0.669	0.518	-0.001	0.073	0.943
			±0.094		0.001	±0.321			±0.013		
Ga-0	Metapopulation	0.891	33.100	<	-0.339	1.142	0.280	0.011	2.829	0.018	
		±0.027		0.001	±0.297			±0.004			
	Well-mixed	0.952	39.516	<	-0.205	0.659	0.524	0.006	1.678	0.124	
		±0.024		0.001	±0.311			±0.003			
Gy-0	Metapopulation	0.617	14.671	<	-0.482	1.750	0.111	0.037	6.110	< 0.001	
		±0.042		0.001	±0.275			±0.006			
	Well-mixed	0.625	7.330	<	0.171	0.542	0.600	0.026	2.150	0.057	
		±0.085		0.001	±0.316			±0.012			
Oy-0	Metapopulation	0.566	8.007	<	-0.010	0.029	0.977	0.039	3.865	0.003	
		±0.071		0.001	±0.339			±0.101			
	Well-mixed	0.610	5.204	<	0.326	0.850	0.415	0.028	1.617	0.137	
		±0.117		0.001	±0.383			±0.017			

Ta-0	Metapopulation	0.283	1.243	0.242	0.830	4.013	0.002	0.071	2.265	0.047
		$\pm 0.228$			$\pm 0.207$			$\pm 0.032$		
	Well-mixed	0.595	4.078	0.002	0.204	0.660	0.524	0.035	1.718	0.117
		$\pm 0.146$			$\pm 0.309$			$\pm 0.020$		
Wt-1	Metapopulation	0.586	10.665	<	-0.376	1.264	0.235	0.042	5.366	< 0.001
		$\pm 0.055$		0.001	$\pm 0.298$			$\pm 0.008$		
	Well-mixed	0.652	5.386	<	0.157	0.501	0.627	0.023	1.376	0.199
		$\pm 0.121$		0.001	$\pm 0.314$			$\pm 0.017$		
$\mathcal{R}$	Col-0	0.682	11.109	<	-0.496	1.839	0.096	0.022	2.480	0.033
		$\pm 0.061$		0.001	$\pm 0.270$			$\pm 0.009$		
	Well-mixed	0.905	17.584	<	0.212	0.684	0.510	0.010	1.353	0.206
		$\pm 0.051$		0.001	$\pm 0.310$			$\pm 0.007$		
Ga-0	Metapopulation	0.863	12.923	<	-0.245	0.801	0.442	0.010	1.012	0.335
		$\pm 0.067$		0.001	$\pm 0.306$			$\pm 0.010$		
	Well-mixed	1.000			0.922			0.000		
		$\pm 0.000$			$\pm 0.000$			$\pm 0.000$		
Gy-0	Metapopulation	0.707	8.659	<	-0.248	0.811	0.436	0.029	2.483	0.032
		$\pm 0.082$		0.001	$\pm 0.306$			$\pm 0.012$		
	Well-mixed	0.965	62.739	<	-0.326	1.101	0.297	0.003	1.403	0.191
		$\pm 0.015$		0.001	$\pm 0.296$			$\pm 0.002$		
Oy-0	Metapopulation	0.841	15.712	<	-0.034	0.108	0.916	0.012	1.618	0.137
		$\pm 0.054$		0.001	$\pm 0.316$			$\pm 0.008$		
	Well-mixed	0.932	28.631	<	-0.038	0.119	0.907	0.006	1.332	0.212
		$\pm 0.033$		0.001	$\pm 0.316$			$\pm 0.005$		
Ta-0	Metapopulation	0.845	11.210	<	-0.136	0.432	0.675	0.008	0.703	0.498
		$\pm 0.075$		0.001	$\pm 0.316$			$\pm 0.011$		
	Well-mixed	0.965	62.739	<	-0.326	1.101	0.297	0.003	1.403	0.191
		$\pm 0.015$		0.001	$\pm 0.296$			$\pm 0.002$		
Wt-1	Metapopulation	0.816	9.096	<	0.060	0.188	0.855	0.002	0.192	0.852
		$\pm 0.090$		0.001	$\pm 0.317$			$\pm 0.013$		
	Well-mixed	0.873	11.154	<	0.389	1.272	0.232	0.009	0.836	0.423
		$\pm 0.078$		0.001	$\pm 0.306$			$\pm 0.011$		

## REFERENCES

---

- Abedon, S. T., Murray, K. L. 2013. Archaeal viruses, not archaeal phages: an archaeological dig. *Archaea*. 2013, 1-10.
- Adams, S. et al. 2018. Circadian control of abscisic acid biosynthesis and signaling pathways revealed by genome-wide analysis of *LHY* binding targets. *New Phytologist*. 220, 893-907.
- Afgan, E. et al. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*. 46, 537-544.
- Agbeci, M., Grangeon, R., Nelson, R.S., Zhang, H., Laliberté, J. F. 2013. Contribution of host intracellular transport machineries to intracellular movement of Turnip mosaic virus. *PLoS Pathogens*. 9, e1003686.
- Agrios, G. N. 2005. *Plant Pathology*, fifth ed. Academic Press, San Diego.
- Agudelo-Romero, P., Carbonell, P., Pérez-Amador, M. A., Elena, S. F. 2008. Virus adaptation by manipulation of host's gene expression. *PLoS ONE*. 3, e2397.
- Aguilar, E. et al. 2015. Effects of elevated CO<sub>2</sub> and temperature on pathogenicity determinants and virulence of potato virus X/potyvirus-associated synergism. *Molecular Plant-Microbe Interactions*. 28, 1364-1373.
- Aguilar, E. et al. 2017. Virulence determines beneficial trade-offs in the response of virus-infected plants to drought via induction of salicylic acid: trade-offs in virus-infected plants to drought. *Plant, Cell and Environment*. 40, 2909-2930.
- Aguirre, J., Manrubia, S. C. 2008. Effects of spatial competition on the diversity of a quasispecies. *Physical Review Letters*. 100, 038106.

- Ahn, S., Vikalo, H. 2018. aBayesQR: a Bayesian method for reconstruction of viral populations characterized by low diversity. *Journal of Computational Biology*. 25, 637-648.
- Ahnert, S. E. 2017. Structural properties of genotype-phenotype maps. *Journal of the Royal Society Interface*. 14, 20170275.
- Alazem, M., Lin, N. S. 2015. Roles of plant hormones in the regulation of host-virus interactions: plant hormone-virus interrelations. *Molecular Plant Pathology*. 16, 529-540.
- Altizer, S. 2006. Seasonality and the dynamics of infectious diseases. *Ecology Letters*. 9, 467-484.
- Andrews, C. J., Paliwal, Y. C. 1983. The influence of preinfection cold hardening and disease development period on the interaction between barley yellow dwarf virus and cold stress tolerance in wheat. *Canadian Journal of Plant Pathology*. 67, 1935-1940.
- Andrews, C. J., Paliwal, Y. C. 1986. Effects of barley yellow dwarf virus-infection and low-temperature flooding on cold stress tolerance of winter cereals. *Canadian Journal of Plant Pathology*. 8, 311-316.
- Anttila, J. et al. 2015. Environmental variation generates environmental opportunist pathogen outbreaks. *PLoS ONE*. 10, e0145511.
- Aprile, A. et al. 2013. Different stress responsive strategies to drought and heat in two durum wheat cultivars with contrasting water use efficiency. *BMC Genomics*. 14, 821.
- Arnholdt-Schmitt, B. 2004. Stress-induced cell reprogramming. A role for global genome regulation?. *Plant Physiology*. 136, 2579-2586.
- Atkinson, N. J., Urwin, P. E. 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. *Journal of Experimental Botany*. 63, 3523-3543.

