Application of genome-wide association studies (GWAS) to identify host targets of viral adaptation

Anamarija Butković

Programa de Doctorado en Biodiversidad y Biología Evolutiva

Director:

Prof. Santiago F. Elena Fito

Tutor:

Dr. Rafael Sanjuán Verdeguer

May 2021

D. 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 (I^2SysBio), centro mixto del CSIC y de la Universitat de València (UV) y **D. Rafael Sanjuán Verdeguer**, doctor en Ciencias Biológicas y Profesor Titular de la UV en el I^2SysBio,

CERTIFICAN:

Que **Doña Anamarija Butković**, Máster en Biología Molecular por la Universidad de Zagreb (Croacia), ha realizado bajo mi supervisión la Tesis Doctoral titulada "*Application of genome-wide association studies (GWAS) to identify host targets of viral adaptation*" y que, por la presente, dan la conformidad a su depósito para su evaluación y posterior defensa pública.

Y para que así conste, en cumplimiento de la legislación vigente, firman el presente certificado en Valencia a 17 de Mayo de 2021.

ELENA FITO SANTIAGO FRANCISCO - DNI 25396106N

Firmado digitalmente por ELENA FITO SANTIAGO FRANCISCO - DNI 25396106N
Fecha: 2021.05.17 11:43:00 +02'00'

Fdo. Santiago F. Elena Fito

Firmado digitalmente por SANJUAN VERDEGUER, RAFAEL (FIRMA)
Fecha: 2021.05.17 19:32:45 +02'00'

Fdo. Rafael Sanjuán Verdeguer

Prof. Santiago F. Elena, PhD
E-mail: santiago.elena@uv.es
Web: http://sfelenalab.csic.es
Twitter: @SFElenaLab

i2sysbio

Parc Científic UV
Catedrático Agustín Escardino 9

46980 Paterna (Valencia)
Tel: +34 963 544 779

Table of Contents

Acknowledgements

My PhD was the hardest and the best thing I have ever done. There were moments where I felt overwhelmed and frustrated but I always found the strength to continue and do the best I can. Of course, this would not be possible without the help of amazing people around me. There were many but for the sake of being short I want to focus more closely on those that had given me great inspiration and showed me how to be a better version of myself. The biggest influence during my PhD was my mentor, Santiago Elena. He is everything I always wanted to be as a scientist and more. Full of great ideas and always ready to help even with a packed schedule, always asking how you are and making jokes to brighten the mood, always pushing us to be better and just being there when needed. Thank you for being like a second father to me when I was far away from home and pushing me to be a better scientist. I would never have become a scientist I am today or would have achieved what I did without your help and guidance. I hope one day I could be half as good a scientist and mentor as you. Thank you! Another person who helped shape me into who I am today and was always there when I felt sad or angry or unmotivated was my husband, Rubén. You are an amazing person and scientist and you have showed me such love and patience during these 4 years that no words of thank you will ever be enough. Thank you for lifting me up when I was down and for looking after me. You were always a friend I could have asked anything and you gave me the best advice. No bad result or rejection ever felt too bad with you around. Thank you for always being there for me! A person that was a constant support to me and everybody else in the lab was Paqui. Someone who always had time for other people and who made sure everything was perfect. You were always so kind to me and genuinely interested in my well-being. Thank you for making my experience comfortable, easy and memorable. And thank you for being there for me like a second mother who I could always ask anything. You deserve the best in this world and I hope you always receive it. Thank you! Now I want to thank the people who have always loved me with all their heart and gave me the best they could, my parents. Thank you for always doing what's best for me, thank you for making sure I had all I needed, thank you for doing the hard things so I could have the best. And thank you for supporting me and being proud of me when I wanted to pursue science and move away from home. I think of you every day and always miss you. Thank you, mom, for taking care of my dad and letting me follow my dreams. Thank you, dad, for always asking how I am in your lucid moments. I love you both very much and I would never be the person I am today without your support and help. Thank you! And

last but not least I want to thank all my lab members, Régis, Paula, Rebecca, Julia, José Luis, Paul, Susana, and Silvia for their help with experiments and for all the good laughs we had during coffee breaks. I especially want to thank Christina for being such a good friend, for always having time for me and for all the wonderful memories. I did not manage to get to know our new PhD students Izan and María José too well, but from the little time I know them I think they are great people and I wish them all the best in their career. Don’t give up and never be afraid to ask for help.

Abbreviations

(−)ssRNA: negative sense single-stranded RNA viruses

(+)ssRNA: positive sense single-stranded RNA viruses

6K1: 6 kDa protein 1

6K2: 6 kDa protein 2

ABA: abscisic acid

AGO: argonaute proteins

AMM: accelerated mixed model

AUDPS: area under the disease progression stairs

AUSIPS: symptoms intensity progression steps curve

BAK1: (*BRI1*)-*ASSOCIATED KINASE 1*

BR: brassinosteroids

BRI1: *BRASSINOSTEROID INSENSITIVE 1*

cDNA: complementary DNA

CI: cylindrical inclusion protein

CK: cytokinin

CP: coat protein

DCL: dicer-like proteins

DEGs: differentially expressed genes

dsDNA-RT: double-stranded DNA viruses with an RNA intermediate

dsDNA: double-stranded DNA viruses

dsRNA: double-stranded RNA viruses

DUB: viral deubiquitinating enzyme

eds8-1: *enhanced disease susceptibility 8*

ET: ethylene

FDR: false discovery rate

GA: gibberellic acid

GWAS: genome-wide association studies

HC-Pro: helper component protease

hpRNA: hairpin RNA

HR: hypersensitive response

ICTV: International Committee on Taxonomy of Viruses

ISR: induced systemic resistance

JA: jasmonic acid

jin1: *jasmonate insensitive 1*

LD: linkage disequilibrium

LMM: linear mixed model

LOF: loss-of-function mutants

MAPK: mitogen-activated protein kinase

MCMC: Markov chain Monte Carlo

miRNA: microRNA

mRNA: messenger RNA

NIa-Pro: nuclear inclusion a-protease domain

NIa: nuclear inclusion protein A

NIb: nuclear inclusion b protein

ORF: open reading frame

P1: protease 1

P3: protein P3

P3N-PIPO: movement protein or pretty interesting *Potyviridae* open reading frame

PAMP: pathogen-associated molecular patterns

PRRs: pattern recognition receptors

PTI: PAMP-triggered immunity

RdRp: RNA-dependent RNA polymerase

RISC: RNA-induced silencing complex

RLKs: receptor-like kinases

RNAi: RNA silencing or RNA interference

ROS: reactive oxygen species

RT: reverse transcriptase

SA: salicylic acid

SAR: systemic acquired resistance

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

SNPs: single nucleotide polymorphisms

sRNA: small RNA

ssDNA: single-stranded DNA viruses

ssRNA-RT: single-stranded RNA viruses with a DNA intermediate

tasiRNA: trans-acting small interfering RNA

TEV: tobacco etch virus

TMV: tobacco mosaic virus

TuMV: turnip mosaic virus

TYMV: turnip yellow mosaic virus

Ub: ubiquitin

UPS: ubiquitin proteasome complex

VPg: viral genome-linked protein

VSV: vesicular stomatitis Indiana virus

Abstract

Viruses are the most abundant entities on Earth and have a great capacity for evolution and adaptation. Some viruses are able to infect a wide range of hosts causing damage in a number of important plants while others infect one host species really well and cause severe detrimental symptoms in a short time span. With time viruses can adapt well to novel hosts and increase their infectivity, virulence and therefore provoke more damage to the host. Yet we still lack knowledge about how plants respond to viral infection with viruses that have different adaptation histories or host ranges, or how viruses that are differently adapted to the host respond to distinct environmental challenges. This thesis tried to answer these questions with the help of turnip mosaic virus (TuMV) and plant host *Arabidopsis thaliana*. Four different strains of TuMV were used; two with different adaptation histories (1) one naïve to arabidopsis, (2) one preadapted to arabidopsis, and two with different host ranges (3) a virus able to infect different genotypes of arabidopsis equally well (generalist) and (4) a virus able to infect only one specific genotype of arabidopsis well (specialist). In the first experiment, a method called genome-wide association studies (GWAS) was used to associate arabidopsis genes involved in viral infection with the naïve and preadapted TuMV strains. Shared and specific host genes for the naïve or preadapted viruses were identified as potential drivers/targets of viral adaptation. Their role in infection was further corroborated with the help of loss-of-function (LOF) mutants. Gene *AT2G14080* showed a strong potential role in resistance to pathogens in arabidopsis. In the second experiment, using the same GWAS method, arabidopsis genes that responded differentially to a generalist and a specialist virus were identified and characterized. The generalist virus manipulated a similar set of host genes in order to infect a wide host range successfully. While the specialist virus manipulated more heterogeneous genes because of host-specific selective pressures that modulated the evolution of the specialist virus. Selected genes were characterized further with the help of LOF mutants. In the final experiment, arabidopsis was inoculated with the naïve and preadapted virus and their genetic robustness (the constancy of the phenotype under mutational changes) and environmental robustness (the constancy of the phenotype under environmental changes) were tested. In agreement with the plastogenetic congruence hypothesis, mutational and environmental robustnesses went hand by hand: the naïve virus proved to be more robust both to mutational and environmental perturbations than the preadapted one. These results show how adaptation to one environment limits

evolvability in alternative ones, thus restricting the capacity of the preadapted virus to quickly respond to future changes in temperature.

To my parents and my husband

Introduction

1.1. Discovery of viruses

The first mention of an agent that could pass through the bacterial filter happened in 1892 and was proposed by Dimitri Ivanovsky. He observed that the filtered tobacco leaf sap containing an agent smaller than bacteria could cause disease in healthy plants. This observation was further corroborated by Martinus Beijerinck six years later when he independently proved Ivanovsky's theory and further observed that after dilution the infected sap can regain its "strength" after replicating in living plant tissue. This put forth the theory that the small agent could reproduce only in living tissue and not the cell-free plant sap. Martinus described this new infectious agent as *contagium vivum fluidum* or *virus*. The term "virus" (Latin for poison) was used for these filterable infectious entities thus leading to the name of the first described plant virus, tobacco mosaic virus (TMV) (Liu, 2014).

1.2. General introduction into viruses

Viruses are the smallest known self-replicating organisms consisting only of a nucleic acid enveloped in a protein shell. They are parasites of the hosts translational machinery which they use to make more infectious virions and infect other cells or hosts. We can argue if viruses are alive or not but we cannot deny that they show compelling complexity and diversity that allowed them to infect virtually every living organism on Earth. Even in their "simplicity" they challenge us every day to try and understand them by changing our perspective and our current knowledge. This diversity came to recognition with the explosion of technologies, where electron microscopy proved to be important by providing a first-ever glance into the various shapes of virions. Scientist noticed that some virions were similar while some completely distinct, making it possible to group certain viruses together. This abundance of data led to the need for more thorough classification which gave birth to the International Committee on Taxonomy of Viruses (ICTV) composed of an international group of scientists that classify and name viruses in an organized manner. In the 2019 release of Virus Taxonomy, 55 orders, 168 families, 1421 genera and 6590 species were described (https://talk.ictvonline.org/taxonomy/). Another widely used representation of different viruses in nature was the Baltimore classification system (Fig. 1) that divides viruses into seven categories as presented in. This classification is based on the type of genome (DNA or RNA) that is packed in the virion and the pathway that every nucleic acid takes to synthesize the messenger RNA (mRNA).

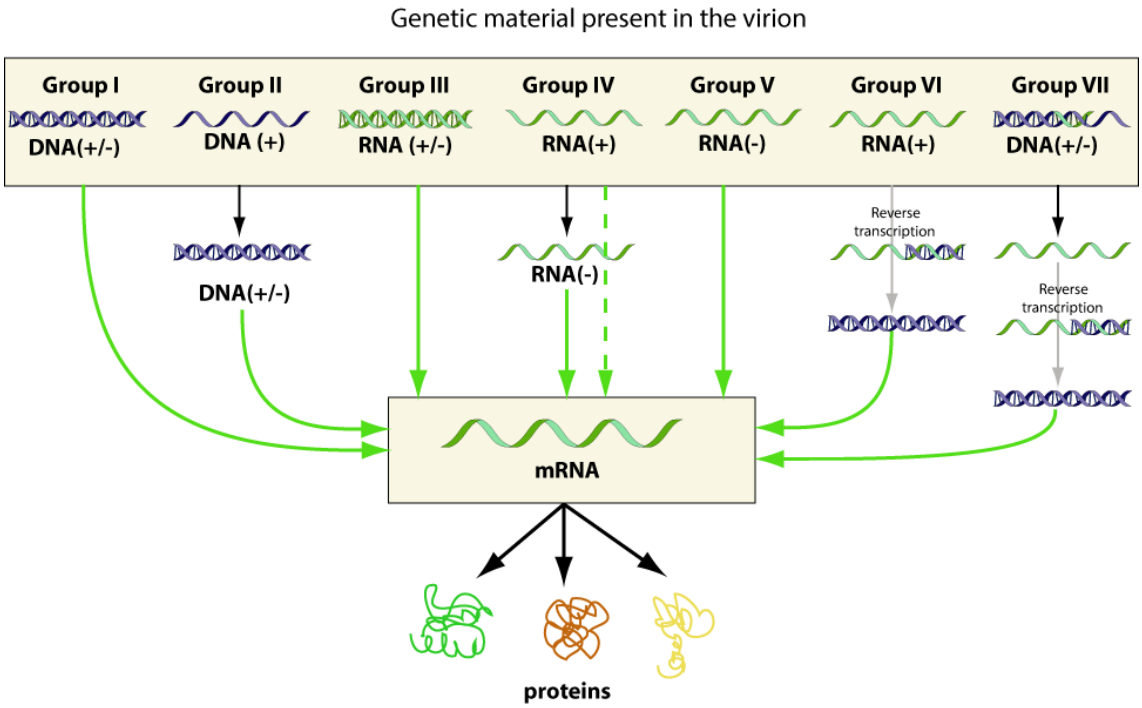


Figure 1.1. The Baltimore classification depending on the type of viral genome. The present virus classification comprises seven trees of life. Reprinted from ViralZone (source: https://viralzone.expasy.org/).

There are seven groups of viruses according to the Baltimore scheme: (I) dsDNA, (II) ssDNA, (III) dsRNA, (IV) (+)ssRNA, (V) (−)ssRNA, (VI) ssRNA-RT and (VII) dsDNA-RT viruses. DNA viruses of the group I and II replicate in the nucleus and use host cellular proteins. RNA viruses belonging to groups III, IV and V are translated and replicated in the cytoplasm. Group IV is translated directly in the cytoplasm with the aid of host translation machinery while group V first has to be transcribed in its complementary strand to be translation ready. Single-stranded RNA viruses are the most abundant and the majority of them can be found infecting vertebrates and plants, while a smaller number infects invertebrates and eukaryotic microorganisms. Group VI consists of positive sense ssRNA viruses with a dsDNA intermediate and group VII of dsDNA viruses with an ssRNA intermediate.

2. Factors influencing the evolution and adaptation of RNA viruses

2.1. Factors intrinsic to viruses

RNA viruses have short generation times, high mutation rates and compact genomes which gives them great adaptive potential.

2.1.1. Mutation rate

Viruses have high evolutionary potential in comparation with cellular organisms. Their short generation times and high mutation rates make them the perfect model organisms to study molecular evolution. Mutation rates or evolutionary rates in viruses depend on different aspects of their biology such as replication speed, polymerase fidelity and genomic architecture. When we take these aspects into account, we see that RNA viruses mutate faster than DNA viruses. This in part is due to the higher error rate of RNA-dependent RNA polymerases (RdRp) in comparation to DNA polymerases. RdRps and RNA-dependent DNA polymerases (retrotranscriptases, RT) have higher mutation rates because there is no proofreading or base excision repair in these enzymes. Comparing the mutation rates between these two enzymes we can see that RdRps have a far higher rate of ~1 mutation per genome per replication compared to the RTs that have 0.1-0.3 mutations per genome per replication (Holmes, 2009). Therefore, due to their high mutation rates RNA viruses live in the error threshold and accumulate more deleterious mutations. Their size is limited to max ~30 kb because larger genomes would mean that the lethal mutations would accumulate more frequently thus leading to greater instability (Duffy et al., 2008). Yet this type of fast and error-prone replication allows them to generate large heterogeneous populations, called quasispecies, and live in a mutation-selection balance (Domingo and Holland, 1997; Holmes, 2009; Elena et al., 2014). Quasispecies is defined as a large viral population where the variants are linked by very high rates of mutation which in turn means that natural selection maximizes the average fitness of the population as a whole (Holmes, 2009). Accessing new host populations puts viral populations under different selection pressures compared to the native viral host and this might allow certain neutral mutations to become beneficial and selected as the fittest. Thus, leading to the adaptation of a virus population in a previously non-susceptible host. All these characteristics increase viral evolutionary potential and genetic variation making them more likely to adapt faster to new hosts.

2.1.2. Epistasis

Epistasis is the interaction between genes or mutations that results in a specific phenotype. It can be antagonistic or synergistic depending on the background. Antagonistic epistasis occurs when mutations have a smaller combined effect than the sum of the effect of individual interactions, while synergistic epistasis occurs when the sum of combined effect of mutations is larger than the sum of their individual effect. Epistasis is the main determinant in adaptive processes as it controls the effects of the interactions of genes or mutations (Whitlock et al., 1995). It seems that antagonistic epistasis, where positive fitness effect in one host can be deleterious in an alternative one, is a major driving force behind across-host trade-offs (Elena et al., 2009). This seems as the most logical situation since in small and compacted RNA genomes, with functional secondary RNA structures, presence of overlapping genes and multifunctional proteins, an improvement of one function without disruption of another seems unlikely (Elena et al., 2014). Epistasis also becomes important when trying to explain persistence and emergence of generalist and specialist viruses. Generalist viruses are able to successfully infect hosts from different species, sometimes even distantly related ones. Specialist viruses specialize in one or very few host species and have high virulence and fitness in them (Elena et al., 2009). Since generalist viruses are able to infect various hosts equally well, we can wonder why some viruses specialize in certain hosts. The answer lies in antagonistic epistasis where the fitness of a generalist virus is limited by trade-offs in different hosts because adaptation in one host is accompanied by fitness loses in other hosts ("jack-of-all-trades" is a master of none). This limits adaptation and promotes specialization in no-host fluctuation conditions, however when hosts fluctuate, selective pressures change leading to the emergence of generalist viruses (Elena et al., 2009). Examples of these situations can often be found in nature. When a virus faces a single host such as the monocultures of various agronomically important crops (maize, rice, wheat, etc.) it will become specialized in it. Generalist viruses would find themselves in different conditions where they often have to switch between various hosts, for example in nature in a heterogeneous host environment that is visited by the same aphid vector or monocultured crops coming into contact with infected endemic plants. Therefore, it seems that generalist viruses would be more capable to cross the species boundary and infect novel hosts.

2.1.3. Recombination and segment reassortment

Recombination is an event in which two RNA viruses coinfect a single host cell and a hybrid RNA molecule is produced when RdRp jumps from one template to another. While reassortment can only occur in segmented viruses during coinfection of a single host cell where different viruses exchange genome segments and molecules with different origin get packed into a single virion (Holmes, 2009). Recombination occurs more frequently in (+)RNA viruses and retroviruses while very rarely in (–)RNA viruses. These differences are due to the different biology of each virus group and they are determined by different genomic architectures. Examples of this biology are evident in retroviruses that usually pack two RNA molecules in their virions and therefore increase the probability of different RNA molecules being packed together. In (–)RNA viruses the genomic constraints do not permit frequent recombination because negative-sense molecules are quickly bound to nucleocapsid subunits limiting the recombination events (Holmes, 2009). The evolutionary advantages of recombination in RNA viruses are still unclear but it is thought as a sort of sexual reproduction. Recombination is important because it has a major impact on RNA virus evolution and epidemiology. It has been associated with increase in virulence, host range expansion, evasion of host immunity and resistance to antivirals. All these features are important as they can lead to emergence of novel viruses able to replicate in new hosts or lead to increase of virulence causing more detrimental symptoms in current hosts (Bentley and Evans, 2018).

2.2. Extrinsic factors

Environmental factors such as, temperature changes, water changes and biodiversity changes put different selective pressures on viruses.

2.2.1. Environmental stressors

Abiotic stressors such as, water levels, temperature changes, altered CO_2 and O_3 levels, cause detrimental effects on the plant and affect the virus' replication cycle and transmission. Stressful conditions modify hormone levels, gene responses, signaling pathways and sap composition, therefore affecting plant virus replication, virulence and fitness level. Abiotic stress and viral infection activate the same signaling pathways in the plant and often interfere with one another (van Munster, 2020). In previous studies it has been shown that temperature affects symptoms expression and viral accumulation further affecting the host plant - virus interaction (Harrison, 1956; Kassanis, 1957;

Obrępalska-Stęplowska et al., 2015; Chung et al., 2016). It appears that virus accumulation is temperature-dependent and the virus lifecycle is affected by seasonal temperature fluctuations in natural hosts (Honjo et al., 2020).

2.2.2. Natural selection and genetic drift

The survival of the most fit variant or the survival of the fittest is called natural selection and it is the main driving force of viral adaptive evolution. Survival of the fittest variant depends on the environment and the genotype of the virus. Different variants of the same virus specie will have different fitness under different environmental conditions. RNA viruses have very large and heterogeneous populations and they constantly face changing environments. Therefore, existence of a heterogeneous population of genotypes gives them an advantage, because a population of different variants increases the chances of one variant having high fitness in the given environment. Positive and negative selection play an important role in selection of high and low fitness variants in the viral population. Positive selection favors variants with mutations with an adaptive value while variants with deleterious mutations get purged by negative selection (Manrubia and Lazaro, 2006). The frequency of a genetic variant in a population can also be influenced by genetic drift, or the change of the frequency of a variant in a population over generations due to chance. Genetic drift leads to random sampling of the genetic variants thus rendering a new variant composition in a population which can affect the fitness and adaptive potential of that population. Vector transmission, transmission between hosts, natural disasters or human influence can lead to genetic drift in viral populations thus changing the composition of viral populations.

2.2.3. Changes in biodiversity

Loss of biodiversity can promote disease emergence and therefore viral adaptation to a new host population. This simplification of ecosystems happens often with agricultural practices and leads to increased incidence of new emerging diseases that sometimes leads to epidemics (Pagán et al., 2012; Lacroix et al., 2014). For a virus to become emergent three ecological and genetic factors need to be satisfied: (1) the virus must come in contact with the new host, (2) the virus must adapt to the new host well enough to ensure successful replication and between host transmission and (3) epidemiological dynamics must optimize between host transmission in the new host population establishing it as a permanent pathogen (Roossinck and García-Arenal, 2015). Spillovers can happen from

wild plants where the multi-host viruses can adapt to the new host fast and spread rapidly among the same plant species. It is hypothesized that viruses with a wide host range (generalists) are more likely to emerge than viruses with a narrow host range (specialists) (Roossinck and García-Arenal, 2015).

2.3. Host factors

The host responds to the viral infection by activating defense responses which creates a very selective environment for the virus.

2.3.1 Host genetics

In the gene-for-gene hypothesis for every gene for virulence in the pathogen there is a corresponding gene for susceptibility in the susceptible plant species (Flor, 1971). Based on this hypothesis an infection can occur if the plant host is susceptible to a certain pathogen or when a pathogen has a matching virulence gene for a plant resistance gene. The plant host has a resistance (*R*) gene that matches the avirulence (*Avr*) gene in the pathogen. The mechanism of pathogen detection works on the principle of the receptor-ligand model where the R-protein-mediated recognition of *Avr* genes triggers the defense mechanisms in the plant. The most widespread family of plant resistance genes which has hundreds of protein variants in *A. thaliana* is the NB-LRR family of proteins. They are adaptable surveillance molecules that can recognize rapidly evolving pathogens amongst other functions (Van Der Biezen and Jones, 1998). The recognition of host surveillance molecules puts high selection pressures on viruses and they change avirulence genes in order to avoid being recognized by the host.

3. Plant response to viral infection

3.1. Importance of plant defense

Plants are one of the most important food sources on the planet. They use sunlight and CO_2 to produce complex carbon-containing molecules that animals eat, providing a food source for humans in the form of plant-based and animal-based foods. Great losses in plant production caused by plant disease are quite devastating on the ecosystem. One of the major causes for such destructive changes in the plant community are viruses. Plants have evolved mechanisms of defense against pathogens which restrict viral replication

and movement, such as: gene silencing, NB-LRR domain-containing resistance proteins which recognize viral effectors and activate effector-triggered immunity, ubiquitin–proteasome pathway, hormone-mediated defense and metabolism regulation (Calil and Fontes, 2016). These mechanisms are often counter attacked by co-evolving viral suppressors that enhance viral pathogenicity in a continuous coevolutionary arms race for dominance. This evolutionary race between plant hosts and viruses can lead to changes in the viral genome that might grant novel advantages to the virus. This may cause an expansion of the virus' host niche, which can lead to epidemics in previously naive hosts. Therefore, the continuous evolutionary competition between plant immunity responses and viral suppressors is a constant threat to agriculture and demands further studies.

3.2. Different types of plant responses to viruses

3.2.1. Effector triggered immunity

The first line of defense against viruses are receptors that recognize specific viral molecular patterns. There are two types of receptors, first ones called pattern recognition receptors (PRRs) triggered by perception of pathogen-associated molecular patterns (PAMP) on the plasma membrane. If activated by viral molecular patterns they trigger a PAMP-triggered immunity (PTI) that activates mitogen-activated protein kinase (MAPK), callose deposition at the cell wall, salicylic acid (SA) accumulation and expression of defense related genes (Zhang and Zhou, 2010). PRRs are represented by two types of proteins on the cell membrane, receptor-like kinases (RLKs) and receptor-like proteins (RLPs) that often require a co-receptor to initiate signaling. Other components of the PTI signaling pathway have also been shown to play an important role in antiviral defense: brassinosteroid insensitive 1 (BRI1)-associated kinase 1 (BAK1) which acts as a positive regulator of plant defense and MAPK4 acting as a negative regulator (Yang et al., 2010; Liu et al., 2011).

Second type of receptors are resistance proteins (R) that act as intracellular immune receptors which recognize virulence effectors secreted by the pathogens and activate a defense response (Zhang and Zhou, 2010; Calil and Fontes, 2016). Majority of the resistance proteins involved in antiviral resistance belong to the coiled-coil (CC)-NB-LRR or Toll/interleukin-1 receptor (TIR)-NBS-LRR class and have been identified in tobacco, tomato, cucumber, potato and arabidopsis (Zhu et al., 2013). They activate

defense response-associated genes, oxygen species (ROS) production, MAPK activation, SA accumulation and calcium ion influx. Defense in resistance genes gets triggered when pathogen-encoded Avr factor is recognized by a plant *R* gene protein. Activation of resistance proteins often leads to hypersensitive response (HR) where infected and adjacent cells activate programmed cell death to restrict the pathogen to the primary infection site. Symptomatic manifestations of the local HR response are chlorotic or necrotic lesions and spots on leaves, stems and fruits of the plant (Mandadi and Scholthof, 2013). Much later in the infection and presumably after the local HR response fails to limit virus spread, systemic necrosis is activated and primarily manifested in the upper non-inoculated tissues. The difference between the HR-induced necrosis and systemic necrosis is that the latter is a lethal response that can ultimately lead to plant death.