- Atkinson, N. J., Lilley, C. J., Urwin, P.E. 2013. Identification of genes involved in the response of *Arabidopsis* to simultaneous biotic and abiotic stresses. *Plant Physiology*. 162, 2028-2041.
- Aucique-Pérez, C. E., de Menezes Silva, P. E., Moreira, W.R., DaMatta, F. M., Rodrigues, F. A. 2017. Photosynthesis impairments and excitation energy dissipation on wheat plants supplied with silicon and infected with *Pyricularia oryzae*. *Plant Physiology and Biochemistry*. 121, 196-205.
- Auld, J. R., Agrawal, A. A., Relyea, R. A. 2010. Re-evaluating the cost and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society B*. 277, 503-511.
- Baltimore, D. 1971. Expression of animal virus genomes. *Bacteriology Reviews*. 35, 235-241.
- Bascompte, J., Jordano, P., Melián, C. J., Olesen, J. M. 2003. The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences of the USA*. 100, 9383-9387.
- Bass, D., Stentiford, G. D., Wang, H. C., Koskella, B., Tyler, C. R. 2019. The pathobiome in animal and plant disease. *Trends in Ecology and Evolution*. 34, 996-1008.
- Basu, A., Chowdhury, S., Chaudhuri, T. R., Kundu, S. 2016. Differential behavior of sheath blight pathogen *Rhizoctonia solani* in tolerant and susceptible rice varieties before and during infection. *Plant Pathology*. 65, 1333-1346.
- Baylis, M. 2017. Potential impact of climate change on emerging vector-borne and other infections in the UK. *Environmental Health*. 16, 112.
- Bedoyam L. C, Daròs, J. A. 2010. Stability of tobacco etch virus infectious clones in plasmid vectors. *Virus Research*. 149, 234-240.
- Bell, P. 2001. Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus? *Journal of Molecular Evolution*. 53, 251-256.

- Belliure, B., Janssen, A., Maris, P. C., Peters, D., Sabelis, M. W. 2004. Herbivore arthropods benefit from vectoring plant viruses: thrips benefit from vectoring TSWV. *Ecology Letters*. 8, 70-79.
- Belshaw, R., Gardner, A., Rambaut, A., Pybus, O. G. 2008. Pacing a small cage: mutation and RNA viruses. *Trends in Ecology and Evolution*. 23, 188-193.
- Bennett, A. F., Lenski, R. E., Mittler, J. E. 1992. Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. *Evolution*. 46, 16-30.
- Bergès, S. E. et al. 2020. Natural variation of *Arabidopsis thaliana* responses to cauliflower mosaic virus infection upon water deficit. *PLoS Pathogens*. 16, e1008557.
- Bergès, S.E. et al. 2018. Interactions between drought and plant genotype change epidemiological traits of cauliflower mosaic virus. *Frontiers in Plant Science*. 9, 703.
- Bergruber, T. W., Lion, S., Gandon, S. 2015. Spatial structure, transmission modes and the evolution of viral exploitation strategies. *PLoS Pathogens*. 11, e1004810.
- Bhattacharya, D., Medlin, L. 1998. Algal phylogeny and the origin of land plants. *Plant Physiology*. 116, 9-15.
- Blüthgen, N., Menzel, F., Blüthgen, N. 2006. Measuring specialization in species interaction networks. *BMC Ecology*. 6, 9.
- Boots, M., Sasaki, A. 2002. Parasite-driven extinction in spatially explicit host-parasite systems. *American Naturalist*. 159, 706-713.
- Boots, M., Meador, M. 2007. Local interactions select for lower pathogen infectivity. *Science*. 315, 1284-6.
- Boots, M., Sasaki, A. 1999. Small worlds' and the evolution of virulence: infection occurs locally and at a distance. *Proceedings of the Royal Society B*. 266, 1933-1938.

- Bowman, C. 2019. Plant Virology. Scientific e-Resources.
- Boyes, D. C. 2001. Growth stage-based phenotypic analysis of arabidopsis: a model for high throughput functional genomics in plants. *The Plant Cell*. 13, 1499-510.
- Boyko, A., Kovalchuk, I. 2011. Genetic and epigenetic effects of plant-pathogen interactions: an evolutionary perspective. *Molecular Plant*. 4, 1014-1023.
- Bradshaw, A.D. 1965. Evolutionary significance of phenotypic plasticity in plants. *Advanced Genetics*. 13, 115-155.
- Brockhurst, M. A., Buckling, A., Rainey, P. B. 2006. Spatial heterogeneity and the stability of host-parasite coexistence. *Journal of Evolutionary Biology*. 19, 374-9.
- Browder, L. E. 1985. Parasite host environment specificity in the cereal rusts. *Annual Review of Phytopathology*. 23, 201-222.
- Brown, J. K. M., Tellier, A. 2011. Plant-parasite coevolution: bridging the gap between genetics and ecology. *Annual Review of Phytopathology*. 49, 345-67.
- Bruinsma, J. 2003. World agriculture towards 2015/2030: a FAO perspective. Earthscan Publications Ltd. London.
- Bull, J. J., Meyers, L. A., Lachmann, M. 2005. Quasispecies made simple. *PLoS Computational Biology*. 1, 61.
- Bull, J. J, Sanjuán, R., Wilke, C. O. 2007. Theory of lethal mutagenesis for viruses. *Journal of Virology*. 81, 2930-2939.
- Cabanillas, D. G. et al. 2018. Turnip mosaic virus uses the SNARE protein VT111 in an unconventional route for replication vesicle trafficking. *Plant Cell*. 30, 2594-2615.
- Callaway, R.M. et al. 2002. Positive interactions among alpine plants increase with stress. *Nature*. 417, 844-848.
- Carr, J. P. 2017. Exploring how viruses enhance plants' resilience to drought and the limits to this form of viral payback. *Plant, Cell and Environment*. 40, 2906-2908.



- Carrasco, P., De la Iglesia, F., Elena, S. F. 2007. Distribution of fitness and virulence effects caused by single-nucleotide substitution in tobacco etch virus. *Journal of Virology*. 81, 12979-12984.
- Cervera, H., Lalić J., Elena, S. F. 2016a. Effect of host species on the topography of the fitness landscape for a plant RNA virus. *Journal of Virology*. 90, 10160-10169.
- Cervera, H., Lalić, J., Elena, S. F. 2016b. Efficient escape from local optima in a highly rugged fitness landscape by evolving RNA virus populations. *Proceedings of the Royal Society B*. 283, 20160984.
- Cervera, H., Elena, S.F. 2016. Genetic variation in fitness within a clonal population of a plant RNA virus. *Virus Evolution*. 2, 006.
- Chabas, H. et al. 2018. Evolutionary emergence of infectious diseases in heterogeneous host populations. *PLoS Biology*. 16, 2006738.
- Chen, C. C. et al. 2003. Identification of turnip mosaic virus isolates causing yellow stripe and spot on calla lily. *Plant Disease*. 87, 901-5.
- Cheng, X., Wang, A. 2017. The *Potyvirus* silencing suppressor protein VPg mediates degradation of SGS3 via ubiquitination and autophagy pathways. *Journal of Virology*. 91, e01478-16.
- Chisholm, S. T., Coaker, G., Day, B., Staskawicz, B. J. 2006. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*. 124, 803-814.
- Chung, B. Y. W., Miller, W. A., Atkins, J. F., Firth, A. E. 2008. An overlapping essential gene in the *Potyviridae*. *Proceedings of the National Academy of Sciences of the USA*. 105, 5897-5902.
- Claverie, J. M. 2006. Viruses take center stage in cellular evolution. *Genome Biology*. 7, 110.

- Comins, H. N., Hassell, M. P., May, R. M. 1992. The spatial dynamics of host parasitoid systems. *Journal of Animal Ecology*. 61, 735-48.
- Cornwall, D. H. et al. 2018. Experimental manipulation of population-level mhc diversity controls pathogen virulence evolution in *Mus musculus*. *Journal of Evolutionary Biology*. 31, 314-22.
- Corrêa, R.L. et al. 2020. Viral fitness determines the magnitude of transcriptomic and epigenomic reprogramming of defense responses in plants. *Molecular Biology and Evolution* 37, 1866-1881.
- Cotton, S. et al. 2009. Turnip mosaic virus RNA replication complex vesicles are mobile, align with microfilaments, and are each derived from a single viral genome. *Journal of Virology*. 83, 10460-10471.
- Cowperthwaite, M. C., Bull, J. J., Ancel Meyers, L. 2006. From bad to good: fitness reversals and the ascent of deleterious mutations. *PLoS Computational Biology*. 2, 141.
- Cramer, G. R., Urano, K., Delrot, S., Pezzotti, M., Shinozaki, K. 2011. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biology*. 11, 163.
- Cuevas, J. M., Moya, A., Elena, S. F. 2003. Evolution of RNA virus in spatially structured heterogeneous environments. *Journal of Evolutionary Biology*. 16, 456-66.
- Cui, H., Wang, A. 2016. Plum pox virus 6K1 protein is required for viral replication and targets the viral replication complex at the early stage of infection. *Journal of Virology*. 90, 5119-5131.
- Dáder, B., Then, C., Berthelot, E., Ducouso, M., Ng, J.C.K., Drucker, M. 2017. Insect transmission of plant viruses: multilayered interactions optimize viral propagation. *Insect Sciences*. 24, 929-946.

- Dastogeer, K. M. G., Li, H., Sivasithamparam, K., Jones, M. G. K., Wylie, S. J. 2018. Fungal endophytes and a virus confer drought tolerance to *Nicotiana benthamiana* plants through modulating osmolytes, antioxidant enzymes and expression of host drought responsive genes. *Environmental and Experimental Botany*. 149, 95-108.
- Davis, T. S., Bosque-Pérez, N. A., Foote, N. E., Magney, T., Eigenbrode, S. D. 2015. Environmentally dependent host-pathogen and vector-pathogen interactions in the barley yellow dwarf virus pathosystem. *Journal of Applied Ecology*. 52, 1392-1401.
- de Ronde, D., Butterbach, P., Kormelink, R. 2014. Dominant resistance against plant viruses. *Frontiers in Plant Science*. 5, 307.
- de Visser, J. A. G. M., Krug, J. 2014. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics*. 15, 480-490.
- de Visser, J. A. G. M. et al. 2003. Evolution and detection of genetic robustness. *Evolution* 57, 1959-1972.
- Delaney, T. P. et al. 1994. A central role of salicylic acid in plant disease resistance. *Science*. 266, 1247-1250.
- Den Boon, J. A., Ahlquist, P. 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annual Review of Microbiology*. 64, 241-256.
- Denancé, N., Sánchez-Vallet, A., Goffner, D., Molina, A. 2013. Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs. *Frontiers in Plant Science*. 4, 155.
- Desnues, C., Boyer, M., Raoult, D. 2012. Sputnik, a virophage infecting the viral domain of life. *Advances in Virus Research*. 82, 63-89.

- Dessau, M., Goldhill, D., McBride, R. L., Turner, P. E., Modis, Y. 2012. Selective pressure causes an RNA virus to trade reproductive fitness for increased structural and thermal stability of a viral enzyme. *PLoS Genetics*. 8, e1003102.
- DeWitt, T. J., Sih, A., Wilson, D.S. 1998. Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution*. 13, 77-81.
- Di Giallonardo, F., Holmes, E. C. 2015. Viral Biocontrol: grand experiments in disease emergence and evolution. *Trends in Microbiology*. 23, 83-90.
- Dickins, T.E., Rahman, Q. 2012. The extended evolutionary synthesis and the role of soft inheritance in evolution. *Proceedings of the Royal Society B*. 279, 2913-2921.
- Diener, T.O. 1963. Physiology of virus-infected plants. *Annual Review of Phytopathology*. 1, 197-218.
- Diezma-Navas, L. et al. 2019. Crosstalk between epigenetic silencing and infection by tobacco rattle virus in *Arabidopsis*. *Molecular Plant Pathology*. 20, 1439-1452.
- Dolan, P.T., Whitfield, Z.J., Andino, R. 2018. Mechanisms and concepts in RNA virus population dynamics and evolution. *Annual Review of Virology*. 5, 69-92.
- Dombrovsky, A., Huet, H., Chejanovsky, N., Raccah, B. 2005. Aphid transmission of a potyvirus depends on suitability of the helper component and the N terminus of the coat protein. *Archives of Virology*. 150, 287-298.
- Domingo, E., Perales, C. 2019. Viral quasispecies. *PLoS Genetics*. 15, e1008271.
- Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. *Microbiology and Molecular Biology Reviews*. 76, 159-216.
- Dormann, C. F., Gruber, B., Freund, J. 2008. Introducing the Bipartite Package: Analyzing Ecological Networks. *R News*. 8, 8-11.
- Duffy, S., Shackelton, L. A., Holmes, E.C. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*. 9, 267-276.

- Durgbanshi, A. et al. 2005. Simultaneous determination of multiple phytohormones in plant extracts by liquid chromatography-electrospray tandem mass spectrometry. *Journal of Agricultural and Food Chemistry*. 53, 8437-8442.
- Elena, S. F., and Sanjuán, R. 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *Journal of Virology*. 79, 11555-11558.
- Elena, S.F., Carrasco, P., Daròs, J.A., Sanjuán, R. 2006. Mechanisms of genetic robustness in RNA viruses. *EMBO Reports*. 7, 168-173.
- Elena, S. F. 2012. RNA virus genetic robustness: possible causes and some consequences. *Current Opinion in Virology*. 2, 525-530.
- Elena, S. F., Solé, R. V., Sardanyés, J. 2010. Simple genomes, complex interactions: epistasis in RNA virus. *Chaos*. 20, 026106.
- Engering, A., Hogerwerf, L., and Slingenbergh, J. 2013. Pathogen-host-environment interplay and disease emergence. *Emerging Microbes and Infections*. 2, 1.
- Falconer, D.S., Mackay, T.F.C. 1995. *Introduction to Quantitative Genetics*, fourth ed. Longman, New York.
- Fernández-Calvino, L. et al. 2014. Virus-induced alterations in primary metabolism modulate susceptibility to tobacco rattle virus in *Arabidopsis*. *Plant Physiology*. 166, 1821-1838.
- Finazzi, G., Minagawa, J. 2014. High light acclimation in green microalgae. *Advances in Photosynthesis and Respiration*. 40, 445-469.
- Fitter, A., Hay, R. 2002. *Environmental Physiology of Plants*, third ed. Academic Press, San Diego.
- Fitzgera, P.J., Stoner, W.N. 1967. Barley yellow dwarf studies in wheat (*Triticum aestivum* L.). I. Yield and quality of hard red winter wheat infected with barley yellow dwarf virus. *Crop Sciences*. 7, 337-341.