After the HR response, systemic acquired resistance (SAR) is often triggered by an interaction of Avr and R proteins leading to the activation of defense signaling at distant tissues. This often leads to the accumulation of SA and jasmonic acid (JA) in the plant leading to gene expression changes. SAR is a long-lasting immune response meant to provide resistance from future infections (Mandadi and Scholthof, 2013). The long lasting effect of SAR is maintained through DNA methylation and chromatin remodeling (Spoel and Dong, 2012). Another form of induced resistance that renders uninfected parts of the plant more resistant to pathogens, similar to SAR, is induced systemic resistance (ISR). Unlike SAR that is induced by pathogens and insects, ISR is potentiated by beneficial microbes living in the rhizosphere, like bacteria and fungi that promote plant growth. ISR depends on the pathways regulated by JA and ethylene (ET), which are different from those activated in SAR (Choudhary et al., 2007).

3.2.2. RNA silencing

RNA silencing or RNA interference (RNAi) is an antiviral defense mechanism in plants in which viruses are both inducers and targets. It is triggered by viral dsRNA or hairpin RNA (hpRNA) which are formed during the viral replication process. Both types of trigger RNAs are processed by Dicer-like (DCL) and Argonaute (AGO) family proteins into small 20-24 nucleotide (nt) RNA (sRNA) duplex with 2-nt 3' overhangs at both ends. One strand of the sRNA duplex becomes the guide RNA forming the RNA-induced silencing complex (RISC) with AGO. This complex binds to the complementary region of the viral ssRNA where AGO cleaves it at the overlapping central region (Guo et al.,

2016). RNA silencing pathways have diversified their roles in the plant and there are multiple copies of DCL's, AGO's and other factors. Arabidopsis has 4 DCL proteins and 10 AGO proteins involved in a variety of pathways, such as microRNA (miRNA) pathway, trans-acting small interfering RNA (tasiRNA) pathway, RNA-directed DNA methylation pathway and exogenic RNA silencing pathway. DCL's 2 and 4 perform the majority of the viral RNA processing in the plant resulting in the overrepresentation of 21 and 22-nt siRNAs (Guo et al., 2016). As a counter defense measure to plant RNAi, well adapted plant viruses encode silencing-suppressor proteins (Wieczorek and Obrępalska-Stęplowska, 2015). These silencing-suppressor proteins inhibit RNAi at various steps of the pathway, for example inhibiting DCL proteins and the co-factors, destabilizing AGO or sequestering siRNA. One of the best characterized silencing suppressors is the potyviral helper component proteinase (HC-Pro), a protein with very diverse silencing suppressor activities such as, ds-siRNA binding, blocking of primary siRNA biogenesis and downregulation of RISC components (Calil and Fontes, 2016).

3.2.3. Role of plant hormones in plant defense

Plant hormones are important regulators of defense responses as well as responses to abiotic and biotic stresses, development and signaling. Viruses can manipulate hormone signaling for their benefit through different molecular interactions. Key players in the defense response to pathogens are SA, JA, brassinosteroids (BR), ET, abscisic acid (ABA), auxin, gibberellic acid (GA) and cytokinin (CK) (Bari and Jones, 2009). SA is a key player in the plant response to pathogens and establishes local and systemic resistance (Pieterse et al., 2012; Vlot et al., 2009). Both DNA and RNA viruses activate the SA pathway. In transgenic lines deficient in SA accumulation there is unrestricted viral spread and development of disease symptoms because a lack of SA accumulation leads to a delayed activation of defense genes and no SAR (Baebler et al., 2014). SA also enhances RNAi and activates DCL's thus inducing resistance to TMV (Campos et al., 2014). On the other hand, the role of JA is controversial because it can suppress viral infection or aid it. Exogenously applied treatment of JA to arabidopsis disrupted geminivirus infection showing that suppression of JA is crucial for infection (Lozano-Durán et al., 2011). While in *Nicotiana tabacum* plants exogenously applied methyl jasmonate (MeJA) permitted systemic viral movement by reducing the plants resistance to TMV (Oka et al., 2013). BRs help plants induce defense responses when infected with viruses. Tobacco plants treated with a brassionsteroid brassinolide exhibited enhanced

resistance to TMV and did not show SA accumulation suggesting that BL-induced resistance is different from SAR (Nakashita et al., 2003). ET alters plants susceptibility to viruses as shown in the study by Fischer and Dröge-Laser (2004) where overexpression of *NtERF5* (ET-responsive transcription factor) showed reduced size of local HR lesions and impaired systemic spread of TMV. While TuMV suppressed the defense response by disrupting the ET pathway, where NIa-Pro (nuclear inclusion a-protease domain) suppressed aphid-induced callose formation in an ET-dependent manner (Casteel et al., 2015). ABA also plays a key role in plant response to different abiotic and biotic stresses. Exogenous application of ABA reduced systemic accumulation of TMV, while disruption of the ABA pathway accelerated systemic accumulation of TMV in arabidopsis (Chen et al., 2013). ABA is also involved in callose deposition on plasmodesmata which may restrict cell to cell movement of viruses and enhance resistance (Mauch-Mani and Mauch, 2005). Auxins appear to have an important role in viral infection, an example is the replicase protein of TMV that interacts with auxin proteins and leads to modifications in auxin gene regulation (Padmanabhan et al., 2005). Cytokinin role was proved in an experiment with geminiviruses where geminiviral AC2/AL2 protein interacted with an adenosine kinase in arabidopsis leading to an increased expression of cytokinin responsive genes (Baliji et al., 2010). P2 protein of *Rice dwarf phytoreovirus* interacts with a key factor in the biosynthesis of gibberellins leading to a dwarf phenotype in rice (Zhu et al., 2005).

3.2.4. Ubiquitin proteasome complex

An important component of the plant and animal viral defense is the ubiquitin proteasome complex (UPS). It has a dual role in infection by aiding viruses to establish a successful infection or in defense by eliminating viral components (Alcaide-Loridan and Jupin, 2012). Most important roles of UPS in the cell are the regulation of the cell cycle, transcription, cell death, development and signal transduction (Hershko and Ciechanover, 1998). The main component of UPS is ubiquitin (Ub), whose attachment to cell proteins regulates protein homeostasis and regulation of signaling pathways. Proteins that control the activation and transfer of ubiquitin are E1 (Ub-activating enzyme), E2 (Ub-conjugating enzyme) and E3 ligase (Alcaide-Loridan and Jupin, 2012). Various plant viruses have evolved proteins that interact with UPS components. For example, potyvirus' HC-Pro interacts directly with subunits of the proteasome thus inhibiting them and increasing their viral load and symptoms (Jin et al., 2007; Dielen et al., 2011; Sahana

et al., 2012). An interesting experiment using turnip yellow mosaic virus (TYMV) shows how a virus can develop counter measures against plant UPS defenses. The RdRP of TYMV gets targeted and degraded by UPS in infected cells creating an impact on the infection rate, but the virus in turn stabilizes the RdRP by using a viral deubiquitinating enzyme (DUB), thus promoting infection (Chenon et al., 2012). Another example of plant viruses using proteasome machinery to promote virulence can be seen in TMV and TuMV, where downregulation of *RPM9*, a 26S proteasome subunit, inhibits systemic spread of the two viruses (Jin et al., 2006).

4. Genetic and environmental robustness

Genetic robustness refers to the constancy of the phenotype in the face of heritable perturbations (genetic or epigenetic) (Visser et al., 2003). Environmental robustness is the buffering against non-heritable perturbations such as external stressors (heat, light changes) or developmental noise. Main factors influencing robustness are: (1) large population sizes where robustness acts at the population level and preserves the invariance of the phenotype and (2) high mutational rates where robustness increases the tolerance of the viral genome to mutations. One way to buffer the effect of each new mutation is to become robust thus dealing with genome instability while also generating huge population sizes. High mutational pressures favor mechanisms that promote mutational robustness in RNA viruses (Montville et al., 2005; Codoñer et al., 2006; Sanjuán et al., 2007; Stern et al., 2014; Thyagarajan and Bloom, 2014; Visher et al., 2016). Considering mutations and large viral population sizes which contribute to viral evolvability or the capacity of a virus to increase its fitness through adaptation, will robustness promote viral evolvability or not? Since robustness buffers mutational and external effects on viral phenotype this can reduce phenotypic variation and the effect of natural selection acting on it. However, robustness can also lead to an increase in genetic variation which may lead to new epistatic interactions thus increasing the range of possible adaptive phenotypes available (Lauring et al., 2013). In a study by Draghi et al. (2010), it was shown that neutral diversity in a robust population accelerates adaptation if the number of phenotypes accessible through mutation is smaller than the total number of phenotypes in the fitness landscape. Robustness is an essential fitness component in RNA viruses because of their small and compact genome size, high mutation rates and ever-changing environmental conditions. Since robust phenotypes arise under the

selective pressures of highly deleterious mutation rates and changing environmental conditions, they can be an important aspect of virus adaptation and survival. Being robust is also important for pathogenicity of viruses because they find themselves in a range of environmental conditions and having an assortment of preadapted variants that are close to the fitness optimum might be advantageous.

5. Methods to measure viral adaptation and the host response

All the mechanisms of virus adaptation and evolution along with plant defense are main forces acting on viral population leading to more/less pathogenic viruses or the emergence/extinction of new virus species. Some of the basic mechanisms of these interactions have been described but general knowledge is still lacking. Another problem is that different virus families can induce different defense responses and symptomatology in the plant. So, conclusions extrapolated by studying a member of one viral family might not hold true for the rest of them even though the general principles can apply. In order to fully understand how virus adapt and cause disease in plant hosts we also need to pay attention to their evolutionary history, host niche and their resilience to environmental changes. Are the viruses well adapted to the host or not? Do they have a wide or narrow host niche and can they handle external changes well? How do all these different viruses affect the host response and defense? All these questions are of serious concern and need more thorough answers. Luckily with the development of new methods this is becoming a feasible task. A great way to measure how viruses with different evolutionary histories affect plant immune response and which genes can possibly have important roles in disease development is genome-wide association studies (GWAS). In order to answer how external changes, for example temperature, affect viruses with different evolutionary histories we can measure their robustness using mutagens. Let's describe these two approaches more closely.

5.1. GWAS

GWAS are becoming increasingly popular over the last 20 years (Bush and Moore, 2012). One of the reasons behind this is that large-scale sequencing is becoming more financially feasible and more organisms are becoming genotyped (Cantor et al., 2010). Thus, making GWAS a powerful method that connects the phenotype and the genotype allowing us to

predict genetic risk factors for disease as well as important agronomic traits, such as viral infections (Korte and Farlow, 2013).

During the genotyping process we are trying to capture most of the single nucleotide polymorphisms (SNPs) that act as units of genetic variation between individuals of the same species. Some SNPs can have no biological impact on the organism but some can have functional consequences such as amino acid changes, transcription changes and binding affinity changes. Single nucleotide polymorphisms are more commonly occurring than genetic variants implicated in rare genetic disorders (cystic fibrosis). Therefore, in the literature it is referred to common base pair changes when talking about SNPs and to rare changes when talking about mutations. This leads to the common disease/common variant hypothesis that is behind the GWAS logic. This hypothesis states that common disorders are influenced by common genetic changes in the population (Bush and Moore, 2012).

One of the most popular model organisms that has vast genome data available with a large number of different individuals is the plant *Arabidopsis thaliana* (L.) Heynh. As part of the 1001 genome project there is a detailed variation map available of 1,135 natural inbred lines from Eurasia, North Africa and North America (1001 Genomes Consortium, 2016). The main advantages are the inbred lines maintained by self-fertilization that make it possible to maintain a phenotype from genetically identical individuals (Korte and Farlow, 2013). Arabidopsis is also small in size and has a short generation time making it faster and easier to work with. Despite all the positive attributes the arabidopsis system has some problems, the main one being a common problem in GWAS; the polygenic nature of many traits measured. This can be circumvented by increasing the sample size and therefore improving the power to recover meaningful associations. When designing the experiment, one has to take the sample size into account and the geographical distribution. Analyzing geographically distant accessions might solve the problem of the polygenic effect but can also introduce genetic heterogeneity. Population structure or genetic heterogeneity refers to variants that are more related because they are geographically close to each other. These variants form subpopulations that have fixed certain genetic variants that differ compared to the variants of another subpopulation. This leads to population stratification, which was proven to be a problem in arabidopsis where a non-causative marker can prove to be a better descriptor of the phenotype than a

causative one (Bush and Moore, 2012; Korte and Farlow, 2013). Fortunately, this problem can be solved by using liner mixed models where the covariates and the SNPs are modelled as fixed effects and the population structure as a random effect (Lippert et al., 2014). Population structure in the mixed model is dealt with by estimating the phenotypic covariance that is due to genetic relatedness between individuals (Korte et al., 2012; Segura et al., 2012; Lippert et al., 2014;). When deciding which linear mixed model to choose (EMMAX, GenABEL, FaST-LMM, Mendel, GEMMA, or MMM) it appears that it does not matter too much since they are all in concordance with the results. The choice of the precise program to use should be made based on speed and convenience (Eu-ahsunthornwattana et al., 2014).

When analyzing results, we often wonder how to decide which associations are true positives. The most common method to select true positives is the 5% Bonferroni correction. But this method is too stringent since GWAS studies have a large number of SNPs that are being analyzed at the same time which leads to the multiple testing problem. Because the assumption that each test performed is independent of others is often not true due to linkage disequilibrium between genetic markers, we may have a problem of false negatives and might miss out on an important gene related to our phenotype (Bush and Moore, 2012; Korte and Farlow, 2013). A threshold that deals better with false positives and false negatives is the false discovery rate (FDR) that is an estimate of the proportion of significant results that are false positives (Bush and Moore, 2012).

Out of the quite large number of GWAS studies performed so far, only a few have been focused on the study of plant - virus pathosystems. These studies looked at genetic determinants of crop plants, such as maize, wheat, soybean, and pepper, in correlation with infections with *Maize rough dwarf fijivirus*, *Barley yellow dwarf luteovirus*, *Tobacco ringspot nepovirus*, and *Potato potyvirus Y* (Chen et al., 2015; Chang et al., 2016; Choudhury et al., 2019; Tamisier et al., 2020). In addition, two studies have worked with arabidopsis and *Turnip mosaic potyvirus* (Rubio et al., 2019) and *Cucumber mosaic cucumovirus* (Montes et al., 2021). Alas, the number of genotypes used in these studies was limited, smaller than 200.

In the light of the emerging interest for GWAS studies that are mainly focused on the phenotypic differences in multicellular organisms and the lack of GWAS focusing on

plant - virus pathosystem, we wanted to focus on the virological aspect and identify host genes that are involved in the virus infection.

5.2. Mutagenesis and environmental fluctuations

Many studies done so far have shown that robustness allows an increase in genetic diversity for the viral populations while maintaining their phenotype. These studies were performed by quantifying the mutational fitness effect of point mutations on the replicative efficiency of the viruses (Sanjuán et al., 2004; Domingo-Calap et al., 2009; Cuevas et al., 2012) or by using mutagens such as nucleoside analogues to quantify the sensitivity of the viruses (Graci et al., 2012; Willemsen et al., 2018). Viruses that after a treatment with a mutagen exhibit a small effect on fitness are considered mutationally robust, whereas those exhibiting a large effect are considered fragile (or brittle) (Lauring et al., 2013).

Environmental robustness can be measured as the persistence of the viral phenotype/fitness in the face of environmental changes. For example, causing temperature changes during a certain time period we can observe if there will be changes in the viral phenotype. Again, viruses that have a small change in their phenotype after the perturbating treatment will be robust, while those that display a large change are brittle.

6. The studied pathosystem

We decided to focus on a pathogen that belongs to a virus family that is widespread in cultivated and wild plants around the globe (Ivanov et al., 2014). The virus we decided to use in these studies is turnip mosaic virus (TuMV; species *Turnip mosaic potyvirus*, genus *Potyvirus*, family *Potyviridae*, order *Patatavirales*, class *Stelpaviricetes*, phylum *Pisuviricota*, kingdom *Orthornavirae*, realm *Riboviria*), a (+)ssRNA plant virus. Viruses in the *Potyviridae* are non-enveloped, filamentous and approximately 680-900 nm long with a single core capsid protein. Potyviruses encode a large polyprotein that is self-cleaved into distinct functional proteins (Figure 2).

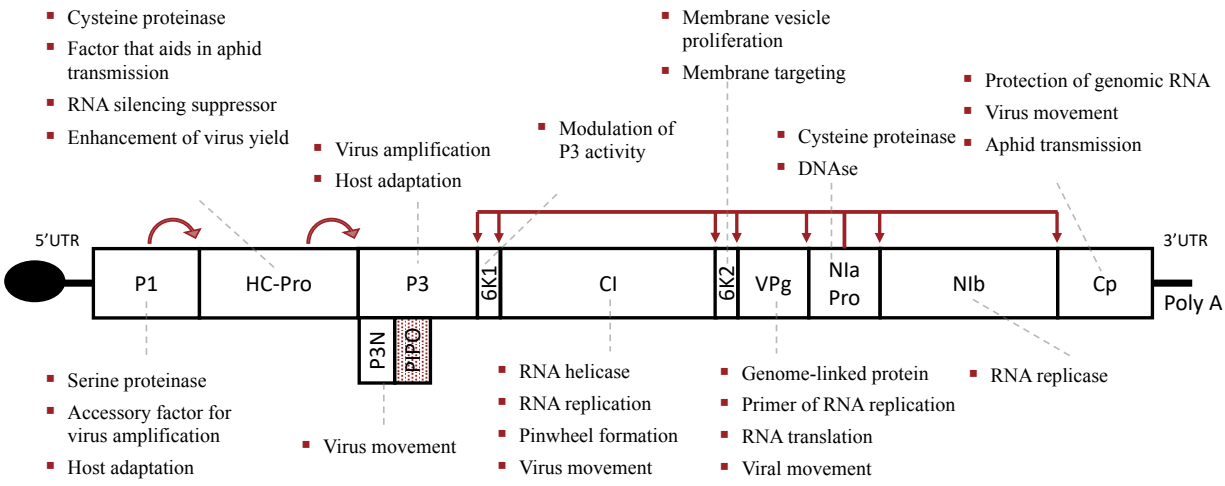


Figure 2. The genome organization of potyviruses and the function of each cleaved protein. Figure adapted from Revers and Garcia (2015).

Potyviridae are divided into 12 genera with 228 species: *Arepavirus*, *Bevemovirus*, *Brambyvirus*, *Bymovirus*, *Celavirus*, *Ipomovirus*, *Macluravirus*, *Poacevirus*, *Potyvirus*, *Roymovirus*, *Rymovirus*, and *Tritimovirus* plus three unassigned species (https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/271/potyviridae). Potyviruses are grouped in different genera based on their sequence similarity, genome composition and vector transmission. Viruses in eleven genera have a monopartite genome except for viruses belonging to the genus *Bymovirus* that have a bipartite genome.

TuMV has a genome size of 9.84 kb and is expressed as a polyprotein where three viral-encoded proteinases (Pl, HC-Pro and NIa-Pro) subsequently cleave it into ten proteins. In addition, there is a smaller ORF within the P3 cistron that is translated in the +2 reading-frame resulting in the P3N-PIPO peptide (Chung et al., 2008). It is geographically widespread by more than 40 different aphid species and capable of infecting various plant hosts, though mostly belonging to the *Brassicaceae* family, causing diverse symptoms like vein mottling, chlorosis, mosaic, necrosis, sterilization and plant death (Guerret et al., 2017). TuMV is one of the most important viruses affecting economically important vegetables and crops (Tomlinson, 1987; Ohshima et al., 2002; Yasaka et al., 2017). It also has a high incidence in wild populations of arabidopsis (Pagán et al., 2010), where the TuMV disease progression and symptomatology is well described. This is important

because arabidopsis is undoubtedly one of the most suitable organisms for GWAS and other experimental studies. It has over 1000 natural accessions genotyped and described so far from Eurasia, North America and North Africa (1001 Genomes Consortium, 2016). Genotypes can be maintained by self-fertilization for an unlimited number of generations, facilitating GWAS and making phenotypization highly reproducible (Korte and Farlow, 2013). In conclusion, the availability of data and a good characterization of the TuMV - arabidopsis pathosystem makes experimental work and interpretation of data easier.

RNA viruses have a great ability to adapt to new hosts, they have high mutation rates, large population sizes and short generation times which leads to high evolutionary potential making them major pathogens responsible for emerging disease (Bordería et al., 2011; Carrasco-Hernández et al., 2017). All these characteristics of RNA viruses can cause large economic losses and difficulties in control and prevention and should be further studied.

Objectives

The main goal of this thesis is to characterize the differences between viral adaptation histories and their interaction with the host or the environment:

1. Map different host response genes to a generalist or a specialist virus.

2. Identify specific genes involved in the response to adapted or naïve virus strains.

3. Evaluate genetic and environmental robustness of an RNA virus with different adaptation histories.

Methods

1. GWA analysis

1.1. Plant material and growth conditions

Four hundred and fifty and 1050 arabidopsis accessions (Supplementary Table S1 and S2) from the 1001 arabidopsis genome collection (https://1001genomes.org; 1001 Genomes Consortium, 2016) were phenotyped in two separate studies. The accessions were representative of the global species distribution. To ensure all the accessions were at a similar growth stage and to reduce the noise that large differences in vegetative development could cause, we confirmed that all selected accessions reached growth stage 3.2 - 3.5 in Boyes et al. (2001) scale ~21 days after germination in our experimental growth conditions [16 h day/8 h night with temperature of 24 ºC day/20 ºC night, 45% relative humidity and 125 µmol m−2s−1 of light intensity (1:3 mixture of 450 nm blue and 670 nm purple LEDs)].

1.2. Virus inoculum

The two strains of TuMV used in Chapter 1 were obtained after twelve passages of experimental evolution in mutant genotypes of the arabidopsis Col-0 accession, as detailed in Navarro et al. (2020). Among all the resulting viral lineages, lineage L4 evolved in the *enhanced disease susceptibility 8* (*eds8-1*) mutant, hereafter referred as TuMV-G, and lineage L4 evolved in the *jasmonate insensitive 1* (*jin1*) mutant, referred as TuMV-S, showed strikingly different host ranges. The *eds8-1* plants lacked the EDS8 protein, causing the reduction of the expression of plant defensin genes and reduced ISR but enhanced SAR. The *jin1* plants lacked the JIN1 protein, causing the loss JA signaling which is a negative regulator of SA-dependent signaling. This results in a constitute expression of SAR. The *eds8-1* plants turned out to be the most resistant ones to TuMV infection while the *jin1* plants were the most susceptible ones. TuMV-G was able to infect all tested plant genotypes with equal fitness, while TuMV-S infected only *jin1* well. Indeed, Navarro et al. (2020) calculated Blüthgen's *d'* specialization indexes (Blüthgen, et al., 2006) for these two strains, finding that TuMV-G had $d' = 0$ (no specialization) while TuMV-S had $d' = 1$ (complete specialization). In agreement with previous potyvirus-arabidopsis studies (Hillung et al. 2014; González et al., 2019), more permissive hosts (here *jin1*) selected for more specialized viruses while more restrictive hosts (in this case, *eds8-1*) selected for more generalist viruses. At the genomic level,

TuMV-G and -S differed in a total of seven point-mutations (Navarro et al. 2020). Relative to the ancestral naïve TuMV strain, TuMV-G contains three nonsynonymous mutations, all affecting the VPg protein (H33Y, D113N and K121E). Likewise, TuVM-S has two synonymous mutations (HC-Pro/C1760U and P3/U3269C) and two nonsynonymous ones (VPg/R118H and CP/S70N).

In Chapter 2, TuMV-AS was obtained from infected *Nicotiana benthamiana* Domin plants inoculated with a transcript product from a p35STunos infectious plasmid that contains TuMV genome cDNA (GenBank accession AF530055.2), corresponding to YC5 isolate from calla lily. This cDNA was under the control of the *Cauliflower mosaic caulimovirus* 35S promoter and a NOS terminator. TuMV-DV was obtained after twelve passages of TuMV-AS in arabidopsis accession Col-0.

TuMV-G and -S, and TuMV-AS and TuMV-DV infected plant tissues was frozen in liquid N_2 and homogenized and mixed with 10 volumes of inoculation buffer (50 mM KH_2PO_4 pH 7.0, 3% polyethylene glycol 6000, 10% Carborundum) right before the mechanical inoculations. The two TuMV strains were mechanically inoculated into healthy arabidopsis plants that were between 21 - 25 days old. The inoculation started from the plants that were the largest (8 - 12 leaves) giving the smaller plants extra time to grow so all the accessions got inoculated at a similar size (Boyes' 3.2 - 3.5). Three middle sized leaves were mechanically inoculated with 5 μl of infectious sap prepared in inoculation buffer. To further minimize differences due to inoculation efficiency all the inoculations were done by the same researcher. Hence, we assume that the inoculation failure rate would be the same among all accessions.

1.3. Inoculation procedure

Eight plants per accession for each TuMV strain were inoculated, resulting in a total of 16 plants phenotyped and two mock-inoculated control plants per accession. Accessions were split into two blocks in Chapter 1 and into four blocks in Chapter 2, because of chamber space and workforce capacity. The inoculation procedure took about 3 - 4 days per block, where in consecutive days different accessions underwent the inoculation procedure because (1) it was not possible to inoculate all the plants in the same day due to the sheer number of them and (2) this way all the plants got synchronized in size at the moment of inoculation.

In the GWAS of Chapter 1 the first block was inoculated from 2019/05/06 to 2019/05/08 and the second block was inoculated from 2019/09/11 to 2019/09/14. Three hundred and 150 accessions were inoculated in each block, respectively. Pot trays contained four accessions inoculated with each viral strain along with their corresponding mocks. To reduce spatial correlations due the relative position of plants in the growth chamber, pots trays were translocated to a new random position every day.

In the GWAS of Chapter 2 all the different accessions were inoculated and phenotyped in four independent blocks consisting of 4800 plants each. The first block was inoculated on 2018/11/27 - 30, the second block on 2019/2/6 - 10, the third block on 2019/3/13 - 16, and the fourth block on 2019/6/4 - 6. The accessions in the fourth block that did not reach the proper size on the day of the inoculation were inoculated a few days later (2019/6/10). Eight plants per accession were inoculated with TuMV-DV and another eight with TuMV-AS, along with two mock-inoculated plants that served as negative controls of infection. Four accessions inoculated with each viral strain and the corresponding mocks were placed in the same tray. Pots trays were also translocated to a new random position every day to reduce spatial correlations.

A replica of all the 51 necrotic and 67 random non-necrotic accessions (118 in total, Supplementary Table S3) from the GWA analysis was done in order to analyze more closely the large peak on chromosome 2 related to the necrosis and symptom severity phenotype. This study was performed in one block with 8 inoculated plants per accessions and viral strain, along with 2 mock-inoculated plants per accession that served as a control for symptomatology. All the plants were grown under the same conditions as mentioned above.

Col-0 loss-of-function (LOF) mutant genotypes that were used to confirm both GWAS results (Table M2, M3) were seeded on 2020/06/03 and inoculated, as described above, with the two TuMV strains on 2020/06/23. All LOF mutants and wild-type (WT) control plants were analyzed in one block in the same growth chamber with 10 plants per virus and per genotype and two mocks per combination.