- Fletcher, S. J. et al. 2016. The tomato spotted wilt virus genome is processed differentially in its plant host *Arachis hypogaea* and its thrips vector *Frankliniella fusca*. *Frontiers in Plant Science*. 7, 1349.
- Fraile, A. et al. 2017. Environmental heterogeneity and the evolution of plant-virus interactions: viruses in wild pepper populations. *Virus Research*. 241, 68-76.
- French, R.K., Holmes, E.C. 2020. An ecosystems perspective on virus evolution and emergence. *Trends in Microbiology*. 28, 165-175.
- Fry, J. D. 1996. The evolution of host specialization: are trade-offs overrated? *American Naturalist*. 148, 84-107.
- Furió, V. et al. 2012. Relationship between within-host fitness and virulence in the vesicular stomatitis virus: correlation with partial decoupling. *Journal of Virology*. 86, 12228-12236.
- Gandon, S. et al. 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. *Proceedings of the Royal Society B*. 263, 1003-1009.
- Gandon, S. 2004. Evolution of multihost parasites. *Evolution*. 58, 455-469.
- Gandon, S., Michalakis, Y. 2002. Local adaptation, evolutionary potential and host-parasite coevolution: interactions between migration, mutation, population size and generation time. *Journal of Evolutionary Biology*. 15, 451-462.
- Ganusov, V. V., Bergstrom, C. T., and Antia, R. 2002. Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. *Evolution*. 56, 213-223.
- Garrett, K. A., Dendy, S. P., Frank, E. E., Rouse, M. N., Travers, S. E. 2006. Climate change effects on plant disease: genomes to ecosystems. *Annual Review of Phytopathology*. 44, 489-509.

- Gavrilets, S., and Gibson, N. 2002. Fixation probability in a spatially heterogeneous environment. *Population Ecology*. 44, 51-58.
- Geng, C. et al. 2017. Tobacco vein banding mosaic virus 6K2 protein hijacks NbPsbO1 for virus replication. *Scientific Reports*. 7, 43455.
- Gibbs, A. 1980. A plant virus that partially protects its wild legume host against herbivores. *Intervirology*. 13, 42-47.
- González, R., Butković, A., Elena, S. F. 2020. From foes to friends: Viral infections expand the limits of host phenotypic plasticity. *Advances in Virus Research*. 106, 85-121.
- González, R., Butković, A., Elena, S.F. 2019. Role of host genetic diversity for susceptibility-to-infection in the evolution of virulence of a plant virus. *Virus Evolution*. 5, vez024.
- Gorovits, R., Sobol, I., Altaleb, M., Czosnek, H., Anfoka, G. 2019. Taking advantage of a pathogen: understanding how a virus alleviates plant stress response. *Phytopathology Research*. 1, 20.
- Grangeon, R. et al. 2013. 6K2-induced vesicles can move cell to cell during turnip mosaic virus infection. *Frontiers in Microbiology*. 4, 351.
- Grundy, J., Stoker, C., Carré, I. A. 2015. Circadian regulation of abiotic stress tolerance in plants. *Frontiers in Plant Science*. 6, 648.
- Gusta, L. V., Wisniewski, M. 2012. Frost tolerance in plants, in: Shabala, S. (Ed.), *Plant Stress Physiology*. CABI Publishing, Wallingford, pp. 132-147.
- Hagemann, N. et al. 2016. Evolution of photorespiration from cyanobacteria to land plants, considering protein phylogenies and acquisition of carbon concentrating mechanisms. *Journal of Experimental Botany*. 67, 2963-2976.

- Hanton, S. L. et al. 2005. Diacidic motifs influence the export of transmembrane proteins from endoplasmic reticulum in plant cells. *Plant Cell*. 17, 3081-3093.
- Haraguchi, Y., Sasaki, A. 2000. The Evolution of parasite virulence and transmission rate in a spatially structured population. *Journal of Theoretical Biology*. 203, 85-96.
- Hatfield, J. L., Prueger, J. H. 2015. Temperature extremes: effect on plant growth and development. *Weather Climate. Extremes*. 10, 4-10.
- Hébrard, E. et al. 2009. Intrinsic disorder in viral proteins genome-linked: experimental and predictive analyses. *Virology Journal*. 6, 23.
- Hecht, M., Bromberg, Y., Rost, B. 2015. Better prediction of functional effects for sequence variants. *BMC Genomics*. 16, 1.
- Hillung, J., Cuevas, J. M., Elena, S. F. 2012. Transcript profiling of different *Arabidopsis thaliana* ecotypes in response to tobacco etch potyvirus infection. *Frontiers in Microbiology*. 3, 229.
- Hillung, J., Elena, S. F., Cuevas, J. M. 2013. Intra-specific variability and biological relevance of P3N-PIPO protein length in potyviruses. *BMC Evolutionary Biology*. 13, 249.
- Hillung, J., Cuevas, J. M., Valverde, S., Elena, S. F. 2014. Experimental evolution of an emerging plant virus in host genotypes that differ in their susceptibility to infection. *Evolution* 68, 2467-2480.
- Hillung, J., Cuevas, J. M., Elena, S. F. 2015. Evaluating the within-host fitness effects of mutations fixed during virus adaptation to different ecotypes of a new host. *Philosophical Transactions of the Royal Society B*. 370, 20140292.
- Hily, J. M. et al. 2016. Environment and host genotype determine the outcome of a plant-virus interaction: from antagonism to mutualism. *New Phytologist*. 209, 812-822.



- Hogenhout S. A., Ammar E., Whitfield A. E., Redinbaugh M. G. 2008. Insect vector interactions with persistently transmitted viruses. *Annual Review of Phytopathology*. 46, 327-359
- Hooke, R. 2007. *Micrographia or some physiological descriptions of minute bodies made by magnifying glasses*. Cosmo Classics, New York.
- Hühne, R., Koch, T., Sühnel, J. 2007. A comparative view at comprehensive information resources on three-dimensional structures of biological macromolecules. *Briefings in Functional Genomics and Proteomics*. 6, 220-239.
- Hughes, W. O. H., Boomsma, J. J. 2006. Does Genetic Diversity Hinder Parasite Evolution in Social Insect Colonies? *Journal of Evolutionary Biology*. 19, 132-143.
- Hull, R. 2002. *Matthews' Plant Virology*, fourth ed. Academic Press, San Diego.
- International Committee on Taxonomy of Viruses, King, A.M.Q. (Eds.), 2012. *Virus taxonomy: classification and nomenclature of viruses: ninth report of the International Committee on Taxonomy of Viruses*. Academic Press, London ; Waltham, MA.
- Iranzo, J., Puigbò, P., Lobkovsky, A. E., Wolf, Y. I., Koonin, E. V. 2016. Inevitability of genetic parasites. *Genome Biology and Evolution*. 8, 2856-2869.
- Jiang, J., Patarroyo, C., Garcia Cabanillas, D., Zheng, H., Laliberté, J. F. 2015. The vesicle-forming 6K2 protein of turnip mosaic virus interacts with COPII coatomer Sec24a for viral systemic infection. *Journal of Virology*. 89, 6695-6710.
- Jiang, J., Laliberté, J. F. 2011. The genome-linked protein VPg of plant viruses - a protein with many partners. *Current Opinion in Virology*. 1, 347-354.
- Jones, R. A. C. 2016. Future scenarios for plant virus pathogens as climate change progresses. *Advances in Virus Research*. 95, 87-147.