Table M2. Selected LOF mutants in the Chapter 1 study, with a description of the corresponding genes functions and a link to reference.

MUTANT ID	*description*	*papers suggesting virus infection function*	*doi*
AT1G57570	Mannose-binding lectin superfamily protein	mannose-binding lectin protein that is involved in pathogen recognition and is a part of plant innate immunity	10.3389/fpls.2014.00397
AT2G04430	nudix hydrolase homolog 5	AtNUDX6 was involved in the plant immune response as a positive regulator of NPR1-dependent SA signaling pathways by modulating NADH levels	https://doi.org/10.1080/09168451.2014.987207
AT2G04450	nudix hydrolase homolog 6	AtNUDX6 was involved in the plant immune response as a positive regulator of NPR1-dependent SA signaling pathways by modulating NADH levels	https://doi.org/10.1080/09168451.2014.987207
AT2G14080	Disease resistance protein (TIR-NBS-LRR class) family	involved in plant defense	10.1038/ni1410
AT3G12850	COP9 signalosome complex-related / CSN complex-like protein	geminiviral C2 protein interacts with CSN5 resulting in a reduction of JA levels and it has been seen that treating A. thaliana plants with exogenous jasmonate disrupts geminivirus infection	10.1105/tpc.110.080267
AT4G10130	DNAJ heat shock N-terminal domain-containing protein	involved in peptidyl-diphthamide biosynthetic process from peptidyl-histidine and tRNA wobble uridine modification	TAIR
AT4G13345	MATERNAL EFFECT EMBRYO ARREST 55, Serine-domain containing serine and sphingolipid biosynthesis protein	Sphingolipids are involved in plant defense and cell death	10.3389/fpls.2012.00068
AT5G08650	Small GTP-binding protein	CPLEPA is a chloroplast translation factor that is essential under suboptimal conditions. Could aid viral translation in the chloroplast	10.1111/mpp.12533
AT5G66750	Protein is similar to SWI2/SNF2 chromatin remodeling proteins. DDM1 is appears to act as a chromatin-remodeling ATPase involved in cytosine methylation in CG	DDM1 deficient mutants showed resistance to TuMV because SA-signaling is important for TuMV-response and hypomethylated mutants induce SA-mediated defense pathways	https://doi.org/10.1093/molbev/msaa091

	and non-CG contexts. Involved in gene silencing and maintenance of DNA methylation and histone methylation. Hypomethylation of many genomic regions occurs in ddm1 mutants, and can cause several phenotypic abnormalities, but some loci, such as BONSAI (At1g73177) can be hypermethylated in ddm1 mutants after several generations, leading to different phenotypes. DDM1 might be involved in establishing a heterochromain boundary. A line expressing an RNAi targeted against DDM1 shows some resistance to agrobacterium-mediated root transformation.		
AT4G02580	NADH-ubiquinone oxidoreductase	Ubiquinone Oxidoreductase Serves as a Susceptibility Factor to Promote Pathogenesis of Rhizoctonia solani in Plants	10.1094/PHYTO-02-19-0055-R

Table M3. Selected LOF mutants in the Chapter 2 study, with a description of the corresponding genes functions and a link to reference.

		papers suggesting virus infection	
MUTANT ID	*description*	*function*	*doi*
AT1G67160	Member of a family of proteins containing an F-box domain at the N-terminal region and three kelch repeats at the C-terminal region. Involved in BR signaling. Co-suppressed KIB1,2,3,4 lines have a dwarf phenotype and resemble BR receptor mutants.	Exogenous applied BRs enhanced plant resistance to virus infection, while application of Bikinin (inhibitor of glycogen synthase kinase-3), which activated BR signaling, increased virus susceptibility.	10.1038/srep20579

AT2G14080	Disease resistance protein (TIR-NBS-LRR class) family	involved in plant defense	10.1038/ni1410
AT2G14120	Encodes a dynamin related protein. DRPs are self-assembling GTPase involved in fission and fusion of membranes. DRP3B functions in mitochondrion and peroxisome fission in combination with DRP3A.	Treatment of plant leaves with a dynamin-specific inhibitor disrupts the delivery of VPg and CI to endocytic structures and suppresses TuMV replication and intercellular movement.	10.1128/JVI.01320-18
AT2G14170	Methylmalonate-semialdehyde dehydrogenase	virus decreases the ALDH expression	https://doi.org/10.1371/journal.pone.0032153
AT2G15320	Leucine-rich repeat (LRR) family protein	NBS-LRR	10.3390/ijms14047302
AT2G19270	mitotic checkpoint protein PRCC-carboxy-term protein	targeted by geminiviruses	10.1104/pp.108.121038
AT3G56560	NAC domain containing protein 65	A NAC Domain Protein Interacts with Tomato leaf curl virus Replication Accessory Protein and Enhances Viral Replication	10.1105/tpc.104.027235
AT5G40450	Encodes a member of a plant gene family, APK_ORTHOMCL5144, of unknown function. RBB1 is localized to the cytosol and involved in vacuolar biogenesis and organization. RBB1 mutants have increased number of vacuolar bulbs and fewer trans-vacuolar strands.	Vacuoles are involved in defense against pathogens and can trigger hypersensitive cell death.	10.1038/cdd.2011.70
AT5G45770	receptor like protein 55	RLP involved in plant defense	10.1104/pp.108.119487

1.4. Phenotyping

In Chapter 1 study, three phenotypic traits were measured: (1) Symptoms severity: on a scale from 0 - 5 (Fig. M1) measured at intermediate (14 days post-inoculation - dpi) and late (21 dpi) infection times so as to explore time-dependent differences in gene expression.

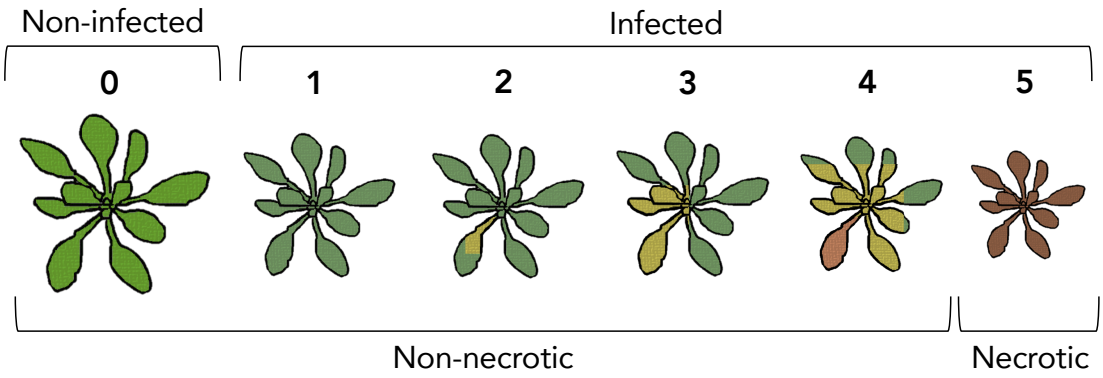


Figure M1. Symptoms scale that was used to evaluate the severity of symptoms in the plants during the 21 days period post inoculation. 0: no symptoms or healthy plant, 1: mild symptoms without chlorosis, 2: chlorosis is visible, 3: advanced chlorosis, 4: strong chlorotic symptoms and beginning of necrosis, 5: clear necrosis and death of the plant.

Therefore, the status of plant infection was assessed by visual inspection for symptoms. The intensity of symptoms was also visually quantified, as the degree of damage TuMV causes correlates with its detrimental effects on the host. For a plant, the intensity of symptoms is an evolutionarily relevant trait, since the degree of damage on the vegetative plant organs and fruit development directly impacts its fitness. (2) Infectivity: number of infected plants out of the total number of inoculated plants after 21 dpi. (3) Disease progression, calculated as the area under the disease progression stairs (*AUDPS*) (Simko and Piepho 2012). The number of infected plants was quantified daily and these values were used to calculate the progression of disease through time.

In the characterization of LOF mutant response to infection in Chapter 1, *AUDPS* and symptoms intensity progression steps curve (*AUSIPS*) (Kone et al. 2017) were measured. *AUSIPS* is calculated using daily symptoms intensity values and, similar to *AUDPS* that summarizes disease progression, it summarizes the progression of the symptomatology through time.

In Chapter 2, five disease-related traits were measured daily during 21 dpi, time when the infection reached a steady plateau. *AUDPS*, infectivity, symptomatology (on a semi-quantitative scale ranging from 0 - 5, Fig. M1), necrosis (binary trait; 0 meant no necrosis and 1 necrosis), and resistance (binary trait; 0 meant none of the plants showed symptoms of infection and 1 obvious symptoms of infection in at least one plant). The intensity of

symptoms and the number of infected plants was visually quantified for 21 dpi. Symptom severity or degree of damage TuMV causes to the plant is an evolutionarily relevant trait, since the degree of damage directly impacts the fitness of the plant.

In the LOF analysis, ten plants per virus per accession were inoculated. Visual inspection of plants was done daily for 21 dpi, annotating the number of infected plants each day and therefore allowing us to calculate *AUDPS*. Symptomatology of each plant was also evaluated during the 21 dpi. A similar calculation to *AUDPS* was done with the symptomatology, obtaining the *AUSIPS*.

For the replica of the 118 necrotic and non-necrotic accessions three phenotypes were measured during 21 days: (1) necrosis, 2) percentage of necrotic accessions per number of infected plants per accessions (3) percentage of necrotic accessions per total number of plants per accession, and (4) *AUSIPS*.

1.5. Genome-wide association mapping

In Chapter 1, association analyses were done with a Python program based on LIMIX (Lippert et al. 2014) written by Prof. Magnus Nordborg's group. LIMIX is a linear mixed model (LMM) that was used for single-trait analysis where SNPs and covariates were treated as fixed effects while the population structure and noise were treated as random effects. The kinship matrix (identical-by-state, IBS matrix) and the genotype data come from the Arabidopsis 1001 Genome Project (1001 Genomes Consortium 2016), consisting of the SNPs for the 1135 genome accessions plus imputed SNPs of a set of accessions that were genotyped with a 250k SNP chip. Kinship measures the degree of genetic relatedness between individuals and is used to remove confounding factors that decrease power and increase the false positive rate in GWAS.

Data normality was checked with SPSS version 25 (IBM Corp., Armonk NY, USA) and deviations from normality between the phenotypic values were observed between the two blocks, therefore the block effect was accounted for in the GWAS analysis through the covariates option in LIMIX. Untransformed phenotypic data was used in the GWAS, since for large sample sizes (450 accessions phenotyped) transformations increase the false positive rate and normalization is not recommended (Goh and Yap, 2009).

Out of ~10 million SNPs (Seren 2018), 1,815,154 had a minor allele frequency higher than 0.05 for all phenotypes. To minimize false positives due to multiple testing (type I errors), we used the false discovery rate (FDR) or the $-\log P \geq 5$ threshold, whatever value was more conservative. FDR was calculated using the fdrBH function (with q = 0.001) of the mSTEM package version 1.0 in R version 3.6.1 in RStudio version 1.2.1335. The exact FDR values used were as follows: for TuMV-G AUDPS 21 dpi FDR = 2.73×10^{-10}, infectivity 14 dpi FDR = 6.32×10^{-13} and 21 dpi FDR = 1.78×10^{-8}, and symptoms 14 dpi FDR = 9.59×10^{-10}. While for TuMV-S it was calculated only for symptoms at 14 dpi FDR = 1.49×10^{-12} and 21 dpi FDR = 1.15×10^{-9}. Manhattan and quantile-quantile (QQ) plots were drawn using rMVP package (Yin et al. 2020) in R version 3.6.1 in RStudio version 1.2.1335. They showed no detectable population structure for phenotyped infection traits (Fig. M2).

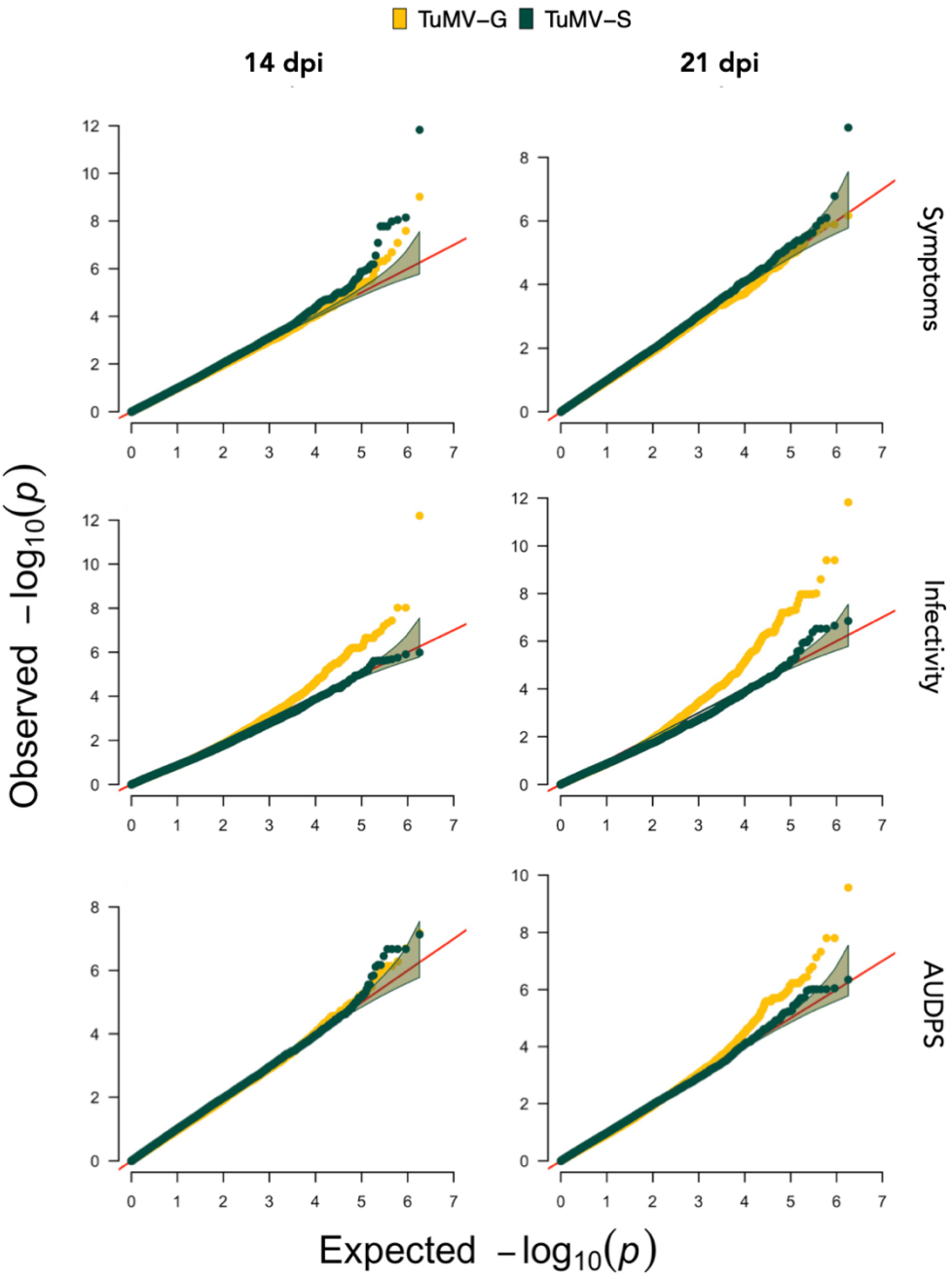


Figure M2. QQ-plots for all the infection traits studied. Results for TuMV-G infected plants are indicated in yellow while results for TuMV-S infected plants are shown in green.

Each significant SNP was tested for linkage disequilibrium (LD) within a 10 kb window by calculating r^2 with the help of PLINK 1.9 (www.cog-genomics.org/plink/1.9/). Furthermore, the r^2 results were examined for indications of any SNPs in strong LD with other significant SNP outside of the region of the significant gene. A 10 kb window was

taken because in arabidopsis LD decays rapidly within 10 kb (Gan et al., 2011; Kim et al., 2007).

In Chapter 2, data normality was checked with SPSS version 25 (IBM Corp., Armonk NY, USA) and deviations from normality between the phenotypic values were observed between the two blocks, therefore the block effect was accounted for in the GWAS analysis through the covariates option in LIMIX. Untransformed phenotypic data was used in the GWAS, since for large sample sizes (450 accessions phenotyped) transformations increase the false positive rate and normalization is not recommended (Goh and Yap, 2009). Each phenotypic trait was standardized by the block mean using a univariate general linear model in SPSS version 26 software (IBM, Armonk, NY). Then, the standardized data were analyzed with a genome-wide efficient mixed model association (GEMMA; Zhou and Stephens, 2012). SNPs with minor allele frequency less than 0.05 were excluded. The genotype data comes from the Arabidopsis 1001 Genomes Project (The 1001 Genomes Consortium, 2016), consisting of the SNPs for the 1001 genome accessions plus imputed SNPs of a set of accessions that were genotyped with a 250k SNP chip. The genotype data was downloaded from https://1001genomes.org/data/GMI-MPI/releases/v3.1/ in a VCF format. Using PLINK 1.9 (Purcell et al., 2007), the VCF file was reformatted into the PED binary format and retained only those SNPs with a 95% genotyping rate. The genotype files were also filtered to keep only the information for the 1050 accessions used. The centered relatedness matrix was computed with GEMMA. Out of ~10 million SNPs (Seren, 2018), 510,485 SNPs had a minor allele frequency > 0.05 for all phenotypes. The threshold was set at the FDR value or at $-\log P \geq 5$, whatever value was more conservative. So, the FDR was used for TuMV-AS in the trait necrosis (1.72×10^{-7}) and for TuMV-DV in the traits necrosis (3.79×10^{-7}) and resistance (1.61×10^{-7}). The FDR was calculated using the package 'fdrtool' version 1.2.15 in R version 3.6.1 in RStudio version 1.2.1335. All the quantile-quantile (QQ) plots were examined for genome-wide inflation of significance that could be caused by population structure (Fig. M3). The heritability (*PVE*) null-model values were extracted from the linear mixed model used to infer SNP associations in GEMMA.

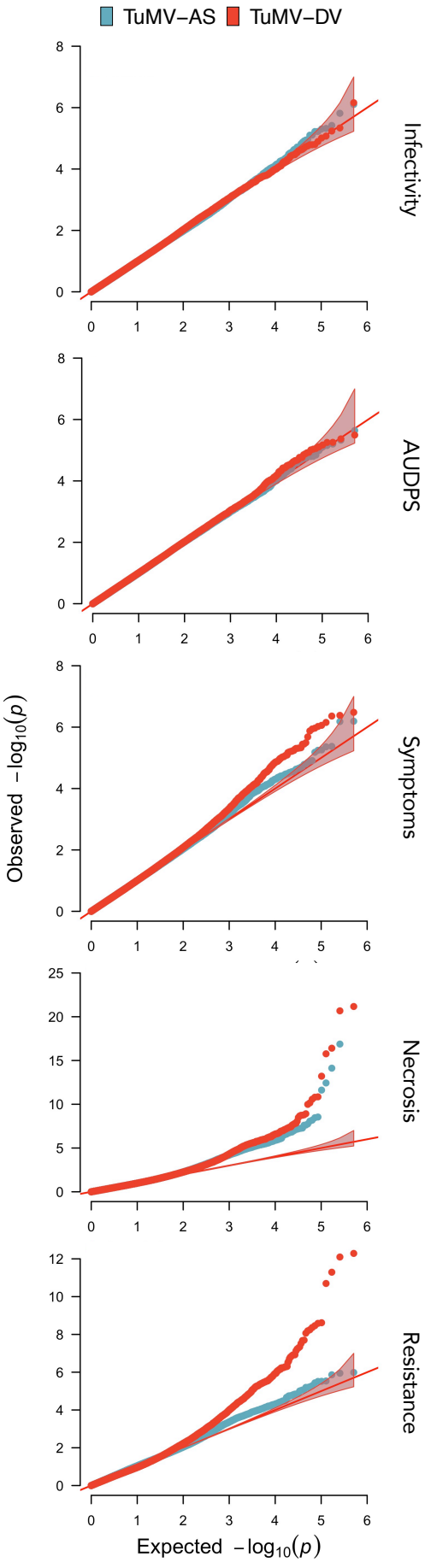

TuMV–AS
TuMV–DV
Infectivity
AUDPS
Symptoms
Necrosis
Resistance
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

Figure M3. The QQ-plots for all the five disease-related traits studied. Results for TuMV-AS infected plants are indicated in blue while results for TuMV-DV infected plants are shown in red.

GWAS analysis of the non-transformed phenotypic values of the 118 necrotic and non-necrotic accessions was done using the Accelerated Mixed Model (AMM) on the GWA-Portal website (https://gwas.gmi.oeaw.ac.at/). SNPs were considered significant if they were above the $-\log P \geq 8$ threshold.

1.6. Multiple-trait GWA analysis in Chapter 2

Multiple-trait association analyses of the disease-related traits for the GWA of 1050 accessions and the replicated study of 118 accessions were done with a Python program based on LIMIX (Lippert et al. 2014) written by Prof. Magnus Nordborg's group. In the multiple-trait GWA analysis for the 1050 accessions only the binary necrosis trait was used and the block effect was accounted for as a cofactor. In the replicate (118 accessions) analysis all four phenotypes were used in the multitrait model; (1) percentage of necrotic accessions per number of infected plants per accessions (2) percentage of necrotic accessions per total number of plants per accession and (3) *AUSIPS*. LIMIX is a linear mixed model (LMM) that treats SNPs and covariates as fixed effects while population structure and noise are treated as random effects. With the help of this package we assessed the extent to which SNPs for each disease-related trait are associated with the two viral isolates. They can be associated in the same way (direct effects) indicating that plants respond to the two viral isolates in the same way or associated in different ways (pleiotropic effects) suggesting different plant response for each viral isolate. Minor allele frequency cut-off of 10% was used and the traits were modelled as Gaussian in order to remove the *p*-value inflation.

1.7. Kruskall-Wallis test and cohort analysis in Chapter 2

The Kruskall-Wallis test was used to evaluate the influence of confounding and the distribution of the response variable on the LMM for the 1050 accessions. This tests if the large peak on chromosome 2 in position 5928864 that is associated with necrosis was an artefact of population structure or distributional assumptions, a Kruskall-Wallis test was performed on the 118 replicated accessions. Using this test the association between the genotype and necrosis was tested for each viral strain separately.

To further explore the significant peak on chromosome 2 for the 1050 accessions, the top SNP 5923326 was included in the multi-locus mixed model analysis (Segura et al., 2012) as a covariate where the appearance of additional significant peaks and presence of allelic heterogeneity was investigated.

1.8. Distribution of major and minor allele in position 5923326 on chromosome 2 in Chapter 2

For the set of 1050 accessions that show necrosis in any plant, the distribution of the major or minor allele in position 5923326 on chromosome 2 was plotted. The presence of major or minor allele in accessions showing the necrotic phenotype from the 1050 accessions was plotted onto a map of the world with a help of packages 'ggplot2', 'maps', and 'mapdata' in R version 3.6.1.

1.9. Bayesian sparse linear mixed model (BSLMM)

To determine whether many variants with small effects or a small number of large effects (sparse) variants were contributing to the disease-related traits variability, the BSLMM method implemented in GEMMA was used to infer the genetic architecture of the measured phenotypic traits (Zhou and Stephens 2012; Zhou, Carbonetto and Stephens 2013). BSLMM models the genetic contribution as the sum of a sparse component and a highly polygenic component. The proportion of genetic variance explained by sparse effects is represented by the parameter $PGE \in [0, 1]$. The second parameter in the model is the total variance explained ($PVE \in [0, 1]$) by additive genetic variants. *PVE* is a flexible Bayesian equivalent of the narrow sense heritability (h^2) estimated by more classical linear mixed models (LMM). BSLMM also outputs a parameter ngamma which is the number of variants with major effect.

In Chapter 1, raw values were used for symptoms severity (discrete variable) while *AUDPS* and infectivity (continuous variable) were normalized by block using a univariate general linear model in SPSS version 25 (IBM Corp., Armonk NY, USA). The block effect on the phenotypes was not incorporated as a covariate in GEMMA, as it was done in LIMIX, because of the optimization algorithm in GEMMA that causes errors if some covariates are identical to some genotypes. In all cases, MCMCs were run with the default settings (burn-in at 100,000, sampling steps at 1,000,000 and recording every ten steps) and minor allele frequency cut-off set at 5%.

In Chapter 2, all five phenotypes were analyzed with BSLMM using MCMC with the default configuration, as described above. *AUDPS* and infectivity were standardized by block as described above and analyzed using linear BSLMM. Severity of symptoms was also analyzed using linear BSLMM. The binary traits necrosis and resistance were analyzed using the probit BSLMM.

The posterior inclusion probability (*PIP*) for a SNP is the probability of including this SNP as causal in the MCMC analysis, estimated from posterior samples of a Gamma distribution that reflects the sparse effects (Schaid, Chen and Larson 2018). This can be used as a measure of the strength of the association that an SNP has with the corresponding phenotype. Variants with a large effect in at least 25% of the MCMC samples were diagnosed as significant ($PIP \geq 0.25$).

1.10. Validation of GWAS associations

In Chapter 1, ten genes identified with the GWAS were selected for further study of their LOF effect on disease progression (Table 2). The 10 chosen Col-0 T-DNA insertion LOF mutants were selected on the criteria that (1) a candidate gene per each of the phenotypic traits per virus was included and (2) they were available as homozygous lines in NASC stock center (https://arabidopsis.info/BrowsePage). The *AUDPS* and the *AUSIPS* were calculated using the number of infected plants and their symptomatology was measured during 21 dpi for each individual plant. For statistical comparisons, a bootstrap approach was taken. One thousand pseudo-replicated matrices, of equal dimensions to the original one (rows representing individual plants and columns representing dpi), were generated per experimental condition. The matrix rows were replaced and thus the temporal correlations across time points were preserved. This algorithm, implemented in R version 3.6.1, generated kernel distributions for *AUDPS* and *AUSIPS*. The 89% highest density intervals (HDI) were calculated using the bayestestR package in R version 3.6.1 in RStudio version 1.2.1335 (Makowski, Ben-Shachar and Lüdecke 2019). To maximize the statistical power of the tests, a difference between two samples was deemed significant if the 89% HDIs did not overlap.

In Chapter 2, nine candidate genes were selected for validation, based on their previously described function and the observed *P*-value among all the significant hits. Five of the

selected genes were mapped for infection with the two viral strains, meanwhile two were associated with TuMV-AS and two with TuMV-DV. The selected genes along with their description and the significant SNP position can be found in Table 3. LOF mutants of the selected genes were ordered from the NASC stock center (http://arabidopsis.info/BrowsePage). Mutants from NASC were chosen on the following criteria: (1) must be in the Col-0 background, (2) must be T-DNA inserts that cause gene LOF and (3) must be homozygous. The number of infected plants and their symptomatology was measured during 21 dpi for each individual plant and *AUDPS* and *AUSIPS* were also calculated. One thousand pseudo-replicated matrices were also calculated for *AUDPS* and *AUSIPS* and their 89% HDIs were calculated using bayestestR package as described above.

2. Evaluation of mutational and environmental robustness

2.1. Viruses, plants, and inoculations

As a source of the inocula for all experiments described below, we used stocks of infectious saps from Arabidopsis Col-0 infected plants. Saps were obtained by grinding the corresponding infected tissues in a mortar with ten volumes of grinding buffer (50 mM KH_2PO_4 pH 7, 3% polyethylene glycol 6000). In the case of TuMV-AS, an arabidopsis-naïve virus, *N. benthamiana* plants were inoculated with the plasmid p35STunos that contains a cDNA of TuMV isolate YC5 from calla lily (*Zantedeschia sp.*; GenBank accession AF530055.2) under the control of the cauliflower mosaic virus 35S promoter and the NOS terminator (Chen et al., 2003). A large stock of viral particles was produced from these plants. In the case of TuMV-DV, the virus was obtained after twelve serial passages of experimental evolution in arabidopsis accession Col-0 of the ancestral TuMV-AS isolate (González et al. ,2019; Navarro et al., 2020), thus representing the case of an arabidopsis-adapted virus.

Arabidopsis plants were always inoculated when they reached growth stage 3.5 in the Boyes' scale (Boyes 2001). Aliquots of 5 µl of 10% Carborundum in grinding buffer were applied onto three different leaves, and inoculation was done mechanically by gentle rubbing with a glass stick.

Unless otherwise indicated, plants were maintained in a BSL-2 growing chamber at 16 h light:8 h dark cycles and temperature variation of 24 °C day:20 °C night. Plants that showed visible symptoms of infection were harvested at 14 dpi.

2.2. Evaluation of mutational robustness

N_2O mutagenesis was done as described in Willemsen et al. (2018). In short, ground-infected tissues were homogenized with DEPC-treated sterile water at 1:1 (w:v) ratio. Diluted saps were centrifuged 2 min at 12,000 rpm at 4 °C and the supernatant was transferred into two different tubes. The first tube contained a control reaction consisting of equal volumes of water and 0.5 M sodium acetate (pH 5.4). The second tube contained the mutagenic reaction consisting in equal volumes of 2 M $NaNO_3$ and 0.5 M sodium acetate (pH 5.4). These tubes were incubated at 26 °C for 3 h. After incubation, $1/10^{th}$ volume of 0.5 M phosphate buffer (pH 7) was added to the tubes to stop the mutagenic reactions.

Four groups of twelve plants were inoculated each with mutagenized and non-mutagenized versions of TuMV-AS and TuMV-DV. Inoculated plants were maintained in the standard growth conditions described in Section 2.1 during 21 dpi.

2.3. Evaluation of thermal robustness

All plants were maintained in the standard cultivation conditions described in Section 2.1 from germination until one week before inoculation. During this week, plants were acclimatized to the thermal conditions corresponding to each of the following four experimental condition (24-four plants each): (1) constant 24 °C; (2) constant 30 °C; (3) sequential changes between 15°C, 24°C, and 30°C every 24 h (median temperature across the entire experiment 24.0 °C, IQR 13.5 °C); and (4) random changes between 15 °C, 24 °C, and 30 °C every 24 h (median temperature across the entire experiment 24.0 °C, IQR 15.0 °C). In all four setups, illumination conditions remained 16 h light and 8 h dark. After this acclimation week, plants were inoculated; twelve with TuMV-AS and twelve with TuMV-DV, and kept in the corresponding thermal regime during 21 dpi. Treatments (3) and (4) were designed to increase the amount of environmental noise to which the replicating TuMV population would be exposed. The possibility of adding an additional constant 15 °C treatment was discarded after some preliminary experiments because infections progressed asymptomatic and with very low viral loads (data not shown).

2.4. Disease progression curves as a proxy to the degree of viral adaptation

All inoculated plants were observed daily for 21 dpi for the presence of symptoms and the number of symptomatic plants recorded. Disease progression curves were characterized by three parameters, the median time to the development of visible symptoms (ST_{50}), the final frequency of infected plants, or infectivity, and *AUDPS*. *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 21 dpi) and twelve (all plants show symptoms at 1 dpi).

In the TuMV/arabidopsis pathosystem, there is a one-to-one match between infection status and the development of symptoms (González et al., 2019; Corrêa et al., 2020); all infected plants develop obvious symptoms at the temperature conditions used in this experiment. Likewise, in this pathosystem the intensity of symptoms is significantly correlated with viral load (Corrêa et al. 2020). Symptoms started with leaf curling and vein clearing (∼5 - 6 dpi) that quickly developed to diverse grades of leaf chlorosis and/or necrosis (∼10 - 12 dpi). Plants also suffered a developmental arrest, with deformed new leaves, siliques abortion, and abnormal growth of the caulinar apex.

2.5. Statistical analyses

The disease progression curves were analyzed using Kaplan–Meier survival regression analyses as implemented in SPSS version 26 software (IBM, Armonk, NY). The significance of factor effects was evaluated using the log-rank Mantel-Cox test statistic that asymptotically follows a χ^2 distribution.

Infection data for each treatment were organized in a 12×22 binary matrix, where rows represent individual plants and columns dpi. Infection status was coded as 1 if plants showed symptoms and 0 otherwise. *AUDPS* values were computed using the 'agricolae' R package version 1.3-2 (https://tarwi.lamolina.edu.pe/∼fmendiburu/). Confidence intervals (95% CIs) were estimated using a bootstrapping method consisting in sampling with replacement the matrix rows, thus preserving the temporal correlations across time points. A thousand pseudo-replicated matrices of equal dimensions to the original one were obtained per experimental condition, thus generating kernel distributions for *AUDPS*. The median *AUDPSs* and their corresponding 95 per cent CIs were estimated

from these distributions. This algorithm was implemented in R version 3.6.1 in RStudio version 1.2.1335.

A measure of environmental robustness is the inverse of the environmental variance, σ_E^2, which results from external environmental perturbations (de Visser et al. 2003). Variance components in a one-way ANOVA model testing for differences among thermal environments were estimated by maximum likelihood techniques as implemented in SPSS version 26 software (IBM, Armonk, NY). Net differences among thermal environments correspond to σ_E^2, whereas differences among replicates within a given environment correspond to random noise.

Results

Chapter 1: Arabidopsis genes contributing to differences in the outcome of infection with generalist and specialist strains of TuMV identified by genome-wide association studies

1. Introduction

Viruses are constantly facing heterogeneity in the hosts they infect. They face species with different response to infection, or in many instances among individuals within the same host species. Some viruses adapt to a particular host species or genotype in which they efficiently complete their reproductive cycle. These viruses are called specialists. Specialist viruses pose a great threat *e.g.* to monocultured crops since well-adapted viruses usually show enhanced within-host replication rates that are often associated with stronger symptoms (Roossinck, 2010; Lacroix et al., 2014; Stobbe and Roossinck, 2016). Examples of specialist viruses are *Dengue flavivirus* and *Mumps orthorubulavirus*, among mammalian viruses, and *Barley stripe hordeovirus* from plants (Elena et al., 2009; Roossinck, 2010). Other viruses infect hosts from widely different genotypes, species, or even higher taxonomical units, and are dubbed generalists (Elena et al., 2009). *Cucumber mosaic cucumovirus* (that infects more than 1000 plant species) and the *Alphainfluenzavirus* (that infects birds, humans and other mammalian species) are examples of generalist viruses (Elena et al., 2009).

Each host range strategy comes with advantages and disadvantages. By specializing in a single host, a virus can limit interspecific competition and better access limited resources (Elena et al., 2009; Bedhomme et al., 2014). The advantage of generalism is the successful infection of multiple hosts. However, there is an obvious limitation to generalism: by being able to infect multiple hosts a virus does not maximize fitness in any particular one (Bedhomme et al., 2015), conforming to the *jack-of-all-trades is a master of none* hypothesis (Whitlock, 1996). It is proposed that selection favors specialist viruses because there is a trade-off limiting the fitness of a generalist virus in any of the alternative hosts and evolution proceeds faster in narrower niches (Woolhouse, 2001). Antagonistic pleiotropy, where beneficial adaptations to a particular host could be

disadvantageous in another (Lalić *et al*., 2011), is the most commonly claimed mechanism to explain this trade-off. Furthermore, to infect multiple hosts, viruses might need to encode for additional genetic information that would slow down their replication and increase their mutational fragility. Also, mutations that are fixed in order to compensate for antagonistic pleiotropy limit access to alternative evolutionary paths towards global maxima in the fitness landscape, reducing evolvability (Cervera et al., 2016). All these characteristics make specialists capable of faster evolution and adaptation than generalists in the face of perturbations or new environments (Bedhomme et al., 2015; Bono et al., 2020). Although specialists tend to adapt faster to single hosts, generalists usually outcompete them in fluctuating environments by being more prepared to survive and reproduce as a consequence of having similar fitness in different hosts (Kassen, 2002; Dennehy et al., 2013). This allows generalist viruses to have higher initial fitness compared to specialists when infecting novel host species and makes them most likely emerging and re-emerging pathogens (Woolhouse and Gowtage-Sequeria, 2005; Turner et al., 2010). Indeed, this theory has widespread support by experiments in which viral lineages being sequentially exposed to different hosts for long periods of time maximize their fitness in all hosts in the same extent as the corresponding specialist, thus overcoming the expected costs of generalism (Turner and Elena, 2000; Deardorff et al., 2011; Bedhomme et al., 2012; Remold, 2012).

The genetic basis of the observed differences between generalist and specialist viruses is actually poorly understood, at least from the perspective of the interaction of these two strategies with the host gene expression. Differences between the genomes of generalist and specialist viruses have been previously described (Takeuchi et al. ,1991; Llamas-Saiz et al. ,1996; Remold et al., 2008; Deardroff et al., 2011; Hillung et al., 2014; Navarro et al., 2020). However, so far just one study has sought to explore differential host responses associated with each virus strategy (Hillung et al., 2016). Here, we aim to explore whether viruses with different host range strategies affect the plant physiology and disease progression in different ways, identifying candidate host genes that differentially respond to a specialist or a generalist virus.

To reach this goal, we have undertaken a GWAS approach. GWAS has gained popularity over the last 20 years due to the increasing number of genome sequences available for a wide range of organisms (Cantor et al., 2010; Bush and Moore, 2012). The basis of

GWAS is capturing single-nucleotide polymorphisms (SNPs) along the genome of an organism and, using statistical methods (such as linear mixed models), to infer the association of SNPs to the trait being analyzed. The common disease-common variant hypothesis posits that common interacting alleles at multiple disease-predisposing loci underlie most common diseases (Bush and More, 2012). This hypothesis would justify the use of GWAS in the identification of alleles associated with specific phenotypes. This connection permits the identification of genetic risk factors for disease, such as susceptibility and resistance to viral infections (Korte and Farlow, 2013). One of the most relevant inferences from GWAS is trait heritability, which indicates how much of the observed phenotypic variation is explained by genotypic variation (SNPs) relative to the contribution of environmental factors (Zaitlen and Kraft, 2012).

Identifying host factors responsible for resistance or permissiveness to infection is the ultimate goal when studying host-pathogen interactions, as this knowledge will help in better management of diseases. Here we have characterized the infection of generalist and specialist strains of TuMV in 450 natural accessions of arabidopsis. The viral strains used in this study were obtained by Navarro et al. (2020) (see section Methods for details on the evolutionary history of these two strains).

In summary, the response to infection of 450 *A. thaliana* natural accessions from different geographic regions was phenotyped in a controlled common garden setting. These accessions were inoculated with two TuMV strains that differ in their degree of specialization. Infection data was analyzed using GWAS, specifically looking for SNPs differentially associated with the infection with generalist and specialist TuMV strains. The genetic architecture of the phenotyped disease-related traits was also studied using the Bayesian sparse linear mixed model (BSLMM).

2. Results

2.1. Characterization of infection traits in natural accessions

The 450 *A. thaliana* accessions (Supplementary Table S1) infected with the generalist (TuMV-G) and specialist (TuMV-S) TuMV strains were phenotyped for disease-related

traits. Three disease-related traits were characterized by visual inspection and are shown in Fig. C1.1.

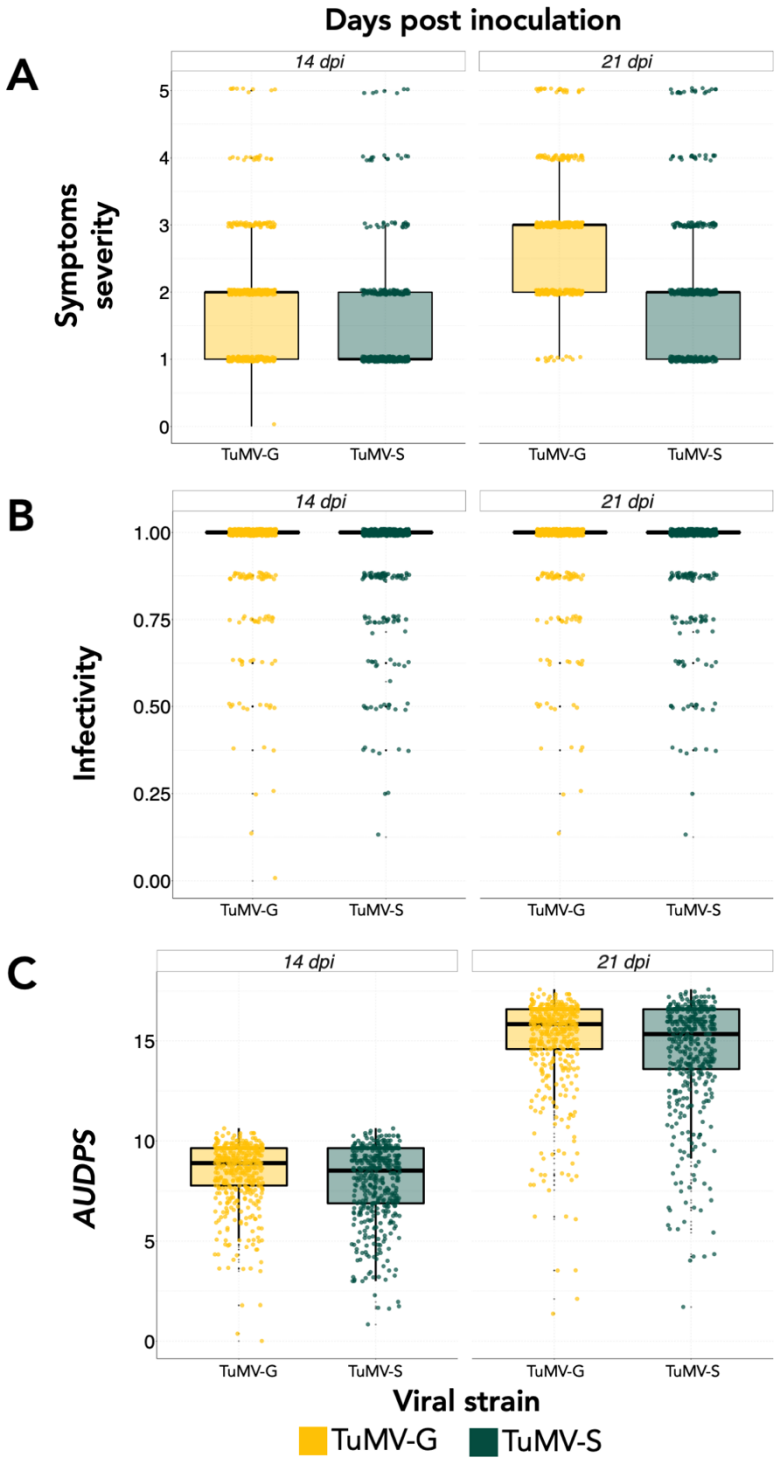


Figure C1.1. Distribution of the three disease-related traits characterized for each viral strain (TuMV-G in yellow and TuMV-S in green) infecting the 450 arabidopsis natural

accessions at 14 (left) and 21 dpi (right). (A) Severity of symptoms. (B) Infectivity. And (C) *AUDPS*.

Table C1.1. shows the results of the Scheirer-Ray-Hare test used to evaluate the effect of virus genotype, dpi and their interaction on each disease phenotype. Firstly, highly significant differences exist between the two viruses, with TuMV-G showing larger median values than TuMV-S for the three traits (median ±IQR for TuMV-G *vs* TuMV-S, respectively: symptoms severity 2.354 ±3.000 *vs* 1.779 ±4.000; *AUDPS* 11.692 ±8.822 *vs* 11.162 ±9.447; and infectivity 0.950 ±0.000 *vs* 0.936 ±0.000) (Fig. 2). Secondly, a highly significant effect has been observed associated to dpi for the severity of symptoms and *AUDPS* (median ±IQR at 14 *vs* 21 dpi, respectively: symptoms severity 1.721 ±4.000 *vs* 2.412 ±3.000; *AUDPS* 8.148 ±3.388 *vs* 14.708 ±3.625) but not for infectivity, indicating that the number of plants diagnosed as infected based on the presence of symptoms did not increase during the last seven days, while symptoms got worse (Fig. 2). Thirdly, a significant interaction between both factors has only been observed in the case of severity of symptoms (Table 1), which in this case suggests that the difference between the two viral strains for this trait was larger at 21 dpi (relative change in means ~40%) than at 14 dpi (relative change in means ~25%) (Fig. 2).

Table C1.1. Non-parametric 2-ways ANOVA (Scheirer-Ray-Hare) test of the two main effects and their interaction for each of the three disease-related traits experimentally determined.

Source of variation	df	Symptoms severity		*AUDPS*		Infectivity	
		H[1]	***P***	***H***	***P***	***H***	***P***
Virus genotype	1	191.414	< 0.001	7.711	0.006	7.183	0.007
dpi	1	250.766	< 0.001	1122.420	< 0.001	0.257	0.612
Virus genotype by dpi	1	14.984	0.001	0.021	0.883	0.037	0.847

[1]*H* statistic follows a χ^2 distribution.

Furthermore, it is well known that in the case of TuMV the set of arabidopsis genes differentially expressed changes along the stage of infection (Sánchez et al., 2015; Corrêa et al., 2020). Guided by these previous experiments, the infection traits were studied both at 14 and 21 dpi to account for potential differences between the viral strains at different stages. Accordingly, GWAS of the infection traits was performed at both time points.

2.2. Genetic architecture of disease-related traits

The 450 accessions accounted for 431,323 SNPs that were tested in both viral strains at both 14 and 21 dpi (Table C1.2.) using the BSLMM analysis. This analysis evaluates how much of the observed phenotypic variance (*PVE*) is explained by the genotyped SNPs and how important are the contributions of sparse effects to the genetic variance (*PGE*).

Table C1.2. Results of the BSLMM analysis on the four disease-progression selected traits and their 89% HDI.

	14 dpi	TuMV-S									
		PVE				*PGE*				ngamma	
		mean	median	lower 80% HDI	higher 80% HDI	mean	median	lower 89% HDI	higher 89% HDI	mean	median
Phenotype	*AUDPS*	0.28	0.28	0.19	0.38	0.48	0.48	0	0.91	111.5	61
	Infectivity	0.09	0.09	0.03	0.15	0.66	0.77	0.12	1	11.3	3
	Symptoms	0.19	0.16	0	0.37	0.38	0.35	0	0.79	9.7	5

	21 dpi	TuMV-S									
		PVE				*PGE*				ngamma	
		mean	median	lower 80% HDI	higher 80% HDI	mean	median	lower 89% HDI	higher 89% HDI	mean	median
Phenotype	*AUDPS*	0.22	0.22	0.12	0.31	0.63	0.71	0.13	1	7.6	5
	Infectivity	0.08	0.08	0.03	0.14	0.65	0.77	0.11	1	19.7	3
	Symptoms	0.15	0.13	0	0.27	0.38	0.33	0	0.81	22.6	8

	14 dpi	TuMV-G									
		PVE				*PGE*				ngamma	
		mean	median	lower 80% HDI	higher 80% HDI	mean	median	lower 89% HDI	higher 89% HDI	mean	median
Phenotype	*AUDPS*	0.09	0.08	0.03	0.14	0.5	0.58	0	0.94	11	5
	Infectivity	0.08	0.08	0.03	0.13	0.7	0.8	0.22	1	5.6	2
	Symptoms	0.11	0.1	0	0.21	0.38	0.33	0	0.82	20.7	11

	21 dpi	TuMV-G									
		PVE				*PGE*				ngamma	
		mean	median	lower 80% HDI	higher 80% HDI	mean	median	lower 89% HDI	higher 89% HDI	mean	median
Phenotype	*AUDPS*	0.1	0.1	0.04	0.15	0.76	0.86	0.38	1	2.6	2
	Infectivity	0.12	0.11	0.05	0.22	0.86	0.92	0.71	1	3.3	2
	Symptoms	0.09	0.07	0	0.17	0.38	0.33	0	0.81	32.6	15

AUDPS, infectivity and symptoms severity had low *PVE* values (Table C1.2.). The lowest *PVE* value was obtained for *AUDPS* [median 0.08 and 89% HDI (0.03, 0.14)] and infectivity [median 0.08 and 89% HDI (0.03, 0.13)] for TuMV-G at 14 dpi and for TuMV-S at 21 dpi [median 0.08 and 89% HDI (0.03, 0.14)], while the largest value was obtained also for *AUDPS* measured at 14 dpi but for TuMV-S [median 0.28 and 89% HDI (0.19, 0.38)]. In all other instances, *PVE* values were similar for both TuMV strains and between the two time points.

Regarding *PGE* (Table C1.2.), on the one hand the smallest value was observed for the severity of symptoms induced by TuMV-G at 14 dpi [median 0.33 and 89% HDI (0, 0.82)] and 21 dpi [median 0.33 and 89% HDI (0, 0.81)] and by TuMV-S at 21 dpi [median 0.33 and 89% HDI (0, 0.81)]. On the other hand, the largest *PGE* value was estimated for TuMV-G infectivity measured at 21 dpi [median 0.92 and 89% HDI (0.71, 1)]. The percentage of *PVE* explained by large sparse effect variants (*PGE*) indicates that major effect loci account for between 50 - 90% of phenotypic variance in *AUDPS* and infectivity traits, in both time points for both viruses (median values reported in Table C1.2.). The number of variants with large effect size, the SNPs that explain most of the phenotype among the 431,323 SNPs, was low for infectivity and severity of symptoms at 14 dpi for both viruses as well as for *AUDPS* and infectivity at 21 dpi also for both viruses (Table C1.2.). To detect large-effect SNPs that might be contributing the most to the variance in disease-related phenotypes a $PIP \geq 0.25$ threshold was imposed in the BSLMM model in GEMMA. With this constrain, three highly significant SNPs have been detected. The first was detected for TuMV-S *AUDP*S estimated at 21 dpi. This SNP was mapped within the gene encoding for *AT2G04440*, a MutT/Nudix family protein. The second significant SNP was found for TuMV-G infectivity at 21 dpi within locus *AT3G19350*, that corresponds to the gene *MATERNALLY EXPRESSED PAB C-TERMINAL (MPC)*. The third significant SNP was also observed for TuMV-G infectivity at 21 dpi and corresponds to position 6,685,977 of an intergenic region on chromosome 3.

Chromosome 3 intergenic position 6,685,977 is between loci *AT3G19290*, which corresponds to the gene *ABA-RESPONSIVE ELEMENT BINDING PROTEIN 4 (ABF4)*, and *AT3G19280*, which corresponds to the gene *FUCOSYLTRANSFERASE 11 (FUT11)*. Interestingly, the chromosome 3 intergenic position 6,685,977 shows a strong LD (r^2 = 1; in a 10 kb window) with *FUT11*.

Next, we ran an LD analysis to discover SNPs at different loci that might be significantly associated. A total of four pairs of SNPs located at different loci showed significant LD values ($r^2 > 0.5$ in all cases; Table C1.3.).

Table C1.3. LD analysis of the significant SNPs from the GWAS. Loci pairs that were in $r^2 > 0.5$ are shown.

LOCUS 1	LOCUS 2	VIRUS AND PHENOTYPE
AT3G07470 TRANSMEMBRANE PROTEIN	*AT3G07470* transmembrane protein	TuMV-S, AUDPS 21 dpi
AT3G21660 UBX DOMAIN-CONTAINING PROTEIN	*AT3G21670* Major facilitator superfamily protein	TuMV-G, Symptoms 21 dpi
AT4G02580 NADH DEHYDROGENASE [UBIQUINONE] FLAVOPROTEIN 2	*AT4G02590* Basic helix loop helix class transcriptional regulator	TuMV-S, Infectivity 14 dpi
AT4G10130 DNAJ HEAT SHOCK N-TERMINAL DOMAIN-CONTAINING PROTEIN	*AT4G10130* Encodes a protein with putative sucrose-phosphate synthase activity	TuMV-S, Infectivity 21 dpi

The rest of SNPs showed strong LD only with other SNPs within the same locus. All four significant pairs involved protein coding genes. Three of the four pairs of SNPs in LD were mapped for TuMV-S.

In summary, for this host-pathogen system, the genetic architecture of *AUDPS* and infectivity phenotypes is relatively simple, involving few small-effect SNPs along with one large effect SNP that is being responsible for the majority of variance in the observed phenotypes. Symptoms severity, however, is genetically more complex and involves many more small effect SNPs. For both viral strains, all the disease phenotypes have a similar genetic architecture between the two temporal stages (14 and 21 dpi).

2.3. GWAS identifies genetic loci associated with disease-related phenotypes differentially induced by specialist and generalist viral strains

The significantly associated SNPs for the three disease-related traits were visualized using Manhattan plots in (Fig. C1.2.).