- Jones, R. A. C., Barbetti, M. J. 2012. Influence of climate change on plant disease infections and epidemics caused by viruses and bacteria. *CAB Reviews*. 7, 22.
- Källberg, M. et al. 2012. Template-based protein structure modeling using the RaptorX web server. *Nature Protocols*. 7, 1511-1522.
- Kassen, R. 2002. The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of Evolutionary Biology*. 15, 173-190.
- Kassen, R., Bell, G. 1998. Experimental evolution in *Chlamydomonas*. IV. Selection in environments that vary through time at different scales. *Heredity*. 80, 732-741.
- Kathiria, P. et al. 2010. Tobacco mosaic virus infection results in an increase in recombination frequency and resistance to viral, bacterial, and fungal pathogens in the progeny of infected tobacco plants. *Plant Physiology*. 153, 1859-1870.
- Kawabata, T., Nishikawa, L. 2000. Protein tertiary structure comparison using the Markov transition model of evolution. *Proteins*. 41, 108-122.
- Kawabata, T. 2003. MATRAS: a program for protein 3D structure comparisons. *Nucleic Acids Research*. 31, 3367-3369.
- Khan, M. I. R., Fatma, M., Per, T. S., Anjum, N. A., Khan, N. A. 2015. Salicylic acid-induced abiotic stress tolerance and underlying mechanisms in plants. *Frontiers in Plant Science*. 6, 462.
- Kim, D., Paggi, J. M., Park, C., Bennett, C., Salzberg, S. L. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*. 37, 907-915.
- Kleczkowski, A. 1950. Interpreting relationships between the concentrations of plant viruses and number of local lesions. *Journal of General Microbiology*. 4, 53-69.

- Knies, J. L., Izem, R., Supler, K. L., Kingsolver, J. G., Burch, C. L. 2006. The genetic basis of thermal reaction norm evolution in lab and natural phage populations. *PLoS Biology*. 4, e201.
- Knipe, D. M., Howley, P. M. 2013. *Fields virology*, 6th ed. ed. Wolters Kluwer/Lippincott Williams and Wilkins Health, Philadelphia, PA.
- Koonin, E. V. et al. 2008. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nature Reviews Microbiology*. 6, 925-939.
- Koonin, E.V., Dolja, V. V. 2013. A virocentric perspective of the evolution of life. *Current Opinion in Virology*. 3, 546-557.
- Koonin, E.V. 2016. Viruses and mobile elements as drivers of evolutionary transitions. *Philosophical Transactions of the Royal Society B*. 371, 20150442.
- Koonin, E.V., Dolja, V. 2013. A virocentric perspective on the evolution of life. *Current Opinion in Virology*. 3, 546-557.
- Krämer, U. 2015. Planting molecular functions in an ecological context with *Arabidopsis thaliana*. *eLife*. 4, e06100.
- Krasensky, J., Jonak, C. 2012. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of Experimental Botany*. 63, 1593-1608.
- Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E. L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*. 305, 567-580.
- Kubinak, J. L. et al. 2014. Serial infection of diverse host (Mus) genotypes rapidly impedes pathogen fitness and virulence. *Proceedings of the Royal Society B*. 282, 20141568.

- Lacroix, C., Seabloom, E. W., Borer, E. T. 2014. Environmental nutrient supply alters prevalence and weakens competitive interactions among coinfecting viruses. *New Phytologist*. 204, 424-433.
- Lafforgue, G., Tromas, N., Elena, S. F., Zwart, M. P. 2012. Dynamics of the establishment of systemic *Potyvirus* infection: independent yet cumulative action of primary infection sites. *Journal of Virology*. 86, 12912-12922.
- Laine, A. L. 2007. Pathogen fitness components and genotypes differ in their sensitivity to nutrient and temperature variation in a wild plant-pathogen association. *Journal of Evolutionary Biology*. 20, 2371-2378.
- Lalić, J., Elena, S. F. 2012. Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. *Heredity* 109, 71-77.
- Lalić, J., Elena, S. F. 2015. The impact of high-order epistasis in the within-host fitness of a positive-sense plant RNA virus. *Journal of Evolutionary Biology*. 28, 2236-2247.
- Lalić, J., Agudelo-Romero, P., Carrasco, P., Elena, S. F. 2010. Adaptation of tobacco etch potyvirus to a susceptible ecotype of *Arabidopsis thaliana* capacitates it for systemic infection of resistant ecotypes. *Philosophical Transactions of the Royal Society B*. 365, 1997-2007.
- Lazzaro, B. P., Little, T. J. 2009. Immunity in a variable world. *Philosophical Transactions of the Royal Society B*. 364, 15-26.
- Lê, S., Josse, J. and Husson, F. 2008. FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*. 25, 1-18.
- Lefevre, P. et al. 2019. Evolution and ecology of plant viruses. *Nature Reviews Microbiology*. 17, 632-644.

- Leggett, H. C. et al. 2013. Generalism and the evolution of parasite virulence. *Trends in Ecology and Evolution*, 10, 592-596.
- Lively, C. M. 2010. The effect of host genetic diversity on disease spread. *American Naturalist*. 175, 149-152.
- Lerich, A., Langhans, M., Sturm, S., Robinson, D.G. 2011. Is the 6 kDa tobacco etch viral protein a bona fide ERES marker? *Journal of Experimental Botany*. 62, 5013-5023.
- Levy, S.F. et al. 2015. Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*. 519, 181-186.
- Lewontin, R. C. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- Lewsey, M., Palukaitis, P., Carr, J. P. 2018. Plant-Virus Interactions: Defence and Counter-Defence, in: Roberts, J.A. (Ed.), *Annual Plant Reviews Online*. John Wiley and Sons, Ltd, Chichester, UK, pp. 134-176.
- Li, H. et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25, 2078-2079.
- Li, P. et al. 2017. Vector and nonvector insect feeding reduces subsequent plant susceptibility to virus transmission. *New Phytologist*. 215, 699-710.
- Listmann, L., LeRoch, M., Schlüter, L., Thomas, M.K., Reusch, T.B.H. 2016. Swift thermal reaction norm evolution in a key marine phytoplankton species. *Evolutionary Applications*. 9, 1156-1164.
- Little, T., Shuker, D., Colegrave, N., Day, T., Graham, A. L. 2010. The coevolution of virulence: tolerance in perspective. *PLoS Pathogens*. 6, e1001006.
- Liu, J. et al. 2016. *Ghd2*, a *CONSTANS*-like gene, confers drought sensitivity through regulation of senescence in rice. *Journal of Experimental Botany*. 67, 5785-5798.

- Löhmus, A., Varjosalo, M., Mäkinen, K. 2016. Protein composition of 6K2- induced membrane structures formed during potato virus A infection. *Molecular Plant Pathology*. 17, 943-958.
- Loreti, E., van Veen, H., Perata, P. 2016. Plant responses to flooding stress. *Current Opinion in Plant Biology*. 33, 64-71.
- Love, M. I., Huber, W., Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 15, 550.
- Machado, J. P. B., Calil, I. P., Santos, A. A., Fontes, E. P. B. 2017. Translational control in plant antiviral immunity. *Genetics and Molecular Biology*. 40, 292-304.
- Madlung, A. 2004. The effect of stress on genome regulation and structure. *Annals of Botany*. 94, 481-495.
- Maere, S., Heymans, K., Kuiper, M. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics*. 21, 3448-3449.
- Malpica, J. M. et al. 2006. Association and host selectivity in multi-host pathogens. *PLoS ONE*. 1, e41.
- Mann, N. H. 2003. Phages of the marine cyanobacterial picophytoplankton. *FEMS Microbiology Reviews*. 27, 17-34.
- Márquez, L. M., Redman, R. S., Rodriguez, R. J., Roossinck, M. J. 2007. A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science*. 315, 513-515.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*. 17, 10.
- Martínez, F. et al. 2016. Interaction network of tobacco etch potyvirus NIa protein with the host proteome during infection. *BMC Genomics*. 17, 87.