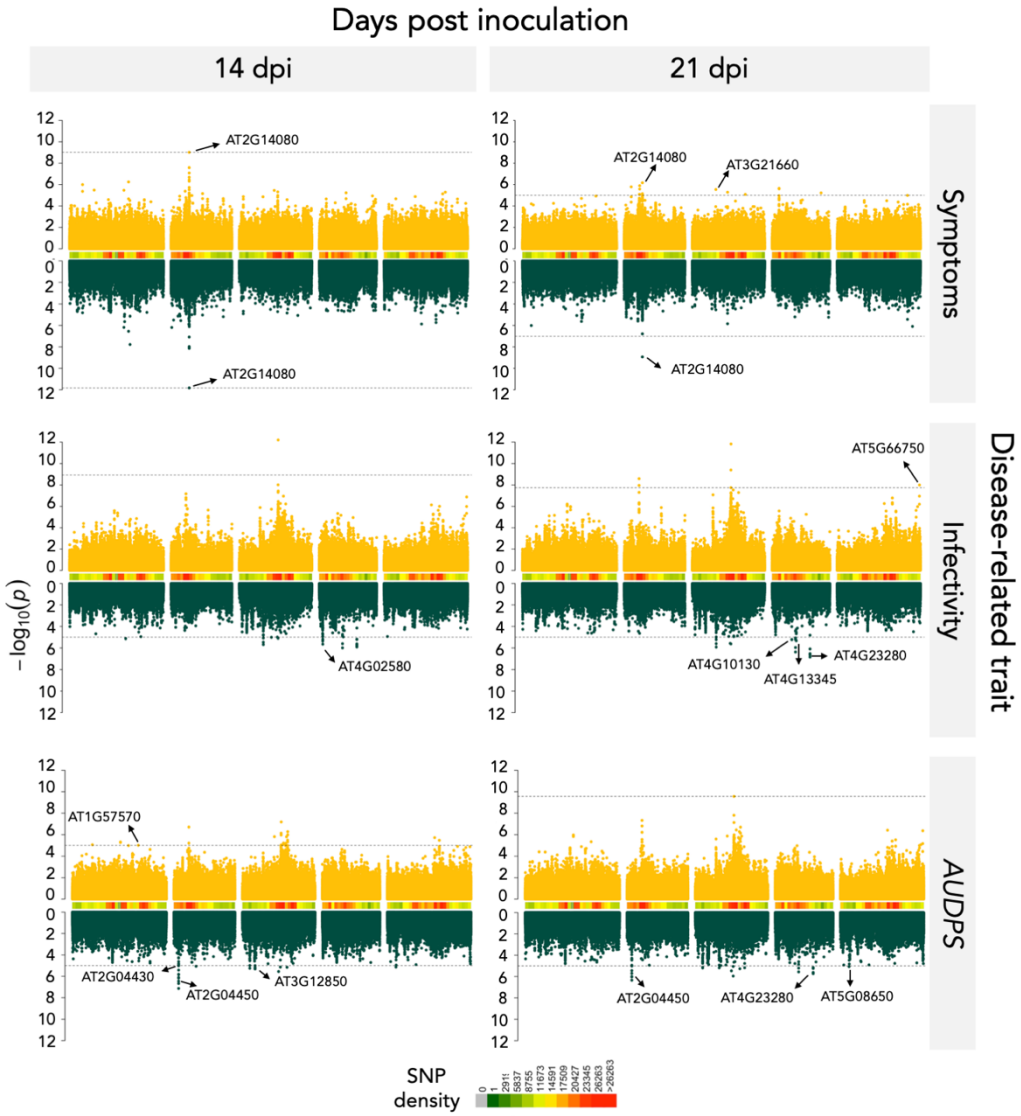


Figure C1.2. Manhattan plots of the analyzed disease-related traits. Data for TuMV-G are indicated in yellow and for TuMV-S in green. Peaks marked on the plots correspond to the most significant SNP values of the genes selected for the mutant analysis. SNP density shows how many SNPs are genotyped for a particular chromosomal region. The dashed lines indicate the significance threshold (FDR or $-\log P = 5$ whenever computing FDR was not possible).

The QQ-plots for infection traits showed no detectable population structure (Figure M2). Using the FDR or the $-\log P \geq 5$ thresholds determined for each of the traits (Methods, section 1.5), a total of eight significant SNPs were identified for TuMV-G and 19 for TuMV-S infection (Table C1.4.).

Table C1.4. Significant genes detected using GWAS for the three disease-related traits during the course of infection with each viral strain.

TUMV-G

TRAIT	dpi	name	gene	description	chr	-log
AUDPS	14	*JAL14*	*AT1G57570*	Jacalin-related lectin 14	1	5.01
SYMPTOMS	14 / 21		*AT2G14080*	Disease resistance protein (TIR-NBS-LRR class) family	2	9.02
SYMPTOMS	21	*PUX6*	*AT3G21660*	Plant UBX domain-containing protein 6	3	5.54
SYMPTOMS	21	*TRP2*	*AT3G46590*	Telomere repeat-binding protein 2	3	5.09
SYMPTOMS	21	*CRK39*	*AT4G04540*	Putative cysteine-rich receptor-like protein kinase 39	4	5.65
SYMPTOMS	21	*AFC3*	*AT4G32660*	Serine/threonine-protein kinase AFC3	4	5.22
INFECTIVITY	21	*DDM1*	*AT5G66750*	ATP-dependent DNA helicase DDM1	5	8.00

TUMV-S

TRAIT	dpi	name	gene	description	chr	-log
SYMPTOMS	14 / 21		*AT2G14080*	Disease resistance protein (TIR-NBS-LRR class) family	2	11.83
INFECTIVITY / *AUDPS*	(14, 21) /21	*CRK20*	*AT4G23280*	Putative cysteine-rich receptor-like protein kinase 20	4	6.52
AUDPS	14 / 21	*NUDT6*	*AT2G04450*	Nudix hydrolase 6	2	6.45
AUDPS	14 / 21		*AT2G04440*	MutT/nudix family protein	2	6.11
INFECTIVITY	14 / 21	*CRRSP27*	*AT3G21980*	Putative cysteine-rich repeat secretory protein 27	3	5.91
INFECTIVITY	14		*AT4G02580*	NADH dehydrogenase [ubiquinone] flavoprotein 2	4	5.64
INFECTIVITY	21	*MEE55*	*AT4G13345*	Serinc-domain containing serine and sphingolipid biosynthesis protein	4	5.32
INFECTIVITY	21	*T9A4.1*	*AT4G10130*	DNAJ heat shock N-terminal domain-containing protein	4	5.22
AUDPS	21		*AT3G07470*	transmembrane protein, putative	3	5.20
AUDPS	14		*AT3G12850*	COP9 signalosome complex-related / CSN complex-like protein	3	5.18
AUDPS	14 / 21		*AT5G08650*	Translation factor GUF1 homolog, chloroplastic	5	5.14
AUDPS	14	*NUDT5*	*AT2G04430*	Nudix hydrolase 5	2	5.02

Some of these SNPs were positioned within seven genes for TuMV-G and 12 for TuMV-S (Table C1.4.). Most of the identified genes were unique for TuMV-G or TuMV-S

strains, with only one shared locus both at 14 and 21 dpi, for symptoms severity: the aforementioned *AT2G14080* (Fig. C1.3A).

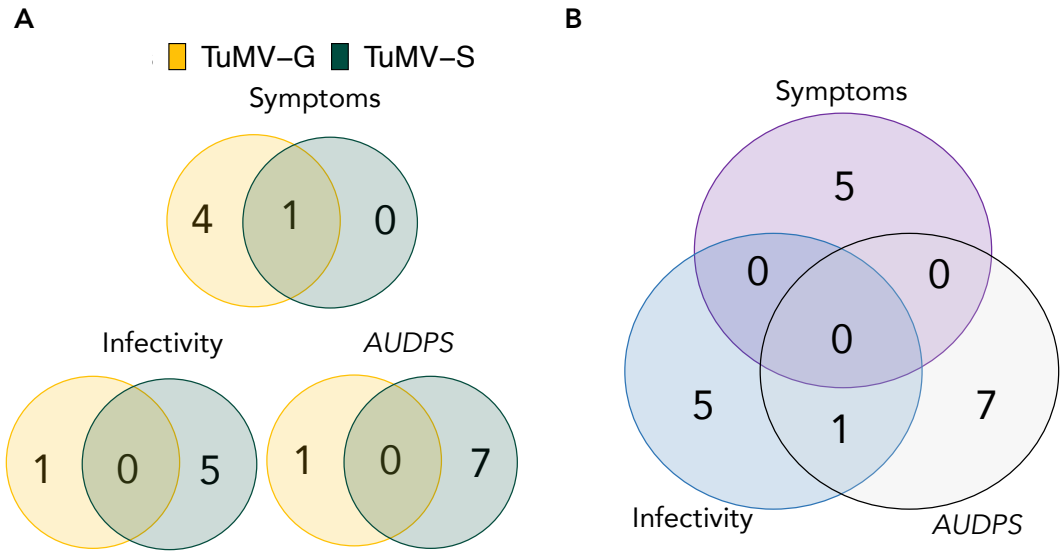


Figure C1.3. Venn diagram showing the number of unique and shared genes. (A) Genes mapped for each viral strain and disease-related traits. (B) Genes mapped for all disease-related traits pooling together both viral isolates.

Comparing the results at 14 and 21 dpi for TuMV-G, genes at 14 dpi seem more related to a general disease response while genes at 21 dpi are more specific and involved in ubiquitin-related processes. Such temporal difference is not seen between for TuMV-S. This may suggest that plants responses to the generalist viral strain change more dynamically than when infected with the specialist strain, in which case the response seems unchanged between the two time points studied. Fig. C1.3B shows that most of the identified genes (17) had an effect only in one of the disease-related traits. However, locus *AT4G23280* that encodes for the putative *CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 20 (CRK20)* was involved both in *AUDPS* and infectivity. Interestingly, locus *AT4G04540* mapped for symptoms severity also encodes for a putative *RICH RECEPTOR-LIKE PROTEIN KINASE 39 (CRK39)*.

2.4. Experimental validation of identified genes

Ten of the identified genes were selected for a validation study in which the corresponding LOF mutants were inoculated with both viral strains and the disease progression was characterized (Fig. C1.4 and Table M4).

Out of the 10 genes, one was shared between the two viral strains, two were unique for TuMV-G and seven were unique for TuMV-S. More genes were validated for TuMV-S because the GWAS mapped more significant SNPs upon infection with this strain. The selected LOF mutants were: *at1g57570*, *nudx5*, *nudx6*, *at2g14080*, *at3g12850*, *at4g10130*, *mee55*, *CPLEPA*, *ddm1*, and *at4g02580*. To evaluate differences in infection dynamics between the mutants and the WT plants, *AUDPS* and *AUSIPS* were calculated using the data collected along the 21 dpi. A comparation between the WT and LOF mutant values for each viral strain was done (Table C1.5 and Fig. C1.4) based on the inferred 89% HDIs obtained with the bootstrap method. Differences in most of the LOFs were found when comparing the *AUDPS* values of the two viral strains with the WT (Table C1.5 and Fig. C1.4).

Table C1.5. 89% HDIs calculated for *AUDPS* and *AUSIPS* for TuMV-G and TuMV-S on each KO mutant and WT plant. A +/– next to a row marks the 89% HDI values that are higher (+) or lower (-) in the mutants compared to the WT.

AUDPS

		TuMV-G				TuMV-S		
	WT	15.70	16.70		**WT**	14.10	15.90	
AT1G57570	***at1g57570***	14.70	16.40		***at1g57570***	14.80	16.00	
AT2G14080	***at2g14080***	13.40	15.60		***at2g14080***	14.00	15.10	
AT3G12850	***at3g12850***	17.20	16.80		***at3g12850***	14.00	16.50	
AT4G02580	***at4g02580***	8.90	15.40	–	***at4g02580***	3.10	11.00	–
AT4G10130	***at4g10130***	12.90	16.70		***at4g10130***	16.10	16.90	+
AT5G08650	***cplepa***	13.80	15.60	–	***cplepa***	4.20	11.80	–
AT5G66750	***ddm1***	16.80	17.00	+	***ddm1***	11.70	16.70	
AT4G13345	***mee55***	12.10	16.50		***mee55***	9.90	14.50	
AT2G04430	***nudt5***	12.80	14.90	–	***nudt5***	7.60	12.70	–
AT2G04450	***nudt6***	13.30	15.00	–	***nudt6***	9.30	13.60	–

AUSIPS

		TuMV-G				TuMV-S		
	WT	32.90	36.50		**WT**	21.10	26.90	

AT1G57570	***at1g57570***	36.20	40.50		***at1g57570***	23.50	28.50	
AT2G14080	***at2g14080***	42.40	47.80	+	***at2g14080***	28.30	35.00	+
AT3G12850	***at3g12850***	35.90	42.30		***at3g12850***	28.10	36.60	+
AT4G02580	***at4g02580***	18.80	32.30	–	***at4g02580***	4.60	16.70	–
AT4G10130	***at4g10130***	28.80	41.00		***at4g10130***	23.20	32.80	
AT5G08650	***cplepa***	31.80	35.90		***cplepa***	7.70	21.70	
AT5G66750	***ddm1***	35.20	42.10		***ddm1***	18.30	27.90	
AT4G13345	***mee55***	32.70	44.30		***mee55***	16.50	24.90	
AT2G04430	***nudt5***	30.50	36.90		***nudt5***	12.20	18.90	–
AT2G04450	***nudt6***	39.90	41.90	+	***nudt6***	20.10	29.60	

Evaluating the mutant *AUDPS* intervals, lower 89% HDIs compared to the WT imply that these mutants have slower disease progression because the LOF gene is positively involved in the viral cycle and the virus uses it to aid its replication or translation. For the TuMV-S and TuMV-G infection, there are four mutants that have lower 89% HDI compared to the WT: *at4g02580*, *cplepa*, *nudt5*, and *nudt6*.

Observing LOF mutants with *AUDPS* intervals higher than the WT suggests that the corresponding genes are involved in plant defense response against infection; removing them enhanced disease progression beyond the one observed for the WT plants. In TuMV-G infection, *ddm1* had 89% HDI higher than the WT. In plants infected with TuMV-S only *at4g10130* had higher 89% HDI compared to the WT.

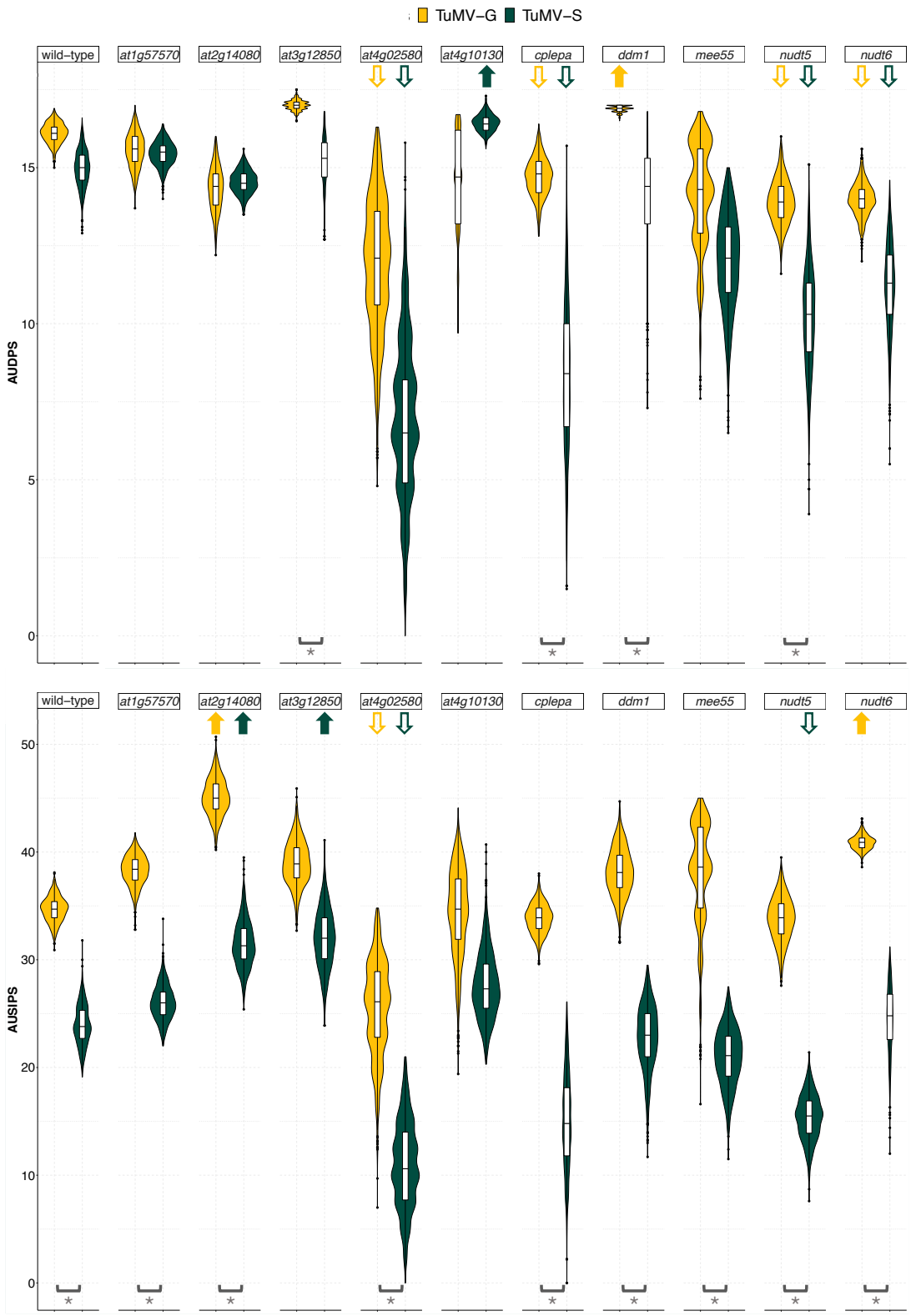


Figure C1.4. 89% HDI calculated for *AUDPS* and *AUSIPS* for each viral strain on each LOF mutant plant genotype. Not overlapping 89% HDI between a given mutant and the WT plants is indicated by an arrow. Arrows pointing up indicate a significant positive difference in medians, while arrows pointing down indicate the opposite trend. Brackets

and asterisks indicate significant differences between TuMV-G and TuMV-S disease progression or severity in the KO mutant plant genotype being considered.

For *AUSIPS*, if mutant values had higher intervals than the WT it meant that the virus is able to cause stronger symptoms in the absence of these host genes. This is the case for *at2g14080* and *nudt6* plants infected with TuMV-G. In the case of TuMV-S this happened in *at3g12850* and, as for TuMV-G, in *at2g14080* plants. An interesting observation was made for nudt6 mutant for *AUDPS*, where it showed an opposite effect in comparation with *AUSIPS*.

Therefore, significant differences in *AUDPS* confidence intervals between WT and mutant plants confirms the role in infection of the genes that were knocked-out. When the mutant had lower *AUDPS* values (e.g., *at4g02580*, *cplepa*, *nudt5*, and *nudt6* for both viral strains) it confirmed the positive function of the gene in the viral replication. While mutants with values higher than those of the WT (e.g., *ddm1* mutants for TuMV-G and *at4g10130* for TuMV-S) confirm the role of the gene in the host defense. Comparations of *AUSIPS* values between WT and mutant plants also confirms the role of most of the studied genes in symptoms severity. The *at4g02580* plants had lower AUSIPS interval in the TuMV-G infection. Therefore, plants defective in a NADH-ubiquinone oxidoreductase susceptibility factor had a milder symptomatology than WT ones. Mutants *at4g02580* and *nudt5* also had lower *AUSIPS* intervals when infected with TuMV-S. Differences in symptoms severity progression were also significant between the two viral strains in the WT, *at1g57570*, *at2g14080*, *at4g02580*, *cplepa*, *ddm1*, *mee55*, *nudt5*, and *nudt6*. This difference indicates that the two viral strains cause different symptomatology in the WT and the majority of the mutants.

3. Discussion

3.1. Comparation with previous studies

Pathogens will have different virulence and induce different responses in their hosts depending on their adaptation history. For example, in the whole-genome transcriptomic study by Hillung et al. (2016), they compared the transcriptomic responses of six arabidopsis accessions infected with generalist or specialist strains of *Tobacco etch*

potyvirus (TEV; genus *Potyvirus*). They showed that the generalist virus manipulated a similar set of host genes across the experimental host range while the specialist virus showed a more heterogeneous response. In our GWAS study, similar conclusions can be reached by comparing genes associated with infection by the generalist or specialist TuMV strains. In the case of the generalist strain TuMV-G, fewer candidate genes were identified compared to the specialist strain TuMV-S. This difference might have emerged as a consequence of the different evolutionary strategies of the two viruses or as a consequence of the GWAS analysis. If we focus on the different evolutionary strategies of the two viruses, we can say that selection has driven the generalist virus to manipulate a similar set of host genes across the host range for successful infection. In contrast, host-specific selective pressures modulated the evolution of the specialist virus, hence, more genes associated with TuMV-S have been found by the GWAS analysis.

A GWAS of TuMV infection in arabidopsis in a natural setting was recently performed by Rubio et al. (2019). None of the genes found by these authors were pinpointed in our study but this could simply reflect three major experimental differences: (1) Rubio et al. (2019) grew their plants in a natural setting where they were exposed to a changing environment. The highly complex natural setting can lead to much more heterogeneous gene regulations, as opposed to a controlled environment that minimizes external abiotic and biotic stressors. It was shown before that differences in temperature, light and water availability influence the response of the plant to a virus (Xu et al., 2008; Hily et al., 2016; González et al., 2021). Multiple stresses affecting the plant at the same time can be problematic when trying to identify genes responsible for the specific response of plants to virus infection. (2) The evolutionary histories of the TuMV strains used in both studies were largely different. While Rubio et al. (2019) used the UK1 isolate, we used strains derived from the YC5 isolate originally obtained from calla lily plants (Chen et al., 2003). (3) In our study, the 450 accession were chosen to represent the world-wide genetic diversity of the species, while French accessions were largely overrepresented in Rubio et al. (2019) study. Rubio et al. (2019) identified six new genes above a threshold of $-\log P \geq 4$ in their GWAS analysis: *RESTRICTION TO TOBACCO ETCH VIRUS MOVEMENT 3*, a DEAD box RNA helicase 1 candidate gene, *EUKARYOTIC TRANSLATION INITIATION FACTOR 3B*, a protein with a pleckstrin homology domain, a protein containing a TIM barrel domain, and a key enzyme involved in the glutamate pathway. Our study identified 13 genes specifically mapped for viral infection response

(Table C1.4), of which eight were experimentally confirmed as having roles in the plant response to TuMV-S and TuMV-G (Fig. C1.4). Despite the lack of matching genes mapped between both studies, there are similarities at the functional level: for both studies there were genes mapped that belonged to ATP-dependent DNA helicase, DnaJ domain superfamily protein and ubiquitin associated proteins.

3.2. Description of significant loci from the GWAS

Mapped genes in the GWAS belonged to categories such as F-box proteins, kinase, hydrolase, LRR family proteins, disease resistance proteins, transcription factors, lectins, helicases, ubiquitin proteases, proteins involved in iron metabolism, pentatricopeptide repeat-containing, GTPases, and berberines, all of them being involved in the plant response to infection, the viral cycle or RNA metabolism (Ge and Xia, 2008; Lee, 2008; Correa et al., 2013; Manna, 2015; Li et al., 2016; Guo et al., 2020; Herlihy et al., 2020; Huang et al., 2020). Locus *AT2G14080* was identified as significant for both viral strains in the analysis of symptoms severity. *AT2G14080* belongs to NBS-LRR genes that are the most numerous class of the *R* genes in arabidopsis. Their effector recognition LRR domains recognize specific pathogens and can lead to a HR immune response (HR) or to an extreme resistance against the virus infection. An HR restricts the pathogen at the primary infection site causing cell death followed by SAR that increases SA accumulation and expression of pathogenesis-related genes (Meyers et al., 2003; Marone et al., 2013). There were also some strain-specific hits that were previously characterized as involved in plant defense or in some important part of the viral cycle. Indeed, genes that differ between the two viruses could be targets of differential selection in evolution of specialist or generalist viruses. For example, ubiquitin protease and *TELOMERE REPEAT-BINDING PROTEIN 2* were specific responses of the plant to TuMV-G infection. While the Nudix hydrolase, NADH dehydrogenase and DNAJ heat shock proteins were specific for plants infected with TuMV-S strain.

Loci *AT4G04540* and *AT4G23280* both encode for cysteine-rich receptor-like protein kinases (*CRK39* and *CRK20*, respectively). *CRKs* are the only genes mapped in common for all three disease-related phenotypic traits. *CRK* genes are induced upon pathogen infection in *A. thaliana* via the SA signaling pathway, resulting in HR (Chen et al., 2004). In general, receptor-like protein kinases (RLKs), a large family with more than 600 members, are central players in the plant receptor kinase-mediated signaling involved in

hormonal responses pathways, cell differentiation, plant growth and development, self-incompatibility, and symbiotic and pathogen recognition (Liang and Zhou, 2018). Given their upstream role in the MAPK signaling cascades, it is not surprising that RLK expression has many pleiotropic effects on diverse plant phenotypes.

3.3. Validation analysis of LOF mutants

In the analysis of the ten selected loss-of function mutants, significant differences can be detected with *AUDPS* and *AUSIPS*, indicating that disease progression was not proportional to symptoms development in the mutants. The reason for this effect could be that the viral load in a given Col-0 mutant, as opposed to the natural accessions (González et al., 2019; Corrêa et al., 2020), is not proportional to symptoms severity. Symptom appearance and progression depends on the viral load but symptom severity might depend on the lack of an essential gene the virus might hijack to evade the defense response, not being directly related to viral load. The virus not being able to evade the defense response might activate stronger SAR which leads to stronger symptoms because of the stronger HR immune response which restricts the pathogen at the primary infection site causing cell death. There was one gene that came up in the GWAS of both strains, *AT2G14080* that had a significant effect in the mutant involved with the two strains and it appears to be involved in plant defense. Two of the ten genes selected for the mutant analysis came from TuMV-G analysis and seven came from the TuMV-S analysis. Eight of the selected genes had a significant effect on the virus disease progression and/or symptoms. *MATERNAL EFFECT EMBRYO ARREST 55* (*MEE55*, encodes for a serine and sphingolipid biosynthesis protein) and *AT1G57570* (encodes for a member of the mannose-binding lectin superfamily protein) apparently had no significant effect on either viral strain under our experimental conditions. There were five genes that had an effect in both viral strains: *AT2G14080*, *AT4G02580*, *cpLEPA*, *NUDX5*, and *NUDX6*. *AT2G14080* is an NBS-LRR resistance gene. These proteins monitor the status of plant proteins targeted by pathogens and activate a series of defense responses (McHale et al., 2006). By removing this gene, viruses managed to induce stronger symptoms. *AT4G02580* is a susceptibility factor and could aid viral pathogenesis (Kant et al. 2019). *CpLEPA* is a highly conserved chloroplastic translation factor that could assist viral transcription in the cytoplasm by enhancing the translation of chrolopastic proteins involved in photosynthesis to compensate for the negative side-effects of infection in chloroplasts activity (Ji et al., 2012; Li et al. ,2013; Sanfaçon, 2015). Mutants *nudt5* and

nudt6 are deficient in proteins that form part of the Nudix hydrolase family, which act as positive regulators in plant immunity (Ge and Xia, 2008; Yoshimura and Shigeoka, 2015), thus leading to a stronger anti-pathogen response. In both viral strains they seem to have important roles for disease progression by enhancing viral replication or gene expression since viruses replicated worse when these two genes where knocked-out. Suggesting the possibility that these two Nudix hydrolases could have additional functions besides the one described in defense. This role for the two hydrolases in viral infection was not described before.

Genes that had an effect in the loss-of function mutant analysis for *AUDPS* for virus TuMV-G were *ddm1* and *AT4G10130*. Corrêa et al. (2020) showed that *ddm1* plants were more resistant to two different strains of TuMV. This might be because induction of SA-mediated defense in *ddm1* mutants may be an explanation of their resistance to TuMV. The opposite has been noticed for geminiviruses where ddm1 mutant showed hypersusceptibility to infection. The reason for this was the methylation of viral genomes which is a plant defense mechanism; when methylation is reduced, plants are more susceptible (Raja et al. 2008). Differences in adaptation history of TuMV-G and the strains studied by Corrêa et al. (2020) might explain why TuMV-G replicates better in this mutant in our study. For TuMV-G the lack of DDM1 might help the virus replicate better since defense genes are not properly methylated and henceforth their expression deregulated. Another significant gene in the mutant analysis was *AT4G10130*, which is involved in peptidyl-diphthamide biosynthetic processes and tRNA wobble uridine modification. Both of these processes are involved in translation modifications and this protein might have a role in an anti-pathogenic response.

For the strain TuMV-S in the *AUSIPS* values, one gene had a significant effect in the mutant analysis, *AT3G12850* which is involved in regulation of JA levels. Viruses infecting *at3g12850* plants replicate better. *AT3G12850*-encoded protein is a COP9 signalosome complex-related/CSN complex-like protein. The tomato yellow leaf curl Sardinia virus (TYLCSV) C2 protein interacts with CSN5 resulting in a reduction of JA levels. As previously shown, treating *A. thaliana* plants with exogenous JA disrupts TYLCSV infection (Lozano-Durán et al., 2011). It is known that plant viruses and herbivores have strategies to manipulate JA levels as this hormone confers defenses to

the plant against biotic and abiotic stresses (Wu and Ye, 2020). This means that in our pathosystem the JA is negatively affecting the viral replication.

3.4. Genetic architecture of disease-related traits

Looking at the analysis of the underlying genetic architecture of each phenotyped trait, it was evident that some disease-related phenotypes were explained by few SNPs (infectivity and symptoms severity at 14 dpi for both viruses and *AUDPS* and infectivity at 21 dpi for both viruses as well), while some traits were highly polygenic and explained by a large number of SNPs (*AUDPS* for TuMV-S at 14 dpi). SNPs that passed the *PIP* threshold were mapped within locus *AT2G04440* (MutT/Nudix family protein) for *AUDPS* of TuMV-S at 21 dpi and position 6,685,977 in an intergenic region on chromosome 3 along with *AT3G19350*, that corresponds to the gene *MPC*, for infectivity of TuMV-G at 21 dpi. All had possible roles in the viral infection. *AT2G04440* was previously characterized as an important player in the plant immune response (Ge and Xia, 2008). *MPC* is an important translation initiation factor that binds to the viral VPg and the RdRP NIb of TuMV, affecting the viral RNA accumulation (Dufresne et al., 2008). The noncoding intergenic region at position 6,685,977 on chromosome 3 could be a promoter region involved in regulation of the expression of both *ABF4* and *FUT11*. *ABF4* controls the ABA-dependent stress response. It was previously shown that *Wheat yellow mosaic potyvirus* disturbs the ABA signaling pathway through the interaction between the viral RdRp and the wheat's light-induced protein TaLIP thus facilitating virus infection (Zhang et al., 2019). There is no clear description of *FUT11* in plant virus infection, but it is involved in protein N-linked glycosylation and the intergenic position 6,685,977 shows a strong LD ($r^2 = 1$; in a 10 kb window) with this gene.

Since the genome of arabidopsis is highly polygenic and is governed by small effect loci (as shown by the BSLMM analysis) our study might have missed some of the genes described in the literature as being involved in the potyvirus infection. Other explanation for the absence of previously described genes would be that they were not important in the context of our virus strains that were preadapted in specific arabidopsis mutants.

Altogether, this work (1) describes differences between a generalist and a specialist pathogen, (2) identifies and characterizes genes involved in a generalist and a specialist

virus infection and (3) illustrates the variability of the genetic elements involved in a viral infection depending on the evolutionary history of the viral strain.

Chapter 2: Arabidopsis genes involved in differential responses to naïve and adapted strains of TuMV

1. Introduction

Plant viruses are a constant threat to plants, causing a complex defense response involving several host genes. When a virus enters an individual host, it needs to evade host's defenses long enough to replicate and establish a successful infection. Plants mount different responses against viral infection, such as gene-for-gene resistance (Van Der Biezen and Jones, 1998; Moffett, 2009), HR local lesion responses (Loebenstein, 2009), active reprogramming of gene expression (Yang et al., 2007; Agudelo-Romero et al., 2008; Corrêa et al., 2020), oxidative bursts (Wojtaszek, 1997), and RNA silencing (Ruiz-Ferrer and Voinnet, 2009). All these responses are being orchestrated by hormonal homeostasis, where various hormones (JA, ET and SA) activate resistance genes of the plant (Soosaar et al., 2005; Carr et al., 2010).

Viruses have their own mechanisms to overcome plant defenses, for example using suppressor proteins that interfere at different levels of the RNA silencing pathway (Cheng and Wang, 2017; Rodamilans et al., 2018). They constantly evolve and manage to infect new cultivated species after spilling over from their wild reservoirs (Lefeuvre et al., 2019). In particular, RNA viruses have high mutation rates, large population sizes and short generation times, thus having great evolutionary potential and causing many new emerging diseases (Woolhouse, 2002; Anderson et al., 2004; Cleaveland et al., 2007; Holmes, 2009; Jones, 2009; Roossinck and García-Arenal, 2015; McLeish et al., 2019). This adaptability, along with the increase in vector dispersal caused by global warming, will make plant virus emergences more frequent and devastating.

A method allowing identification of important genes involved in plant defense or resistance is GWAS. Its hypothesis posits that common interacting alleles at multiple disease-predisposing loci underlie most common diseases (Bush and Moore, 2012), thus justifying the use of GWAS to identify them, connecting phenotypes with genotypes and allowing us to predict genetic risk factors for disease as well as important agronomic

traits, such as susceptibility and resistance to viral infections (Korte and Farlow, 2013). Genes identified in GWAS might have possible roles as drivers/targets of viral adaptation to the plant, determining if the virus-plant interaction will be more or less virulent.

As already mentioned in the Introduction, arabidopsis is one of the best harnessed organisms for GWAS analysis. The availability of genome data and analysis tools, makes it a very attractive organism to study various different traits. Therefore, this study was set out to identify arabidopsis genes with a role in response to viral infections and how these genes may change as a consequence of virus evolution. The pathogen organism used is TuMV. For this particular study, two strains of TuMV were selected: a first one naïve for arabidopsis, TuMV-AS; and a second one, TuMV-DV, adapted to arabidopsis (Butković et al., 2020; Corrêa et al., 2020). Both strains largely differ in the severity of symptoms they induce as well as in the viral load and in the magnitude of the perturbation induced in the plant transcriptome and methylation profiles (Corrêa et al., 2020). Genes that differentially respond to both viral strains would be good candidates to be considered as drivers/targets of viral adaptation. Furthermore, the identification of such genes can potentially lead to better understanding of the infection and ease the management of pests.

In summary, a GWAS evaluating the response of 1050 accessions of arabidopsis that were exposed to the naïve and adapted TuMV strains was conducted. The genetic variation for five disease-related phenotypic traits among the 1050 genotypes was evaluated in an effort to identify genes associated with those traits in response to each or both strains. Finally, the observed genetic associations were confirmed by studying the infection of loss-of-function mutant plants.

Also, the large peak on chromosome 2 for necrosis was further studied with various GWAS methods and its association with the necrotic phenotype in the plant host was confirmed. It appears that the minor allele of SNP 5923326 on chromosome 2 is present in most of the accessions and is the causative allele of necrotic phenotype. However, the lack of patterns in the geographical spread of the necrotic accessions could be consistent with apparent instability of the region around the locus AT2G14080. This makes it difficult to conclude anything about selective sweeps in this region.

2. Results

2.1. Heritability and genetic architecture analysis

In general, for the 510,485 SNPs tested, infectivity and *AUDPS* traits for both viruses had low *PVE* and *PGE* values while necrosis had the highest (Table C3.1).

Table C2.1. Results of the BSLMM analysis and the 89% HDI for *PVE* and *PGE* measures of trait heritability.

TuMV-AS

Phenotype	*PVE*				*PGE*				ngamma	
	mean	median	lower 80% HDI	higher 80% HDI	mean	median	lower 89% HDI	higher 89% HDI	mean	median
AUDPS	0.3801	0.3806	0.27	0.48	0.2783	0.2395	0	0.58	114.9418	119
Infectivity	0.1154	0.1114	0.04	0.19	0.2972	0.2288	0	0.7	25.514	16
Symptoms	0.206	0.2041	0.13	0.28	0.4853	0.494	0	0.89	126.3885	132
Necrosis	0.4113	0.3973	0.14	0.66	0.6599	0.685	0.35	1	5.1967	4
Resistance	0.3003	0.2805	0.06	0.51	0.3284	0.2486	0	0.78	43.5093	14

TuMV-DV

Phenotype	*PVE*				*PGE*				ngamma	
	mean	median	lower 80% HDI	higher 80% HDI	mean	median	lower 89% HDI	higher 89% HDI	mean	median
AUDPS	0.3444	0.3436	0.25	0.44	0.3497	0.3118	0	0.72	100.1864	81
Infectivity	0.1689	0.1668	0.09	0.24	0.324	0.2474	0	0.76	24.9428	17
Symptoms	0.4362	0.4358	0.36	0.52	0.594	0.6253	0.25	1	121.8052	120
Necrosis	0.7328	0.7468	0.55	0.94	0.5134	0.5052	0.1	0.9	14.7954	10
Resistance	0.5151	0.518	0.21	0.86	0.3641	0.3044	0	0.81	37.1895	14

The disease traits with the larger *PVE* heritability in the BSLMM model (~50%) were necrosis and resistance for the 510,485 SNPs tested in both viral strains (Table C2.1). Resistance had very low *PVE* in the null LMM model (Table C2.2). However, in the case of necrosis, the number of variants with large effects identified with the BSLMM among all tested SNPs was five for TuMV-AS [median 0.69 and 89% HDI (0.35 - 1.00)] and 15 for TuMV-DV [median 0.50 and 89% HDI (0.10 – 0.90)]. In the case of resistance, 44 major effect SNPs were found for the TuMV-AS infection [median 0.25 and 89% HDI (0.00 – 0.78)] and 37 for the TuMV-DV one [median 0.40 and 89% HDI (0.00 – 0.81)]. Overall, TuMV-DV had larger *PVE* and *PGE* values than TuMV-AS. Our results suggest that necrosis and resistance traits are genetically less complex and involve fewer medium-large effect SNPs. By contrast, *AUDPS*, infectivity and the severity of symptoms seem to

be far more complex traits and are determined by many small effect SNPs. The lowest null-model *PVE* values were observed for the trait resistance and the largest for *AUDPS* and necrosis (Table C2.2).

Table C2.2. Values of narrow sense heritability from LMM analysis in GEMMA.

	***PVE* estimate in the null model**	
phenotype	**ancestral**	**evolved**
AUDPS	0.44	0.38
Infectivity	0.13	0.17
Symptoms	0.21	0.44
Necrosis	0.25	0.56
Resistance	0.04	0.09

This low null-model *PVE* values mean that resistance is not strongly determined by genetics but influenced by other factors. The discrepancy between the BSLMM and null-model *PVE* values can be explained by different assumptions of the models. BSLMM is a hybrid of linear mixed models and sparse regression models that takes a polygenic background into account when estimating *PVE* for a certain trait. While the null model calculated under linear mixed models assumes that every genetic variant has normally distributed sizes and affects the phenotype. Regarding symptoms severity, it had very similar *PVE* values for the BSLMM and the LMM models, for TuMV-AS [median 0.21 and 89% HDI (0.13, 0.28)] and for TuMV-DV [median 0.44 and 89% HDI (0.36 – 0.52)]. Indicating that the trait is either controlled by many small effect variants or by large effect variants but they are found in closely related accessions.

To detect the large effect SNPs, $PIP \geq 0.25$ was used. The SNPs estimated to have a detectable large effect in necrosis were mapped within the loci *AT2G14080*, that encodes for a disease resistance protein, for TuMV-AS and *AT2G14120*, that encodes for a dynamin-related protein, for TuMV-DV. It appears that these two genes are major determinants in the necrosis development caused by TuMV and will be further discussed in the discussion.

2.2. Results of the GWAS: Arabidopsis genes significantly associated with TuMV infection

In each GWAS analysis the QQ-plots showed no inflation significance that could be caused by population structure (Fig. M3). Fig. C2.1 shows the Manhattan plots generated for both TuMV strains.

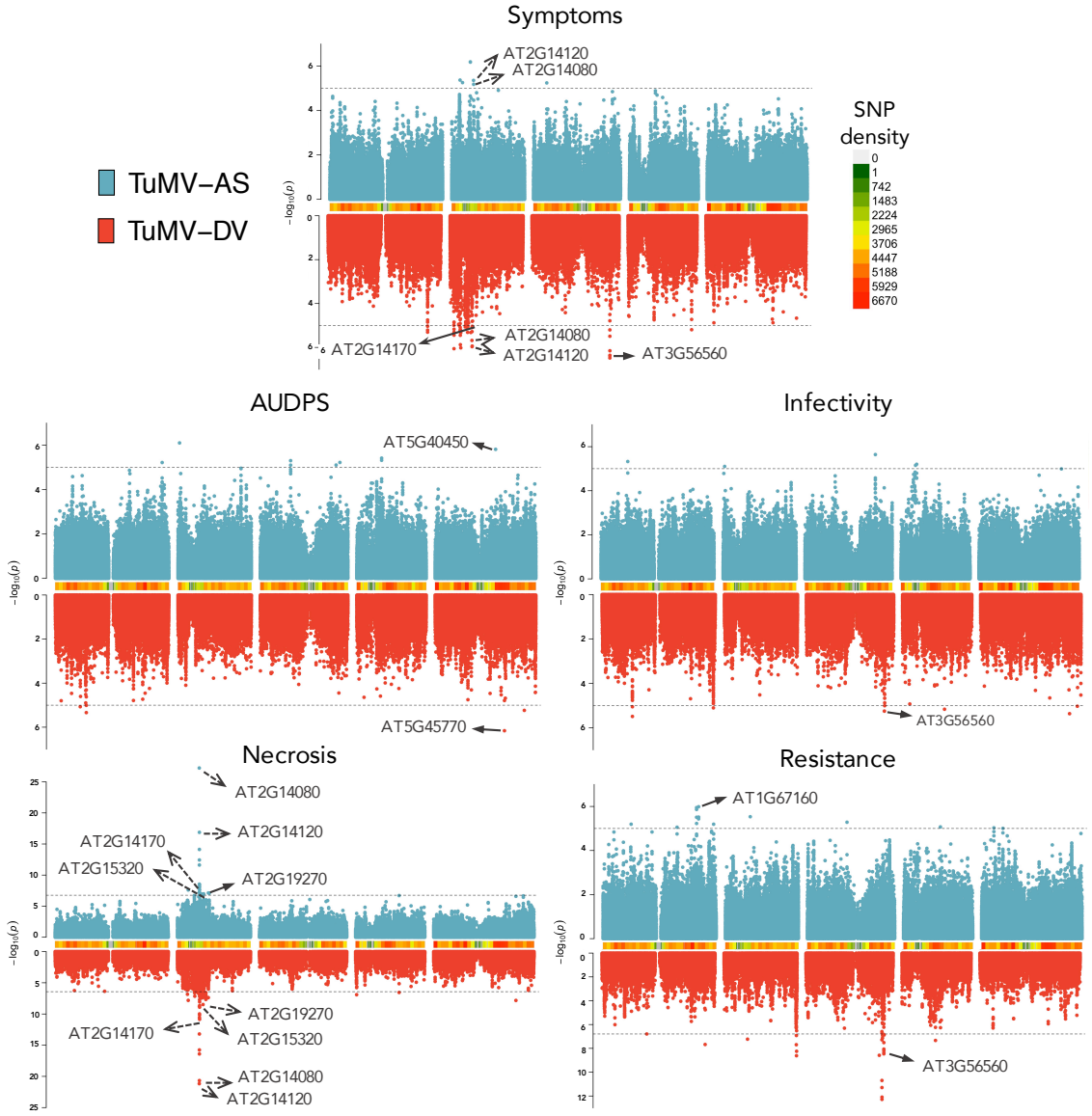


Figure C2.1. Manhattan plots for the five phenotypic traits measured in plants infected with the two TuMV strains. Peaks marked on the plots correspond to the most significant SNPs of the genes selected for the validation experiments. SNP density shows how many SNPs are genotyped for a particular chromosomal region. The dashed lines indicate the calculated cutoff value.

Significant peaks above the FDR threshold for each viral strain and genes that are marked were selected for further analysis (Table C2.3).

Table C2.3. Significant genes detected for the five phenotyped disease-related traits for each viral strain using GWAS. The genes that are shared between the two strains are bolded.

TuMV-AS			TuMV-DV		
phenotype	**gene**	**–log*P***	**phenotype**	**gene**	**–log*P***
Necrosis / Symptoms	***AT2G14080***	27.22	Necrosis / Symptoms	***AT2G14080***	21.17
Necrosis / Symptoms	***AT2G14120***	16.88	Necrosis / Symptoms	***AT2G14120***	20.67
Necrosis	*AT2G14110*	8.54	Resistance	*AT3G55180*	12.29
Necrosis	***AT2G14170***	8.18	Resistance	*AT3G55160*	11.29
Necrosis	***AT2G19270***	7.12	Necrosis / Symptoms	***AT2G14170***	10.57
Necrosis	*AT2G19890*	7.08	Resistance / Symptoms	*AT3G56580*	8.19
Necrosis	*AT2G16400*	6.98	Resistance / Symptoms / Infectivity	*AT3G56550*	6.83
Necrosis	***AT2G15320***	6.95	Resistance	*AT2G47630*	8.62
Necrosis	***AT2G15900***	6.81	Resistance	*AT3G53235*	8.59
Symptoms	*AT2G12475*	6.19	Resistance	*AT3G53240*	8.59
AUDPS /infectivity	*AT2G01990*	6.10	Necrosis	***AT2G15900***	8.48
Resistance	*AT1G67160*	5.95	Resistance / Symptoms / Infectivity	*AT3G56560*	8.46
AUDPS	*AT5G40450*	5.82	Resistance / Symptoms	*AT3G56570*	8.07
Infectivity	*AT3G50960*	5.64	Resistance	*AT2G47770*	8.25
Resistance	*AT1G67170*	5.51	Necrosis	*AT5G54390*	7.83
Resistance	*AT1G68460*	5.44	Resistance	*AT2G47485*	7.69
AUDPS	*AT4G10800*	5.42	Necrosis	*AT2G17700*	7.68
Symptoms	*AT2G05940*	5.37	Necrosis	*AT2G17860*	7.60
Infectivity	*AT1G20735*	5.33	Necrosis	***AT2G15320***	7.42
AUDPS	*AT3G22920*	5.31	Necrosis	***AT2G19270***	7.38
Resistance	*AT3G28415*	5.28	Resistance	*AT4G15880*	7.34
Symptoms	*AT3G11400*	5.24	Necrosis	*AT2G16220*	7.33
AUDPS	*AT3G57570*	5.22	Resistance	*AT3G55150*	7.28
AUDPS	*AT1G76250*	5.22	Necrosis	*AT2G17690*	7.28
Infectivity	*AT4G07410*	5.20	Necrosis	*AT2G17695*	7.28

Resistance	*AT1G78980*	5.19	Resistance	*AT2G14825*	7.23
Infectivity	*AT4G06676*	5.15	Necrosis	*AT2G15220*	7.11
AUDPS	*AT3G54810*	5.11	Necrosis	*AT2G17730*	6.99
Resistance	*AT1G53310*	5.06	Necrosis	*AT2G16380*	6.97
Resistance	*AT5G10490*	5.03	Necrosis	*AT2G17720*	6.97
Resistance	*AT5G17370*	5.00	Necrosis	*AT2G14510*	6.94
			Resistance	*AT2G09935*	6.92
			Resistance	*AT2G47560*	6.92
			Resistance	*AT3G55580*	6.92
			Symptoms	*AT3G56540*	5.20
			Necrosis	*AT2G14530*	6.75
			Necrosis	*AT2G14960*	6.47
			Necrosis	*AT2G11890*	6.46
			Necrosis	*AT2G16592*	6.46
			AUDPS	*AT5G45770*	6.16
			Symptoms	*AT2G03600*	6.06
			Symptoms	*AT2G07050*	6.02
			Infectivity	*AT1G23020*	5.49
			Infectivity / AUDPS	*AT5G59180*	5.37
			Symptoms	*AT2G07020*	5.33
			Symptoms	*AT1G71040*	5.31
			Symptoms	*AT2G06990*	5.23
			Symptoms	*AT4G36160*	5.19
			Infectivity	*AT4G21100*	5.17
			Infectivity	*AT1G79500*	5.11
			AUDPS	*AT1G19310*	5.07
			AUDPS	*AT1G23380*	5.01
			Symptoms	*AT2G10602*	5.01

There were six loci shared between the two viral strains and 47 unique loci for TuMV-DV and 25 for TuMV-AS that were significant. For traits symptoms and necrosis there were loci shared between viral strains. In the case of symptoms, two loci were shared between both viruses: *AT2G14080* and *AT2G14120* (Fig. C2.2A). For necrosis, there were six loci shared; *AT2G14080*, *AT2G14120*, *AT2G14170*, *AT2G19270*, *AT2G15320*, and *AT2G15900* (Fig. C2.2A). Loci shared between different phenotyped traits of both

viruses were *AT2G14080*, *AT2G14120* and *AT2G14170* for traits symptoms and necrosis, *AT5G59180* and *AT2G01990* for *AUDPS* and Infectivity, *AT3G56550* and *AT3G56560* for resistance, infectivity and symptoms, and *AT3G56580*, *AT3G56550*, *AT3G56560*, and *AT3G56570* for symptoms and resistance (Fig. C2.2B).

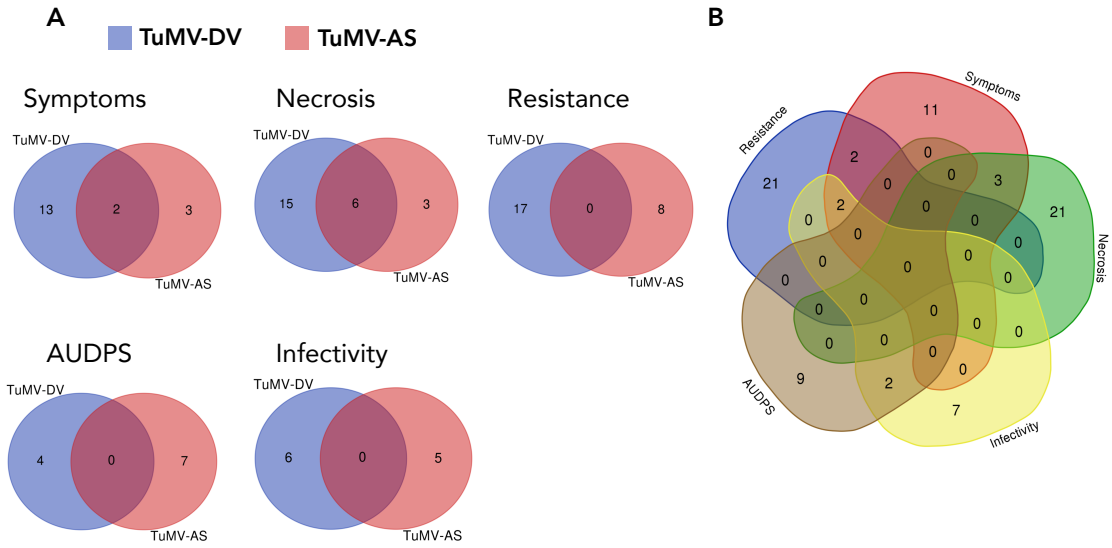


Figure C2.2. Venn diagram showing the number of unique and shared genes. (A) Genes mapped for each viral strain and disease-related traits. (B) Genes mapped for all disease-related traits pooling together both viral isolates.

In total, 31 host genes were significantly correlated with the TuMV-AS infection and 53 genes with the TuMV-DV one (Table C2.1).

There are six loci shared between TuMV-AS and TuMV-DV: *AT2G14080*, *AT2G14120*, *AT2G14170*, *AT2G15320*, *AT2G15900*, and *AT2G19270*, (Table C2.1). All of them were significant in the GWAS analysis of necrosis and two (*AT2G14080* and *AT2G14120*) were also found to be significant in the analysis of severity of symptoms. From the six shared loci five were selected for further experimental validation, they were: *AT2G14080*, *AT2G14120*, *AT2G14170*, *AT2G15320*, and *AT2G19270*.

In the replicated study we performed a GWA on four phenotypes per each virus: (1) necrosis binary (2) percent necrosis per total number of plants (3) percent necrosis per infected plants (4) *AUSIPS*, where necrosis binary trait for TuMV-DV showed too

inflated results (Fig. C2.2). Therefore, results of this GWAS were removed from further analysis. There were mapped SNPs above the threshold of $-\log P > 8$ within 33 genes for both viruses, five of which were shared with the original GWAS (Table C2.4, Fig. C2.2). In the case of the ancestral virus 33 accessions from 51 (64.71%) had at least one necrotic plant at 21 dpi. For the evolved virus 36 accessions from 51 (70.59%) had at least one necrotic plant at 21 dpi. For the 67 previously non-necrotic accessions one had necrotic symptoms (at least in one plant of the accession) for the ancestral virus (1.49%) and nine (13.43%) for the evolved virus (at least in one plant of the accession).

Table C2.4. Significant genes detected above $-\log P > 8$ in the replicated experiment for four disease related phenotypes for each viral strain using GWAS.