- Mauck, K. E., De Moraes, C. M., Mescher, M. C. 2010. Effects of cucumber mosaic virus infection on vector and non-vector herbivores of squash. *Communicative and Integrative Biology*. 3, 579-582.
- Mauck, K. E., Kenney, J., Chesnais, Q. 2019. Progress and challenges in identifying molecular mechanisms underlying host and vector manipulation by plant viruses. *Current Opinion in Insect Science*. 33, 7-18.
- McBride, R. C., Ogbunugafor, C. B., Turner, P. E. 2008. Robustness promotes evolvability of thermotolerance in an RNA virus. *BMC Evolutionary Biology*. 8, 231.
- McLeish, M. et al. 2017. Scale dependencies and generalism in host use shape virus prevalence. *Proceedings of the Royal Society B*. 284, 20172066.
- Merits, A., et al. 2002. Proteolytic processing of potyviral proteins and polyprotein processing intermediates in insect and plant cells. *Journal of General Virology*. 83, 1211-1221.
- Meyerowitz, E. M., Pruitt, R. E. 1985. *Arabidopsis thaliana* and Plant Molecular Genetics. *Science*. 229, 1214-1218.
- Miller, S., Krijnse-Locker, J. 2008. Modification of intracellular membrane structures for virus replication. *Nature Reviews Microbiology*. 6, 363-374.
- Mingot A., Valli A., Rodamilans B., San León D., Baulcombe D. C., García J. A., López-Moya J. J. 2016. The P1N-PISPO trans-frame gene of sweet potato feathery mottle potyvirus is produced during virus infection and functions as an RNA silencing suppressor. *Journal of Virology*. 90, 3543-3557.
- Mittler, R. 2006. Abiotic stress, the field environment and stress combination. *Trends in Plant Science*. 11, 15-19.

- Miura, K., Tada, Y. 2014. Regulation of water, salinity, and cold stress responses by salicylic acid. *Frontiers in Plant Science*. 5, 4.
- Miyashita, S., Kishino, H. 2010. Estimation of the size of genetic bottlenecks in cell-to-cell movement of Soil-borne wheat mosaic virus and the possible role of the bottlenecks in speeding up selection of variations in trans-acting genes or elements. *Journal of Virology*. 84, 1828-1837.
- Moreno-Gómez, S., Stephan, W., Tellier, A. 2013. Effect of disease prevalence and spatial heterogeneity on polymorphism maintenance in host-parasite interactions. *Plant Pathology*. 62, 133-141.
- Morens, D. M., Fauci, A. S. 2013. Emerging infectious diseases: threats to human health and global stability. *PLoS Pathogens*. 9, e1003467.
- Movahed, N. et al. 2017. Cylindrical inclusion protein of turnip mosaic virus serves as a docking point for the intercellular movement of viral replication vesicles. *Plant Physiology*. 175, 1732-1744.
- Munné-Bosch, S., Alegre, L. 2004. Die and let live: leaf senescence contributes to plant survival under drought stress. *Functional Plant Biology*. 31, 203-216.
- Murren, C. J. et al. 2015. Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity*. 115, 293-301.
- Mushegian, A. R. 2020. Are there  $10^{31}$  virus particles on Earth, or more, or less? *Journal of Bacteriology*. 202, 00052-20.
- Newman, M. E. J. 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the USA*. 103, 8577-8582.
- Oliver, J. E., Whitfield, A. E. 2016. The genus *Tospovirus* : Emerging bunyaviruses that threaten food security. *Annual Review of Virology*. 3, 101-124.
- Paez-Espino, D. et al. 2016. Uncovering Earth's virome. *Nature*. 536, 425-430.



- Pagán, I. et al. 2010. *Arabidopsis thaliana* as a model for the study of plant-virus coevolution. *Philosophical Transactions of the Royal Society B*. 365, 1983-1995.
- Pandey P., Ramegowda, V., Senthil-Kumar, M. 2015. Shared and unique responses of plants to multiple individual stresses and stress combinations: physiological and molecular mechanisms. *Frontiers in Plant Science*. 6, 723.
- Papaix, J. et al. 2013. Dynamics of adaptation in spatially heterogeneous metapopulations. *PLoS ONE*. 8, e54697.
- Parrat, S. R., Numminen, E., Laine, A. L. 2016. Infectious disease dynamics in heterogeneous landscapes. *Annual Review of Ecology, Evolution, and Systematics*. 47, 283-306.
- Patel, D., Franklin, K. A. 2009. Temperature-regulation of plant architecture. *Plant Signaling and Behavior*. 4, 577-579.
- Paudel, D. B., Sanfaçon, H. 2018. Exploring the diversity of mechanisms associated with plant tolerance to virus infection. *Frontiers in Plant Science*. 9, 1575.
- Perales, C., Martín, V., Domingo, E. 2011. Lethal mutagenesis of viruses. *Current Opinion in Virology*. 1, 419-422.
- Perales, C., Domingo, E. 2019. Viral quasispecies. *PLoS Genetics*. 15, e1008271.
- Pfennig, K. S. 2001. Evolution of pathogen virulence: the role of variation in host phenotype. *Proceedings of the Royal Society B*. 268, 755-760.
- Prasch, C. M., Sonnewald, U. 2013. Simultaneous application of heat, drought, and virus to *Arabidopsis* plants reveals significant shifts in signaling networks. *Plant Physiology*. 162, 1849-1866.
- Prendeville, H. R., Ye, X., Morris, T. J., Pilson, D. 2012. Virus infections in wild plant populations are both frequent and often unapparent. *American Journal of Botany*. 99, 1033-1042.

- Prokopová, J. et al. 2010. Photosynthetic responses of lettuce to downy mildew infection and cytokinin treatment. *Plant Physiology*. 48, 716-723.
- Prosperi, M. C., Salemi, M. 2012. QuRe: software for viral quasispecies reconstruction from next-generation sequencing data. *Bioinformatics*. 28, 132-133.
- R Core Team. 2016. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- Rajamäki, M. L., Valkonen J. P. T. 1999. The 6K2 protein and the VPg of potato virus A are determinants of systemic infection in *Nicotiana glauca*. *MPMI*. 12, 1074-1081.
- Ramegowda, V., Senthil-Kumar, M. 2015. The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. *Journal of Plant Physiology*. 176, 47-54.
- Raudvere, U. et al. 2019. gProfiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*. 47, 191-198.
- Ravindran, S. 2012. Barbara McClintock and the discovery of jumping genes. *Proceedings of the National Academy of Sciences of the USA*. 109, 20198-20199.
- Regoes, R. R., Nowak, M. A., Bonhoeffer, S. 2000. Evolution of virulence in a heterogeneous host population. *Evolution*. 54, 64-71.
- Rejeb, I., Pastor, V., Mauch-Mani, B. 2014. Plant responses to simultaneous biotic and abiotic stress: molecular mechanisms. *Plants*. 3, 458-475.
- Restrepo-Hartwig, M. A., Carrington, J. C. 1994. The tobacco etch potyvirus 6-kilodalton protein is membrane associated and involved in viral replication. *Journal of Virology*. 68, 2388-2397.
- Revers, F., García, J. A. 2015. Molecular biology of *Potyriviruses*. *Advances in Virus Research*. 92, 101-199.