TUMV-AS

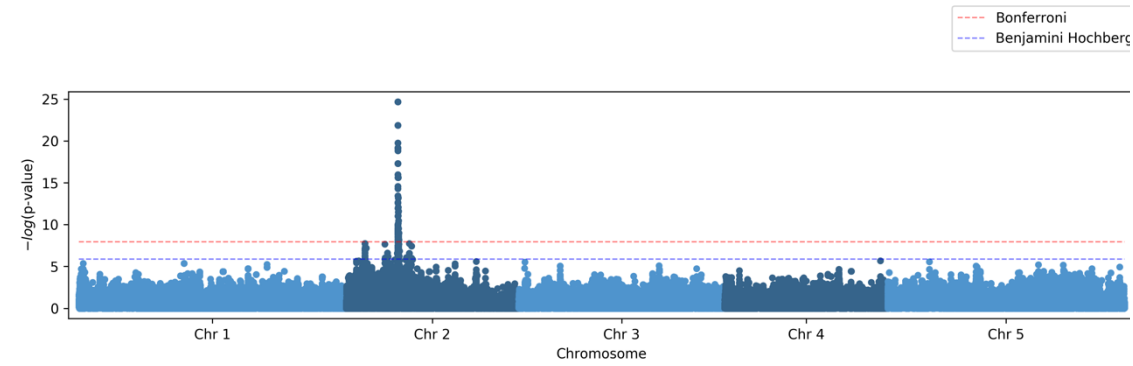
Identifier	Gene
AT2G07042.1	other_RNA
AT2G16490.1	XH domain-containing protein
AT2G07020.1	kinase with adenine nucleotide alpha hydrolases-like domain-containing protein
AT2G14110.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
AT2G14160.1	RNA-binding (RRM/RBD/RNP motifs) family protein
AT2G14080.1	Disease resistance protein (TIR-NBS-LRR class) family
AT2G16405.1	Transducin/WD40 repeat-like superfamily protein
AT2G14060.1	encodes a protein whose sequence is similar to SAM:salicylic acid carboxyl methyltransferase (SAMT) (GI:6002712)(Clarki
AT2G14095.1	hypothetical protein
AT2G07120.1	F-box associated ubiquitination effector family protein
AT2G14170.1	Arabidopsis thaliana methylmalonate-semialdehyde dehydrogenase protein_coding ALDEHYDE DEHYDROGENASE 6B2 (ALDH6B2)
AT2G16400.1	BEL1-LIKE HOMEODOMAIN 7 (BLH7) BEL1-LIKE HOMEODOMAIN 7
AT2G14120.3	Encodes a dynamin related protein. DRPs are self-assembling GTPasse involved in fission and fusion of membranes. DRP3B
AT2G14290.1	F-BOX/DUF295 BRASSICEAE-SPECIFIC 13 (ATFDB13)
AT2G15370.1	Predicted fucosyltransferase, based on similarity to FUT1, but not functionally redundant with FUT1. protein_coding
AT2G14680.2	MATERNAL EFFECT EMBRYO ARREST 13 (MEE13)
AT2G16430.2	Encodes an acid phosphatase involved plant acclimation to Pi deprivation

TUMV-DV

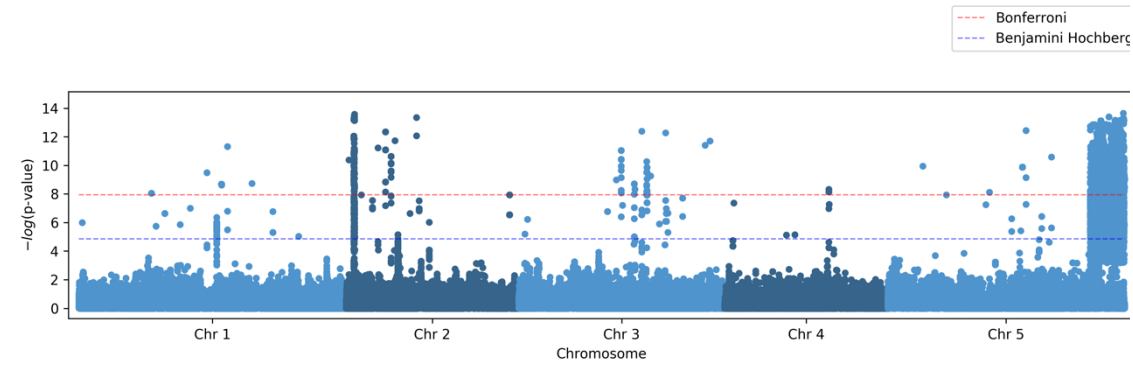
Identifier	Gene
AT2G13900.1	Cysteine/Histidine-rich C1 domain family protein
AT2G14110.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
AT2G13965.1	transmembrane protein
AT2G13640.1	Transcription factor IIS family protein
AT2G14080.1	Disease resistance protein (TIR-NBS-LRR class) family
AT2G14390.1	hypothetical protein

AT2G14095.1	hypothetical protein
AT2G13960.1	Homeodomain-like superfamily protein
AT2G13810.1	ALD1 is a L-lysine alpha-aminotransferase. It is part of the pipecolic acid biosynthetic pathway, where it catalyzes th
AT2G14170.1	Arabidopsis thaliana methylmalonate-semialdehyde dehydrogenase protein_coding ALDEHYDE DEHYDROGENASE 6B2 (ALDH6B2)
AT2G14120.3	Encodes a dynamin related protein. DRPs are self-assembling GTPasse involved in fission and fusion of membranes. DRP3B
AT2G13650.1	Encodes a Golgi-localized GDP-mannose transporter. It can transport ADP-glucose in vitro.
AT3G02990.1	member of Heat Stress Transcription Factor (Hsf) family The mRNA is cell-to-cell mobile
AT2G05990.1	Encodes enoyl-ACP reductase a component of the fatty acid synthase complex.
AT2G06040.1	Contributes to UV tolerance through nucleotide excision repair.
AT2G13800.1	somatic embryogenesis receptor-like kinase 5

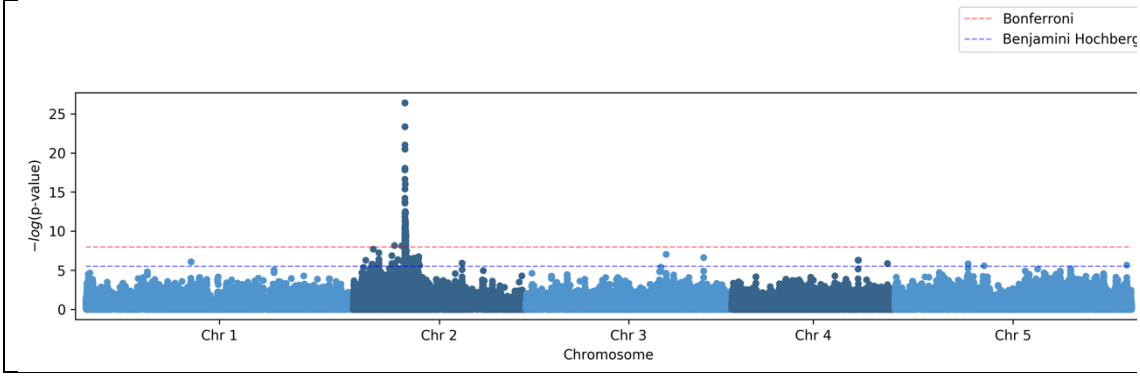
TuMV-AS Necrosis binary



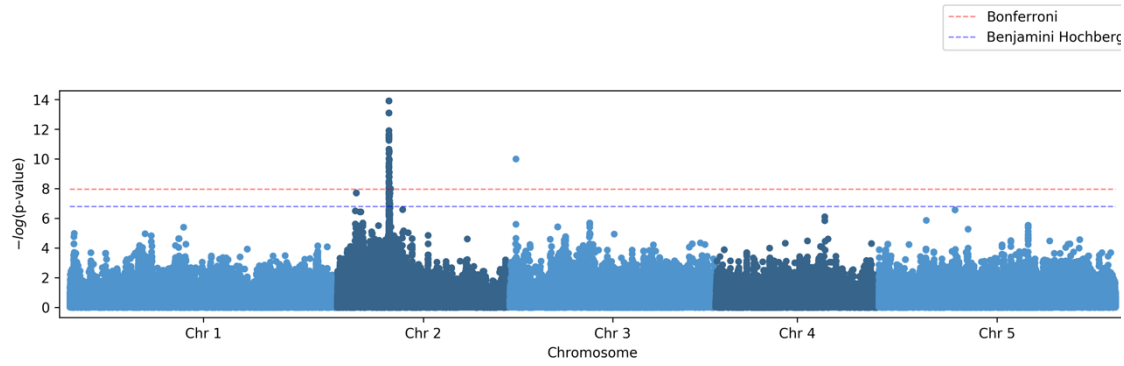
TuMV-DV Necrosis binary



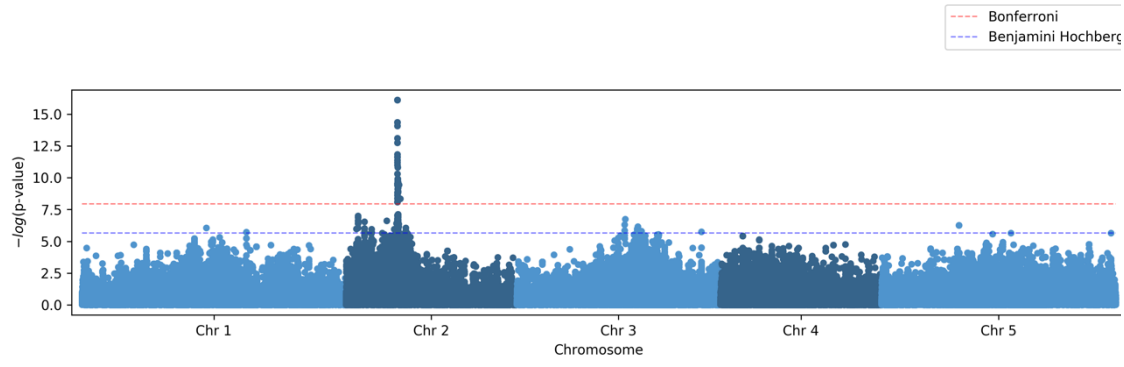
TuMV-AS percent necrosis per infected plants



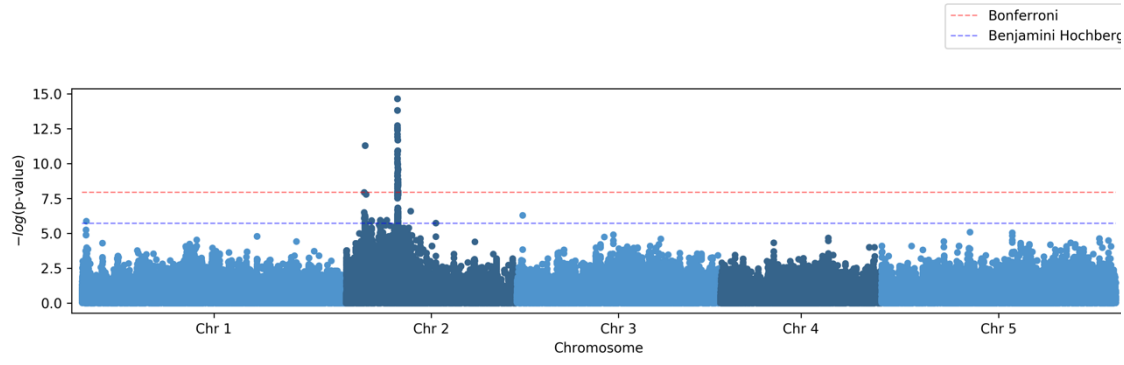
TuMV-DV percent necrosis per infected plants



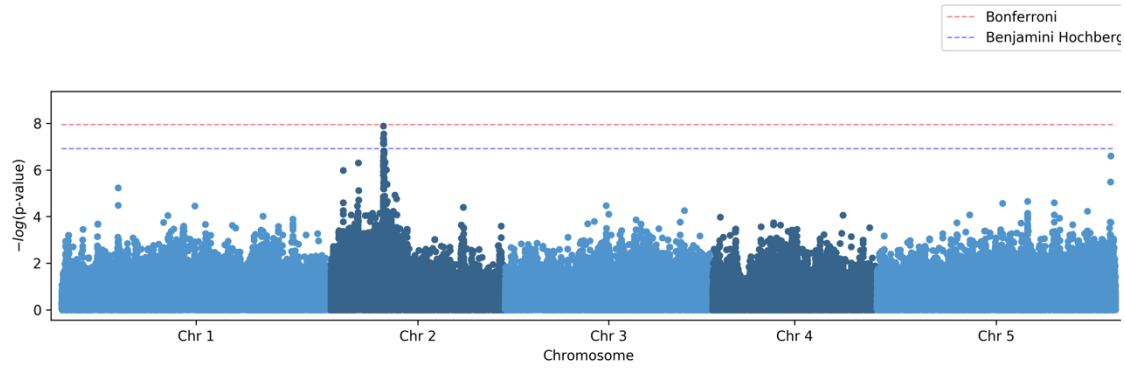
TuMV-AS percent necrosis per total number of plants



TuMV-DV percent necrosis per total number of plants



TuMV-AS AUSIPS



TuMV-DV AUSIPS

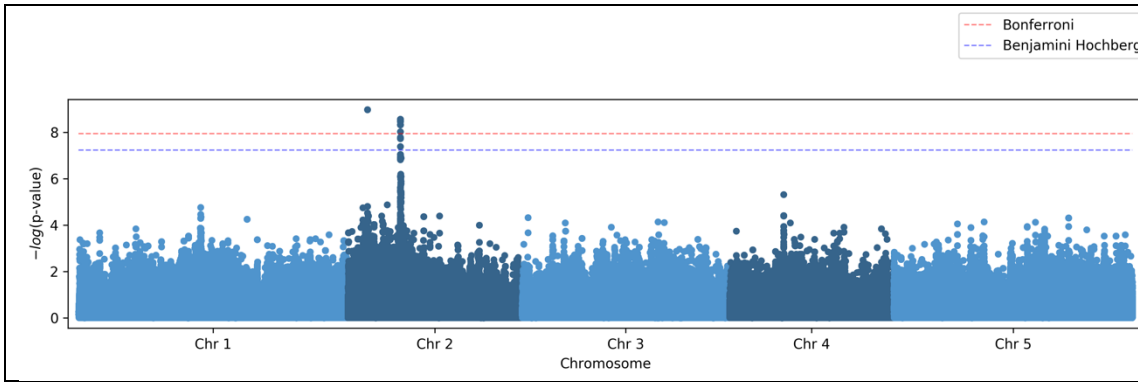


Figure C2.2. Manhattan plots of the replicated GWAS for 118 accessions for four phenotyped traits per virus. The purple line marks the Benjamini-Hochberg threshold and the red line the Bonferroni correction threshold.

2.3. Experimental validation of GWAS results

Nine candidate genes were chosen for experimental validation (Fig. C2.1, Table M3) and the selected loci were involved in different functions assumed to be important for the virus; five were detected for both TuMV strains (*AT2G14080*, *DRP3B*, *ALDH6B2*, *AT2G15320*, and *AT2G19270*) and two were specific to each strain (*FDB5* and *AT5G40450* for TuMV-AS; *NAC065* and *AT5G45770* for TuMV-DV). As expected, TuMV-DV shows significantly larger *AUDPS* and *AUSIPS* values than TuMV-AS in WT plants, thus confirming TuMV-DV is more virulent than its ancestor.

Firstly, let's compare the way disease progressed in the LOF mutant plants *vs* WT plants based on the *AUDPS* and *AUSIPS* 89% HDI intervals (Table C2.5). In the case of the naïve TuMV-AS, both *AUDPS* and *AUSIPS* values were significantly different between WT plants and mutants *at2g14080* and *at2g15320* (non-overlapping 89% HDI; indicated with blue arrows in Fig. C2.3 and +/– sign in Table C2.5), but not so in the other seven LOF mutants.

Table C2.5. 89% HDIs calculated for *AUDPS* and *AUSIPS* for TuMV-AS and TuMV-DV on each KO mutant and WT plant. A +/– next to a row marks the 89% HDI values that are higher (+) or lower (–) in the mutants compared to the WT.

AUDPS

	TuMV-AS			TuMV-DV			
	WT	4.50	11.70	**WT**	16.40	16.90	
KIB3	***at1G67160***	10.20	14.80	***at1G67160***	9.10	15.50	–

AT2G14080	***at2G14080***	11.90	14.20	+	***at2G14080***	11.30	15.70	–
DRP3B	***at2G14120***	2.90	10.50		***at2G14120***	4.40	12.00	–
ALDH6B2	***at2G14170***	7.80	13.70		***at2G14170***	11.30	15.90	–
AT2G15320	***at2G15320***	0.00	2.80	–	***at2G15320***	9.90	14.60	–
AT2G19270	***at2G19270***	0.00	5.80		***at2G19270***	1.40	7.60	–
NAC065	***at3G56560***	2.60	10.20		***at3G56560***	2.80	10.20	–
RBB1	***at5G40450***	5.40	11.60		***at5G40450***	15.50	16.90	
RLP55	***at5G45770***	7.30	13.20		***at5G45770***	12.30	17.00	

				AUSIPS				
		TuMV-AS				**TuMV-DV**		
	WT	7.90	21.20		**WT**	**29.70**	33.30	
KIB3	***at1G67160***	19.10	29.00		***at1G67160***	**14.90**	25.70	–
AT2G14080	***at2G14080***	26.50	33.90	+	***at2G14080***	**21.30**	32.80	
DRP3B	***at2G14120***	5.70	20.80		***at2G14120***	**8.80**	24.50	–
ALDH6B2	***at2G14170***	13.90	23.90		***at2G14170***	**19.90**	28.20	–
AT2G15320	***at2G15320***	0.00	5.00	–	***at2G15320***	**14.40**	23.90	–
AT2G19270	***at2G19270***	0.00	11.40		***at2G19270***	**2.40**	16.20	–
NAC065	***at3G56560***	4.70	17.20		***at3G56560***	**4.20**	16.50	–
RBB1	***at5G40450***	10.90	23.70		***at5G40450***	**32.10**	35.80	
RLP55	***at5G45770***	13.80	24.90		***at5G45770***	**24.10**	34.80	

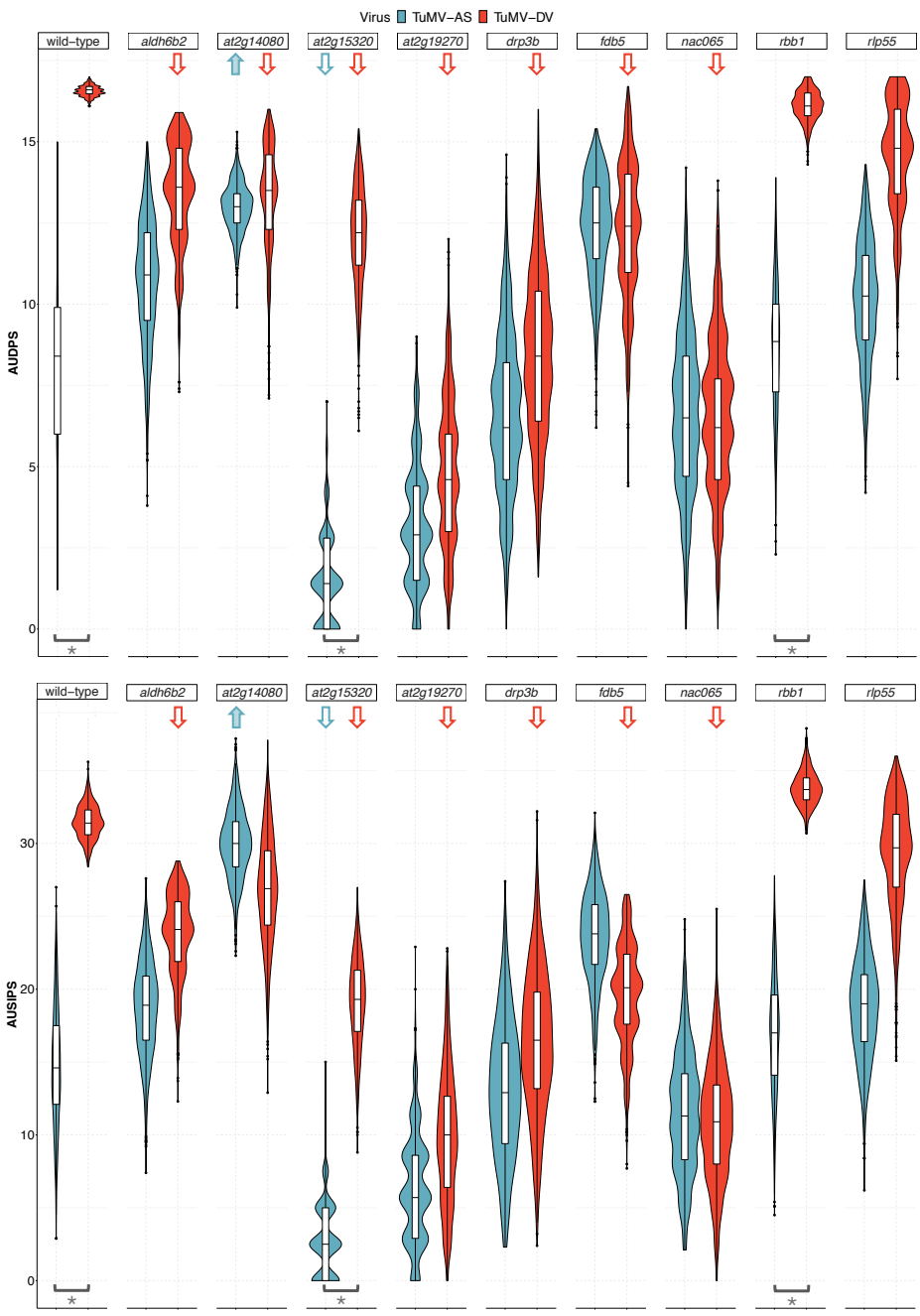


Figure C2.3. 89% HDI calculated for *AUDP*S and *AUSIPS* for each viral strain on each LOF mutant plant genotype. Not overlapping 89% HDI between a given mutant and the WT plants is indicated by an arrow. Arrows pointing up indicate a significant positive difference in medians, while arrows pointing down indicate the opposite trend. Brackets and asterisks indicate significant differences between TuMV-AS and TuMV-DV disease progression or severity in the LOF mutant plant genotype being considered.

Both of these mutants affect proteins of the LRR family involved in disease resistance and were identified in the GWAS of both viral strains. Interestingly, TuMV-AS showed

faster disease progression and stronger symptomatology in *at2g14080* than in the WT plants, supporting the idea that locus *AT2G14080* is involved in the defense response against TuMV. In sharp contrast, TuMV-AS was significantly less virulent in *at2g15320* than in the WT plants, suggesting locus *AT2G15320* is somehow enhancing TuMV-AS replication. Infection of LOF mutants in the two genes selected as specifically involved in TuMV-AS infection (*fdb5* and *rbb1*) showed no significant differences with values observed in WT plants. In the case of the adapted strain, TuMV-DV, six of the LOF mutants tested had a significant negative effect both in *AUDPS* and *AUSIPS* (non-overlapping 89% HDI; indicated with red arrows in Fig. C2.3 and – sign in Table C2.5); in *at2g14080* the negative effect was only significant for *AUDPS*. These observations suggest that all these genes shall be positively involved in infection, as their LOF results in a slowly progressing infection and with weaker symptoms than in the WT plants. Disease progression of LOF mutants *rbb1* and *rlp55* (the latter selected because the GWAS pointed towards a TuMV-DV-specific effect) was not significantly different from what was observed in the WT plants.

Second, lets now compare the effect of the different LOF mutations on the relative performance of the two TuMV strains. Comparing the 89% HDI for the two viral strains on each plant genotype, TuMV-AS and TuMV-DV only significantly differ in the LOF mutant genotypes *at2g15320* and *rbb1*, in both cases TuMV-DV showing faster progression and more severe disease. Above was mentioned that TuMV-AS infection depended on the expression of locus *AT2G15320* (encoding an LRR domain containing protein involved in pathogen recognition), here our hypothesis is that during the course of virus evolution, this dependence has been relaxed, as the adapted TuMV-DV now replicates well in LOF plants for this gene.

2.3. Multiple-trait GWAS analysis of disease-related traits

With the multiple-trait analysis of direct effects (common response to both viruses) for symptoms severity and necrosis traits of the both viral isolates for the 1050 accessions, there is a clear strong association with the SNP 5923326 on chromosome 2 (Fig. C2.4A). It appears that the minor allele of SNP 5923326 on chromosome 2 is present in most of the accessions that showed necrosis (Fig. C2.5) and it is found at a 10,2% global frequency. The position of 5923326 is 2kb from the strongly associated loci *AT2G14080* that was previously found for necrosis and symptoms severity in a GWAS of 1050

accessions. This locus is important because the second strongest hit of this analysis was mapped within *AT2G14080* in position 5928864 on chromosome 2. For traits *AUDPS* and infectivity there were no significant SNPs found in the multiple-trait analysis (not shown). Also, there was no evidence that plants respond differentially to two viral strains since there is a lack of associated SNPs with the two viral strains in the pleiotropic effects analysis (Fig. C2.4B).

In the case of the second experiment with 118 accessions, the multiple-trait analysis recapitulated the second strongest hit in the 1050 accessions analysis, the SNP 5928864 on chromosome 2 (Fig. C2.4C). The rest of the strongly associated SNPs in this GWAS analysis are all mapped within loci *AT2G14080*. Considering how GWAS methods are sensitive to confounding and linear mixed models are sensitive to the assumptions about the distribution of the response variable, a test measuring how much this matters in the case of the necrosis trait was performed. A Kruskall-Wallis test at each locus without consideration for population structure and assumptions about distribution, returned the same significant SNPs as in the previous GWAS analysis for both viral isolates. Thus, confirming the association of the SNP 5928864 with the necrosis trait.

Figure C2.4. Multiple-trait GWAS analyses showing effects of each SNP on necrosis following treatment with the ancestral or evolved strain; A) direct effects or effects common to the plant response to both strains for 1050 accessions, B) pleiotropic effects or different responses of the plant to both strains for 1050 accessions, C) direct effects of each SNP on necrosis following treatment with the ancestral or evolved strain for the 118 replicated accessions, D) multi-locus GWAS conditioning for SNP 5923326 at chromosome 2 for 1050 accessions, showing the direct effects of each SNP on necrosis following treatment with the ancestral or evolved strain.

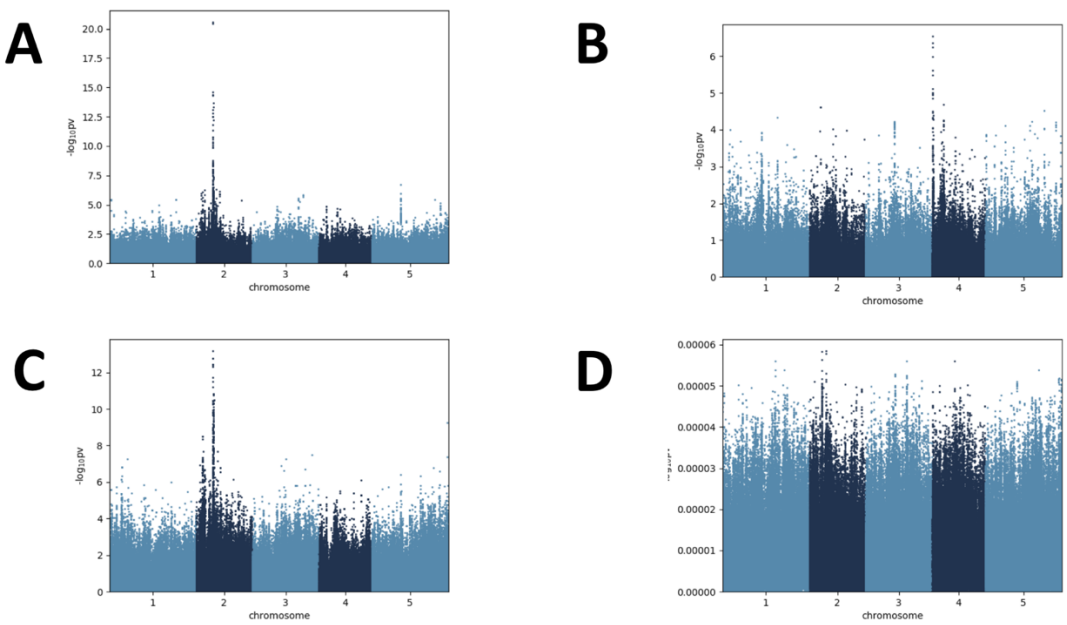
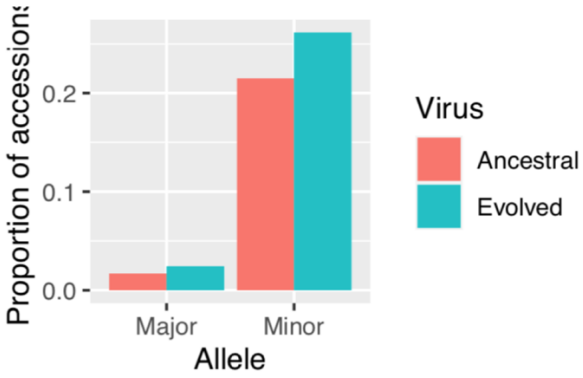


Figure C2.5. Proportion of accessions showing necrosis in any plant for 1050 genotypes at position 5923326 on chromosome 2. The ancestral virus that is marked in red refers to TuMV-AS and the evolved virus marked in green refers to TuMV-DV.