- Rodelo-Urrego, M. et al. 2013. Landscape heterogeneity shapes host-parasite interactions and results in apparent plant-virus codivergence. *Molecular Ecology*. 22, 2325-2340.
- Rodelo-Urrego, M., García-Arenal, F., Pagán, I. 2015. The effect of ecosystem biodiversity on virus genetic diversity depends on virus species: a study of chiltepin-infecting *Begomovirus* in Mexico. *Virus Evolution*. 1, vev004.
- Rodríguez, D. J., Torres, S. L. 2001. Models of infectious diseases in spatially heterogeneous environments. *Bulletin of Mathematical Biology*. 63, 547-571.
- Roossinck, M. J. 2003. Plant RNA virus evolution. *Current Opinion in Microbiology*. 6, 406-409.
- Roossinck, M. J. 2010. Lifestyles of plant viruses. *Philosophical Transactions of the Royal Society B*. 365, 1899-1905.
- Roossinck, M. J. 2011. The good viruses: viral mutualistic symbioses. *Nature Reviews Microbiology*. 9, 99-108.
- Roossinck, M. J. 2012. Plant virus metagenomics: biodiversity and ecology. *Annual Review of Genetics*. 46, 359-369.
- Roossinck, M. J. 2015a. Move over, bacteria! Viruses make their mark as mutualistic microbial symbionts. *Journal of Virology*. 89, 6532-6535.
- Roossinck, M. J. 2015b. Plants, viruses and the environment: ecology and mutualism. *Virology*. 479, 271-277.
- Roossinck, M. J., Bazán, E. R. 2017. Symbiosis: viruses as intimate partners. *Annual Review of Virology*. 4, 123-139.
- Roossinck, M. J. et al. 2010. Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology*. 19, 81-88.
- Rosenthal, S. R. et al. 2015. Redefining disease emergence to improve prioritization and macro-ecological analyses. *One Health*. 1, 17-23.

- Rubio, B. et al. 2019. Genome wide association study reveals new loci involve in *arabidopsis thaliana* and turnip mosaic virus (TuMV) interactions in the field. *New Phytologist*, 221, 2026-2038.
- Safari, M., Ferrari, M. J., Roossinck, M. J. 2019. Manipulation of aphid behavior by a persistent plant virus. *Journal of Virology*. 93, e01781-18.
- Salonen, A., Ahola, T., Kääriäinen, L. 2005. Viral RNA replication in association with cellular membranes. *Current Topics in Microbiology and Immunology*. 285, 139-173.
- Sanjuán, R., Moya, A., Elena, S. F. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proceedings of the National Academy of Sciences of the USA*. 101, 8396-8401.
- Sanjuán, R. 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. *Philosophical Transactions of the Royal Society B*. 365, 1975-1982.
- Sanjuán, R. 2012. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. *PLoS Pathogens*. 8, e1002685.
- Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L. M., Belshaw, R. 2010. Viral mutation rates. *Journal of Virology*. 84, 9733-9748.
- Sardanyés, J., Simó, C., Martínez, R., Solé, R. V., Elena, S. F. 2014. Variability in mutational fitness effects prevents full lethal transitions in large quasispecies populations. *Scientific Reports*. 4, 4625.
- Schaad, M. C., Jensen, P. E., Carrington, J. C. 1997. Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein. *EMBO Journal*. 16, 4049-4059.

- Schmid-Hempel, P., Koella, J. C. 1994. Variability and its implications for host-parasite interactions. *Parasitology Today*. 10, 98-102.
- Schneider, C. A., Rasband, W. S., Eliceiri, K. W. 2012. NIH image to ImageJ: 25 years of image analysis. *Nature Methods*. 9, 671-675.
- Seaton, G.G.R., Hurry, V.M., Rohozinski, J. 1996. Novel amplification of non-photochemical chlorophyll fluorescence quenching following viral infection in *Chlorella*. *FEBS Letters*. 389, 319-323.
- Shabala, S., Babourina, O., Rengel, Z., Nemchinov, L. G. 2010. Non-invasive microelectrode potassium flux measurements as a potential tool for early recognition of virus-host compatibility in plants. *Planta*. 232, 807-815.
- Shabala, S. et al. 2011a. Plasma membrane  $\text{Ca}^{2+}$  transporters mediate virus-induced acquired resistance to oxidative stress: calcium efflux and oxidative stress tolerance. *Plant, Cell and Environment*. 34, 406-417.
- Shabala, S. et al. 2011b. Endomembrane  $\text{Ca}^{2+}$ -ATPases play a significant role in virus-induced adaptation to oxidative stress. *Plant Signaling and Behavior*. 6, 1053-1056.
- Shannon CE. 1948. A mathematical theory of communication. *Bell System Technical Journal*. 27, 379-423.
- Shinozaki, K., Yamaguchi-Shinozaki, K., Seki, M. 2003. Regulatory network of gene expression in the drought and cold stress responses. *Current Opinion in Plant Virology*. 6, 410-417.
- Shinozaki, K., Uemura, M., Bailey-Serres, J., Bray, E. A., Weretilnyk, E. 2015. Responses to abiotic stress, in: Buchanan, B.B., Grissem, W., Jones, R.L. (Eds.). *Biochemistry and Molecular Biology of Plants*. Wiley, Chichester, pp. 1051-1100.

- Short, S. M. 2012. The ecology of viruses that infect eukaryotic algae: Algal virus ecology. *Environmental Microbiology*. 14, 2253-2271.
- Shukla, A., López-González, S., Hoffmann, G., Hafrén, A. 2019. Diverse plant viruses: a toolbox for dissection of cellular pathways. *Journal of Experimental Botany*. 70, 3029-3034.
- Shukla, A., Pagán, I., García-Arenal, F. 2017. Effective tolerance based on resource reallocation is a virus-specific defence in *Arabidopsis thaliana*. *Molecular Plant Pathology*. 19, 1454-1465.
- Simko, I., Piepho, H. P. 2012. The area under the disease progress stairs: calculation, advantage, and application. *Phytopathology*. 102, 381-389.
- Simmons, H. E., Munkvold, G. P. 2014. Seed Transmission in the Potyviridae, in: Gullino, M. L., Munkvold, G. (Eds.), *Global Perspectives on the Health of Seeds and Plant Propagation Material*. Springer Netherlands, Dordrecht, pp. 3-15.
- Singh, J., Zhang, X., Stewart, L. R., Mitchell, T., Qu, F. 2014. Role of double-stranded RNA-binding proteins in RNA silencing and antiviral defense, in: *Plant Virus-Host Interaction*. Elsevier, pp. 1-16.
- Skoog, F., Miller, C. O. 1957. Chemical regulation of growth and organ formation in plant tissues cultured in vitro. *Symposia of the Society for Experimental Biology*. 11, 118-130.
- Solé, R. V., Elena, S. F. 2019. *Viruses as Complex Adaptive Systems, Primers in complex systems*. Princeton University Press, Princeton.
- Sonnhammer, E.L., Von Heijne, G., Krogh, A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology*. 6, 175-182.

- Spetz, C., Valkonen, J. P. 2004. Potyviral 6K2 protein long-distance movement and symptom-induction functions are independent and host-specific. *Molecular Plant-Microbe Interactions*. 17, 502-510.
- Sreenivasulu, N., Harshavardhan, V. T., Govind, G., Seiler, C., Kohli, A. 2012. Contrapuntal role of ABA: does it mediate stress tolerance or plant growth retardation under long-term drought stress? *Gene*. 506, 265-273.
- Suttle, C. A. 2007. Marine viruses - major players in the global ecosystem. *Nature Reviews Microbiology*. 5, 801-812.
- Suzuki, N., Rivero, R.M., Shulaev, V., Blumwald, E., Mittler, R. 2014. Abiotic and biotic stress combinations. *New Phytologist*. 203, 32-43.
- Takanami, Y. 2006. Contributions of plant virus research to the biological sciences. *Journal of General Plant Pathology*. 72, 393-395.
- Takeda, S., Paszkowski, J. 2006. DNA methylation and epigenetic inheritance during plant gametogenesis. *Chromosoma*. 115, 27-35.
- Tellier, A., Brown, J. K. 2011. Spatial Heterogeneity, Frequency-Dependent Selection and Polymorphism in Host-Parasite Interactions. *BMC Evolutionary Biology*. 1, 11.
- Thapa, V., McGlenn, D. J., Melcher, U., Palmer, M. W., Roossinck, M. J. 2015. Determinants of taxonomic composition of plant viruses at the Nature Conservancy's Tallgrass Prairie Preserve, Oklahoma. *Virus Evolution*. 1, vev007.
- The 1001 Genomes Consortium. 2016. 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell*. 166, 481-491.
- Thimm, O. et al. 2004. mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal*. 37, 914-939.

- Thrall, P. H., Burdon, J. J. 2003. Evolution of virulence in a plant host-pathogen metapopulation. *Science*. 299, 1735-1737.
- Thrall, P. et al. 2012. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters*. 15, 425-435.
- Tromas, N., Zwart, M. P., Lafforgue, G., Elena, S. F. 2014. Within-host spatiotemporal dynamics of plant virus infection at the cellular level. *PLoS Genetics*. 10, e1004186.
- Vale, P. F., Salvaudon, L., Kaltz, O., Fellous, S. 2008a. The role of the environment in the evolutionary ecology of host parasite interactions. *Infection, Genetics and Evolution*. 8, 302-305.
- Vale, P. F., Stjenman, M., Little, T. J. 2008b. Temperature-dependent cost of parasitism and maintenance of polymorphism under genotype-by-environment interactions. *Journal of Evolutionary Biology*. 21, 1418-1427.
- Van der Biezen, E. A., Jones, J. D. G. 1998. Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*. 23, 454-456.
- Van Munster, M. et al. 2017. Water deficit enhances the transmission of plant viruses by insect vectors. *PLoS ONE*. 12, e0174398.
- Verchot, J. 2016. Plant virus infection and the ubiquitin proteasome machinery: arms race along the endoplasmic reticulum. *Viruses*. 8, 314.
- Visher, E., Whitefield, S. E., McCrone, J. T., Fitzsimmons, W., Luring, A. S. 2016. The mutational robustness of influenza A virus. *PLoS Pathogens*. 12, e1005856.
- Vurro, M., Bonciani, B., Vannacci, G. 2010. Emerging infectious diseases of crop plants in developing countries: impact on agriculture and socio-economic consequences. *Food Security*. 2, 113.



- Wan, J., Cabanillas, D. G., Zheng, H., Laliberté, J. F. 2015. Turnip mosaic virus moves systematically through both phloem and xylem as membrane-associated complexes. *Plant Physiology*. 167, 1374-1388.
- Wang, Y., Frei, M. 2011. Stressed food - the impact of abiotic environmental stresses on crop quality. *Agriculture, Ecosystems and Environment*. 141, 271-286.
- Wei T, et al. 2010. Sequential recruitment of the endoplasmic reticulum and chloroplasts for plant potyvirus replication. *Journal of Virology*. 84, 799-809.
- Weigel, D. 2012. Natural variation in *Arabidopsis*: from molecular genetics to ecological genomics. *Plant Physiology*. 158, 2-22.
- Weitz, J. S. et al. 2013. Phage-bacteria infection networks. *Trends in Microbiology*. 21, 82-91.
- Westwood, J. H. et al. 2013. A viral RNA silencing suppressor interferes with abscisic acid-mediated signalling and induces drought tolerance in *Arabidopsis thaliana*: virus-induced drought tolerance. *Molecular Plant Pathology*. 14, 158-170.
- Westwood, M. L. et al. 2019. The evolutionary ecology of circadian rhythms in infection. *Nature Ecology and Evolution*. 3, 552-560.
- Whitlock, M. C. 1996. The Red Queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. *American Naturalist*. 148, 65-77.
- Whitlock, M. C. 2002. Selection, load and inbreeding depression in a large metapopulation. *Genetics*. 160, 1191-1202.
- Whitlock, M. C. 2003. Fixation probability and time in subdivided populations. *Genetics*. 164, 767-779.
- Whitlock, M. C., Gomulkiewicz, R. 2005. Probability of fixation in a heterogeneous environment. *Genetics*. 171, 1407-1417.

- Wieczynski, D. J., Turner, P. E., Vasseur, D. A. 2018. Temporally autocorrelated environmental fluctuations inhibit the evolution of stress tolerance. *American Naturalist*. 191, 195-207.
- Wilke, C. O. 2005. Quasispecies theory in the context of population genetics. *BMC Evolutionary Biology*. 5, 44.
- Wilke, C., Forster, R., Novella, I. S. 2006. Quasispecies in time-dependent environments. *Current Topics in Microbiology and Immunology*. 299, 33-50.
- Wolinska, J., King, K. C. 2009. Environment can alter selection in host-parasite interactions. *Trends in Parasitology*. 25, 236-244.
- Woolhouse, M. E. J., Dye, C. 2001. Preface: population biology of emerging and re-emerging pathogens. *Philosophical Transactions of the Royal Society B*. 356, 981-982.
- Wren, J. D. et al. 2006. Plant virus biodiversity and ecology. *PLoS Biology*. 4, 314-315.
- Wu G, et al. 2018. Dynamin-like proteins of endocytosis in plants are coopted by potyviruses to enhance virus infection. *Journal of Virology*. 92, e01320-18.
- Wu, X., Valli, A., García, J., Zhou, X., Cheng, X. 2019. The tug-of-war between plants and viruses: great progress and many remaining questions. *Viruses*. 11, 203.
- Xie, J., Jiang, D. 2014. New insights into mycoviruses and exploration for the biological control of crop fungal diseases. *Annual Review of Phytopathology*. 52, 45-68.
- Xu, P. et al. 2008. Virus infection improves drought tolerance. *New Phytologist*. 180, 911-921.
- Yang, Y., Wang, W., Chu, Z., Zhu, J. K. and Zhang, H. 2017. Roles of nuclear pores and nucleo-cytoplasmic trafficking in plant stress responses. *Frontiers in Plant Science*. 8, 574.

- Yao, Y., Kovalchuck, I. 2011. Abiotic stress leads to somatic and heritable changes in homologous recombination frequency, point mutation frequency and microsatellite stability in *Arabidopsis* plants. *Mutation Research*. 707, 61-66.
- Yarwood, C. E. 1957. Heat tolerance of virus-infected and fungus-infested tissues. *Phytopathology*. 47, 538-539.
- Yates, A., Antia, R., Regoes, R. R. 2006. How do pathogen evolution and host heterogeneity interact in disease emergence? *Proceedings of the Royal Society B*. 273, 3075-3083.
- Zhang, R. et al. 2017. A high quality *Arabidopsis* transcriptome for accurate transcript-level analysis of alternative splicing. *Nucleic Acids Research*. 45, 5061-5073.
- Zhang, Y. Y., Fischer, M., Colot, V., Bossdorf, O. 2012. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist*. 197, 314-322.
- Zhu, J. K. 2016. Abiotic stress signaling and responses in plants. *Cell*. 167, 313-324.