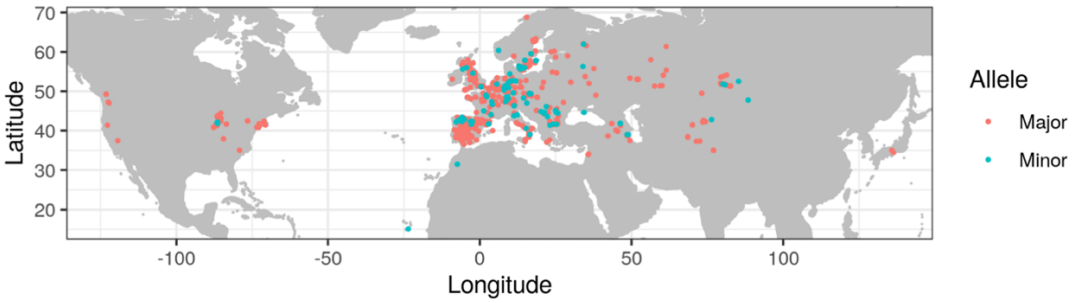
2.4. Multi-locus mixed model of the SNP 5923326 on chromosome 2

By conditioning the GWA analysis on SNP 5923326 on chromosome 2 for necrosis trait in both viral isolates, we did not find any more significant peaks (Fig. C2.4D). Thus, removing evidence of the presence of allelic heterogeneity or presence of other causative alleles at this locus.

2.5. Distribution of SNP 5923326 on chromosome 2

The geographical spread of the necrotic accessions showed no clear pattern (Fig. C2.6). This could be consistent with repeated mutations and the apparent instability of the region around the locus *AT2G14080*.

Figure C2.6. Geographic distribution of major and minor alleles at position 5923326 on chromosome 2 is shown. The major allele is shown in red while the minor allele is shown in green. The minor allele is associated with increased necrosis.



3. Discussion

3.1. Previous studies

GWAS have become an increasingly common approach to identify candidate loci underlying observed phenotypic variability. Considering its popularity this is the first GWAS focusing on arabidopsis response to a viral infection that deals with such a large number of accessions (1050). Notice that the only previous attempt to use GWAS to identify arabidopsis genes interacting with TuMV was rather limited in size, with only 317 accessions (Rubio et al. 2019). Using 1050 accessions gave us a strong resolution to identify interesting novel genes involved in the viral infection response of arabidopsis.

In a previous phenomic and transcriptomic study Sánchez et al. (2015) described that two different strains of TuMV, UK1 and JPN1, widely differ in symptoms severity, flower development and organization of vascular bundles. TuMV-UK1 proved to be more pathogenic, causing more severe symptoms and abnormalities in flower development. Interestingly, a parallelism exists between TuMV-UK1 and TuMV-DV. In both cases, genes related to metabolism, stress responses, chloroplast transport, and calcium-

mediated signaling processes, were involved in the severe phenotype. Furthermore, the number of genes with altered expression was larger for TuMV-UK1 than for the mild TuMV-JP1 strain, also paralleling the difference between TuMV-DV and TuMV-AS.

Transcriptomic studies sought to identify differences in gene expression and generate lists of differentially expressed genes (DEGs) under certain conditions. By contrast, GWAS pursues the identification of SNPs that could have an impact in phenotypic traits but without informing about the expression levels of the corresponding genes. Ultimately, however, both approaches should converge into a common picture of the underlying genetic architecture of complex phenotypes. Our study and experimental approach is complementary to the recent study by Corrêa et al. (2020) in which the transcriptome and methylome of arabidopsis plants infected with TuMV-AS and TuMV-DV were compared. As in here, they also observed many more DEGs for TuMV-DV than for TuMV-AS, as well as suppression of genes involved in biotic stress. The overlap between the lists of loci identified in our study and DEGs in Corrêa et al. (2020) was rather small: seven in the case of TuMV-AS and 31 in the case of TuMV-DV (Table C2.6).

Table C2.6. Comparation of significant genes with the transcriptomic study by Corrêa et al. (2020).

TuMV-AS	
gene	**description**
AT1G53310	Phosphoenolpyruvate carboxylase 1
AT1G67170	Protein FLX-like 2
AT2G05940	Serine/threonine-protein kinase RIPK
AT2G14120	Dynamin related protein (DRP3B)
AT2G15900	Phox domain-containing protein
AT2G19270	mitotic checkpoint protein PRCC-carboxy-term protein;
AT3G50960	Thioredoxin domain-containing protein PLP3A
TuMV-DV	
gene	**description**
AT1G02750	Protein DEHYDRATION-INDUCED 19 homolog 2
AT1G04870	PRMT10
AT1G09820	Pentatricopeptide repeat-containing protein
AT1G09840	Shaggy-related protein kinase kappa
AT1G19310	At1g19310/F18O14_14
AT1G23020	Ferric reduction oxidase 3, mitochondrial
AT1G71040	Multicopper oxidase LPR2

AT1G79500	2-dehydro-3-deoxyphosphooctonate aldolase 1
AT2G14080	Disease resistance protein (TIR-NBS-LRR class) family
AT2G14120	Dynamin related protein (DRP3B)
AT2G15042	Leucine-rich repeat (LRR) family protein
AT2G15220	At2g15220/F15A23.4
AT2G16380	Phosphatidylinositol/phosphatidylcholine transfer protein SFH7
AT2G17720	Prolyl 4-hydroxylase 5
AT2G17730	NEP-interacting protein 2
AT2G19270	mitotic checkpoint protein PRCC-carboxy-term protein;
AT2G47485	At2g47485
AT2G47560	RING-H2 finger protein ATL64
AT2G47630	Alpha/beta-Hydrolases superfamily protein
AT3G53235	unknown protein; Ha.
AT3G53240	Receptor like protein 45
AT3G55580	Regulator of chromosome condensation (RCC1) family protein
AT4G09760	Probable choline kinase 3
AT4G11090	Protein trichome birefringence-like 23
AT4G16680	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT4G28370	Transmembrane E3 ubiquitin-protein ligase FLY1
AT4G35240	Protein of unknown function (DUF630 and DUF632)
AT5G11150	vesicle-associated membrane protein 713
AT5G19850	Alpha/beta-Hydrolases superfamily protein
AT5G46450	Disease resistance protein (TIR-NBS-LRR class) family
AT5G46470	Disease resistance protein RPS6

Transcriptomic analyses will only detect significant DEGs in response to viral infection, regardless this change is a direct response to infection or a pleiotropic consequence of changes in the expression of other genes. Obviously, a host may have genes involved in viral defense that do not significantly modify their expression patterns. As GWAS looks for causal loci responding to a certain trait, GWAS will find genes that are involved in viral infection response independently of their expression changes. Therefore, GWAS may allow identification of genes that would not correspond to DEGs in a transcriptomic analysis. The biological functions of the differentially expressed genes found in the transcriptomic analysis were as hydrolases, kinases, dynamin related proteins, proteins involved in microtubule formation, signaling, tolerance to metal ions, LRR family proteins, pentatricopeptide repeat-containing proteins, and disease resistance proteins. Interestingly, two DEGs were shared between TuMV-AS and TuMV-DV and also mapped in the GWAS: *DRP3B* and *AT2G19270*, already discussed above. Corrêa et al. (2020) transcriptomic analyses also found TuMV-DV-specific differential expression of

locus *AT2G14080* discussed above. All these shared loci were validated with the LOF mutant analysis and showed significant effect on the virus symptom development and disease progression.

In a transcriptomic study comparing two strains of TEV that also differed in their degree of adaptation to arabidopsis, Agudelo-Romero et al. (2008) identified that most of the genes that were up- or down-regulated by the naïve and the adapted TEV strains were genes corresponding to biotic and abiotic stresses and defense responses. These results are in good agreement with all the transcriptomic experiments described in the previous paragraphs comparing naïve and well-adapted TuMV strains, which suggests a common mechanism of potyviruses adaptation to arabidopsis, in all cases adaptation involving evasion of certain plant defense mechanisms. Here, the GWAS approach has shown its potential to highlight similar genes, as genes involved into the biotic, abiotic, and defense responses were also found. Furthermore, our validation experiments have shown that the naïve TuMV-AS strains was controlled by TIR-NBS-LRR genes to which TuMV-DV was, at least, insensitive.

3.2. Significant genes in the GWAS

Let's summarize the results of this study. Necrosis and resistance were the only traits for which large effect SNPs have been detected. The other measured traits were highly polygenic, with low *PVE* and *PGE* values. Looking at the BSLMM heritability estimates in Table C2.1, it can be concluded that the disease traits *AUDPS* and infectivity are governed by many loci of small effect (~100 SNPs for AUDPS and ~25 for infectivity), each explaining a small percentage of *PVE*. Regarding symptoms severity, it had very similar *PVE* values for the BSLMM and the LMM models, [mean 0.21 and 89% HDI (0.13 – 0.28)] for TuMV-AS and [mean 0.44 and 89% HDI (0.36 – 0.52)] for TuMV-DV. Indicating that the trait is either controlled by many small effect variants or by large effect variants but they are found in closely related accessions. In summary, our results suggest that necrosis and resistance traits are genetically less complex and involve fewer medium-large effect SNPs. By contrast, *AUDPS*, infectivity and the severity of symptoms seem to be far more complex traits and are determined by many small effect SNPs.

The goal of our study was two-fold, first to identify new arabidopsis genes that might be involved in the interaction with its natural pathogen TuMV and, second, to highlight

among those genes which ones may respond in a different manner to two strains of the virus that differ in their degree of host adaptation, hence being possible drivers of viral evolution. There were only six arabidopsis mapped genes shared between the naïve TuMV-AS and the adapted TuMV-DV strains; all six common genes have been previously described as involved in plant responses to viral infections (the already mentioned TIR-NBS-LRR and DRP3B proteins, a second LRR protein, ALDH6B2, SNX4, and a mitotic checkpoint protein). The number of TuMV strain-specific genes found was significantly larger for TuMV-DV (47) than for TuMV-AS (25) (Fisher's exact test, $P < 0.0001$). Among the loci that had a pleiotropic effect on more than one disease-related trait and that were specific for TuMV-DV (*i.e.*, those whose implication in the virus' infection cycle was acquired as a consequence of the adaptation), it is worth mentioning methyltransferases, helicases, ubiquitin proteases, NAC-domain containing proteins, and proteins involved in iron metabolism. Interestingly, many well-known interactors of potyviruses (*e.g.*, eukaryotic translation initiation factors eIF4E and eIF4G, several heat-sock proteins or the DNA-binding phosphatase 1) were not mapped for any of the two viral strains. This could simply reflect a lack of allelic variation segregating for those genes in the arabidopsis populations included in the study or that the measured phenotypic traits are highly polygenic and cannot be explained by a few loci of major effect.

The significant SNPs for both viruses were mapped within diverse genes that were mostly involved in functions already described in relation to viral infection cycle. However, two genes have been characterized as related to TuMV infection cycle for the first time: the TIR-NBS-LRR family disease resistance protein encoded by locus *AT2G14080* and gene *DYNAMINE-RELATED PROTEIN 3B* (*DRP3B*). Disease resistance TIR-NBS-LRR class proteins with a N-terminal Toll-domain show similarity to the nucleotide binding site and other domains of plant resistance proteins in the NBS-LRR family. Members of this gene family are found numerously in clusters in the genome followed by duplication and amplification events and are the most numerous class of the *R* genes in arabidopsis (Meyers et al., 2003). NBS-LRR protein family can lead to HR which restricts the pathogen at the primary infection site and leads to cell death following a SAR that increases SA accumulation and expression of *PATHOGENESIS-RELATED* (*PR*) genes. During a virus infection, NBS-LRR genes can lead to a HR or to an extreme resistance against the virus infection. The effector recognition LRR domains are responsible for the

recognition of specific pathogens and are the most variable parts of the protein (Meyers et al., 2003; Marone et al., 2013). All this makes *AT2G14080* a really interesting candidate involved in plant defense against TuMV.

Dynamin-related proteins are self-assembling GTPase involved in fission and fusion of membranes. In particular, DRP3B functions in mitochondrion and peroxisome fission in combination with DRP3A. Treatment of plant leaves with a dynamin-specific inhibitor disrupts the delivery of VPg and CI to endocytic structures and suppresses TuMV replication and intercellular movement (Wu et al., 2018). These two genes that were undescribed before as related to TuMV infection have clear functions related to the plant virus response and are potential important players in plant resistance or susceptibility. An open question to be tackled in future experiments is whether these two genes are also involved in a common plant response to other viruses or they are potyvirus-specific.

Along with *AT2G14080* and *DRP3B* the rest of the significant genes used in the LOF study will be described here. *AT2G14170* corresponds to gene *ALDEHYDE DEHYDROGENASE 6B2* (*ALDH6B2*) that encodes for a methylmalonate-semialdehyde dehydrogenase $NAD(P)^+$-dependent enzyme that catalyzes oxidation of aldehydes. It was shown that ripe grape berries infected with *Grapevine leafroll-associated closterovirus 3* have a significant decrease in the expression of aldehyde dehydrogenase genes (Zhang et al., 2012). *AT2G15320* encoded protein also belongs to the leucine-rich repeat (LRR) family protein involved in plant defense (Marone et al., 2013). *AT2G19270* encodes for a mitotic checkpoint PRCC-carboxy-term protein that was found to be targeted by geminiviruses (Ascencio-Ibáñez et al., 2008) and adenoviruses (Turner et al., 2015) to influence host gene expression. *AT1G67160* was mapped for TuMV-AS and corresponds to the gene *F-BOX/DUF295 BRASSICACEAE-SPECIFIC 5* (*FDB5*) that encodes for a protein which operates as a positive regulator of the BR-mediated signaling pathway and Ub-dependent protein catabolic process. BRs have been proven to be involved in plant resistance to virus infections (Zhang et al., 2015; Deng et al., 2016) as well as Ub-dependent protein catabolic processes (Alcaide-Loridan and Jupin, 2012; Verchot, 2016). The role of the F-box domain has been previously described in the defense mechanism of the host where they are involved in the ubiquitination and degradation of proteins (Correa et al., 2013). *AT3G56560* mapped for TuMV-DV corresponds to gene *NAC DOMAIN CONTAINING PROTEIN 65* (*NAC065*) which encodes for a protein with DNA-binding

transcription factor activity. Proteins with a NAC domain are transcriptional regulators and are important as transcriptional reprogrammers that help with regulation of plant stress response (Nuruzzaman et al., 2013). They also act as positive or negative regulators of plant immunity (Yuan et al., 2019), where they promote virus accumulation, as already seen in geminiviruses (Selth et al., 2005) and TMV (Wang et al., 2009). They also play a role in plant resistance to *Turnip crinkle tombusvirus* where the NAC domain protein interacts with the viral coat protein inducing the hypersensitive response of the plant (Ren et al., 2000). *AT5G40450* corresponds to the *REGULATOR OF BULB BIOGENESIS 1* (*RBB1*) gene that encodes a member of a gene family involved in vacuolar biogenesis and organization (Han et al., 2015). Plants use vacuoles to fight off pathogens, the type of vacuole depends on the type of pathogen (bacteria, fungi or viruses). The vacuolar collapse system which causes rapid degradation of the cellular material is used to fight of viral infection and triggers hypersensitive cell death (Hara-Nishimura and Hatsugai, 2011). *AT5G45770* corresponds to the *RECEPTOR LIKE PROTEIN 55* (*RLP55*) gene whose product is involved in regulation of defense response (Wang et al., 2008). They are cell surface receptors that respond to external and internal stimuli and are involved in growth and development and pathogen defense. It was demonstrated that they respond to abiotic and biotic stresses as well as hormones (Lv et al., 2016).

In the replicated study with 118 accessions there were SNPs mapped within 33 genes, and five of these genes were shared with the previous GWAS. Those five genes were: *AT2G14110* Haloacid dehalogenase-like hydrolase (HAD) superfamily protein, *AT2G14080* disease resistance protein (TIR-NBS-LRR class) family, *AT2G07120* F-box associated ubiquitination effector family protein, *AT2G14170* methylmalonate-semialdehyde dehydrogenase and *AT2G16400* BEL1-like homeodomain 7. All the significant genes were mapped on chromosome 2 except for *AT3G02990*, a member of heat stress transcription factor (Hsf) family, on chromosome 3 for TuMV-DV which was not previously identified. All the newly identified genes have a role that was previously described as involved in infection.

3.3. Validation of significant GWAS hits

A validation of the role of a small subset of the identified candidate genes on TuMV infection was performed. To this end, nine LOF mutants were selected. Seven mutants had a significant negative effect on the rates of disease progression and development and

severity of symptoms for the adapted strain, while only two also had a significant effect in the naïve ancestral strain. This difference highlights the stronger dependence of TuMV-DV to host factors. Regarding the two LRR domain containing genes that are affecting both strains, the LOF of locus *AT2G14080* shows particularly struggling results. One should expect that LOF of a resistance-inducing gene would result in increased susceptibility of plants to infection. This is consistent with the response to TuMV-AS infection, with *at2g14080* plants showing enhanced symptoms. However, TuMV-DV infected *at2g14080* plants showed a reduction in the severity of symptoms, at odds with the expectation. Similarly striking, the LOF of gene *AT2G15320* resulted in attenuation of symptoms and milder diseases for both viral strains. Finally, only two LOFs had no significant effect on any of the strains, *rbb1* and *rlp55*. Actually, this asseveration strictly applies to a lack of net effect on the trait medians, though they significantly increased the variance among plants.

A closer GWA analysis of the locus *AT2G14080* was performed. The two most significant SNPs on chromosome 2, 5923326 and 5928864, were recovered in the multiple-trait analysis continuously when accounting for direct effects. This result means that both viral strains cause the same plant response. Minor allele in the SNP position 5923326 on chromosome 2 is the most represented in accessions showing necrosis, thus contributing the most to the necrotic phenotype. The importance of the SNP 5923326 was further confirmed through the multi-locus analysis where conditioning for this SNP removes all the significant peaks in the GWA analysis. The lack of patterns in the geographical positions of the necrotic accessions makes it hard to derive any further conclusions about selective sweeps in this region.

In this work the infection with an adapted and a naïve TuMV strain of a wide variety of arabidopsis accessions (1050) was characterized. This data allowed the mapping of loci involved in potyvirus infection, informing on the genetic nature of host response and pinpointing genes involved in the response to both viruses along with specific genes for each strain. The results of the present GWAS converge to some degree with previous transcriptomic descriptions of arabidopsis infection with the same TuMV strains (Corrêa et al., 2020). Combining both results, it was possible to highlight essential genes regulating the interplay between the two components of this pathosystem as well as genes that may drive virus adaption. With further GWAS analyses we confirmed the strong

association of the locus *AT2G14080* with the necrotic phenotype but there was no significant pattern in the geographical location of the necrotic accessions making conclusions about selective sweeps hard.

Chapter 3: Adaptation of TuMV to a specific niche reduces its genetic and environmental robustness

1. Introduction

RNA viruses are very successful parasites that infect hosts across all biological kingdoms. This evolutionary success results from their evolvability, which in turn depends on the combination of three factors, namely high mutation rates, short generation times, and very large population sizes. However, these properties also come with costs. First, high mutation rates impose an upper limit to the length of the genome that can be maintained without increasing mutational load, which results in highly streamlined and compacted genomes (Elena and Sanjuán, 2005; Belshaw et al., 2007). Second, most mutations have a deleterious fitness effect, with a large fraction of them being even lethal (reviewed in Sanjuán, 2010), thus jeopardizing the survival of viral populations. How do RNA viruses maintain their functionality under such scenario of strong genomic stress? In the last 15 years or so, several studies have experimentally shown that such mutational pressure favors mechanisms that promote mutational robustness in RNA viruses (*e.g.*, Montville et al., 2005; Codoñer et al., 2006; Sanjuán et al., 2007; Stern et al., 2014; Thyagarajan and Bloom, 2014; Visher et al., 2016). Broadly speaking, genetic robustness refers to the constancy of the phenotype in the face of heritable perturbations (genetic or epigenetic; de Visser et al. 2003). However, the evolutionary origin and maintenance of genetic robustness still remains an unsolved question (de Visser et al., 2003; Elena et al., 2006; Elena, 2012; Lauring et al., 2013). Any mutation increasing genetic robustness will hardly rise in frequency because they have no other phenotypic effect than buffering the effect of other mutations (de Visser et al., 2003). This means that: (1) they will increase in frequency only at very high deleterious mutation rates because genotypes without these robustness-conferring mutations will simply suffer stronger mutational loads. (2) They will slow down the rate of adaptation by buffering the effect of other linked beneficial mutations. In conclusion, at low deleterious mutation rates (which may not be the case of RNA viruses), genetic robustness will not be easily selected. In theory, genotypes that produce more neutral mutations (*i.e.*, they inhabit in neutral network within the genotypic

landscape) could be directly selected (Wilke, 2001; Wilke et al., 2001; Codoñer et al., 2006). However, plenty of mutation-accumulation studies done with different RNA viruses suggest that the fraction of neutral mutations should be relatively small compared with those having deleterious effects (Sanjuán, 2010). Mutation accumulation in small populations may also select for genetic robustness (Krakauer and Plotkin, 2002; Forster et al., 2006; Elena et al., 2007), though a low population size would also reduce the effectiveness of selection (Forster et al., 2006; Elena et al. ,2007).

How to escape from this conundrum? In this context is where Ancel and Fontana (2000) postulated the plastogenetic congruence theory. Rapid environmental fluctuations and environmental unpredictability are quite common selective pressures and, therefore, any mutation conferring environmental robustness will necessarily be efficiently selected. Taken in a broad sense, environmental robustness refers to any kind of buffering against non-heritable perturbations (including both external stresses and developmental noise caused by fluctuations in the concentration of morphogens; de Visser et al., 2003). The plastogenetic congruence theory postulates that genetic robustness will arise as a correlated trait of strong selection for environmental robustness.

Viruses face strongly unpredictable environments during their life cycles: heterogeneity in susceptible host species, differences in cell types and even in the physiological stages of susceptible cells within a host species, the presence of antiviral immune and pharmacological responses, and other environmental factor, being temperature a well-known driver of virus adaptation (González et al., 2020). Experimental support to the plastogenetic congruence hypothesis in viruses was provided by Domingo-Calap et al., (2010), who evolved populations of bacteriophage Qβ under periodic temperature pulses to select for thermotolerant viruses (*i.e.*, environmentally robust) that in a series of subsequent experiments were shown to be also more genetically robust than control viruses.

In this study, we tested the plastogenetic congruence hypothesis using TuMV in its natural host arabidopsis. Specifically, we have used the aforementioned TuMV-AS and TuMV-DV, naïve and well-adapted to arabidopsis, respectively. In our study, we have evaluated the mutational and environmental robustness (thermal stability) of both strains. We found that TuMV-DV was very fragile to the accumulation of random mutations and showed

very little thermostability. In contrast, TuMV-AS was more robust both mutationally and environmentally. We discuss these results in the context of the plastogenetic congruence hypothesis and also in the context of how adaptation to one environment limits evolvability in alternative ones.

2. Results

The three variables measured, *AUDPS*, infectivity, and ST_{50}, were strongly correlated, as indicated by partial correlation analyses controlling for the viral isolate: *AUDPS* and infectivity were positively correlated (r_p = 0.9444, 7 df, $P = 0.0001$), *AUDPS* and ST_{50} were negatively correlated ($r_p = -0.9965$, 7 df, $P < 0.0001$) and infectivity and ST_{50} were negatively correlated too ($r_p = -0.9478$, 7 df, $P = 0.0001$). Fast appearance of symptoms (smaller ST_{50}) and a large number of infected plants (larger infectivity) are thus reflected in larger *AUDPS* values, thus confirming *AUDPS* provides a good proxy to the degree of adaptation of a particular viral genotype to its host and environmental conditions. Therefore, for simplicity, in the following sections, we will only report the results for the analyses done with *AUDPS*.

2.1. Adaptation of TuMV to arabidopsis and standard thermal conditions results in a reduction in genetic robustness

First, we evaluated the degree of adaptation to accession Col-0 in standard growing conditions of both viruses. Figure C3.1A shows the disease progression curves for the naïve TuMV-AS (solid black symbols and lines) and the arabidopsis-adapted TuMV-DV (solid red symbols and lines) viruses. Very significant differences exist between both viruses in the disease progression (χ^2 = 11.9775, 1 df, $P = 0.0005$). Consistently, the median *AUDPS* for TuMV-AS was 1.1667 ±0.0463 (±95% CI), while it was 7.3333 ±0.0785 for TuMV-DV (Fig. C3.1B, green distributions; *i.e.*, 6.29-fold better adapted).

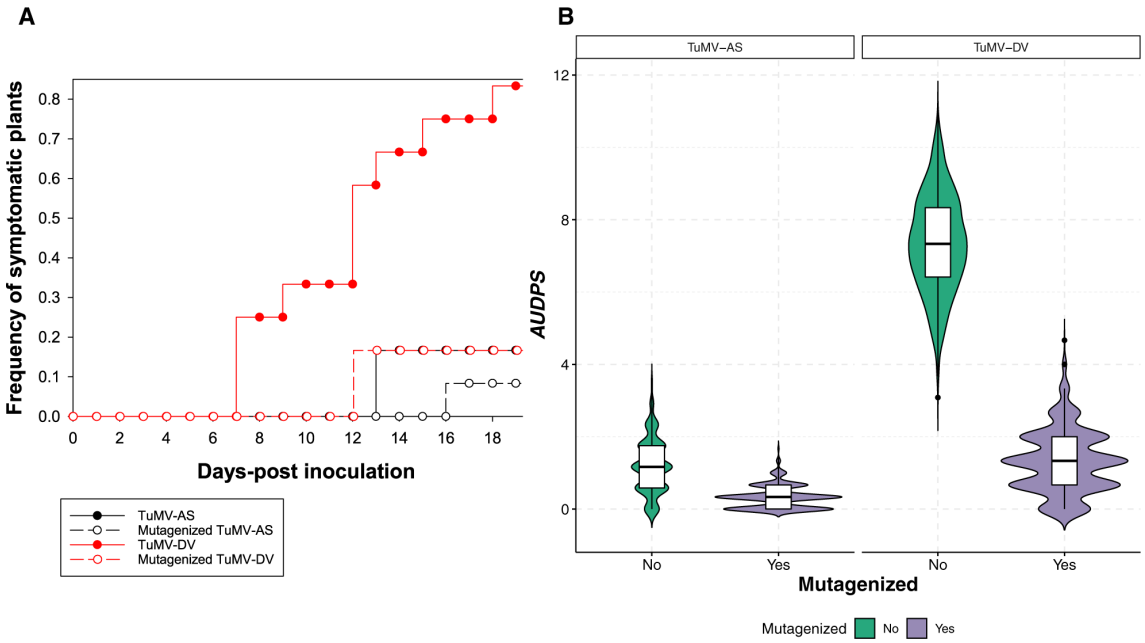


Figure C3.1. Evaluation of genetic robustness for TuMV-AS and TuMV-DV. (A) Disease progression curves for viruses submitted to N_2O-induced mutagenesis (open symbols and dashed lines) and their corresponding non-mutagenized controls (solid symbols and lines). (B) Estimates of *AUDPS* for each experimental condition. The kernel distributions estimated using the bootstrap algorithm are over imposed to the box diagrams.

After confirming the higher degree of adaptation of TuMV-DV to arabidopsis in the standard growing conditions, we sought to evaluate the degree of genetic robustness of each one virus. Fig. C3.1A shows the disease progression curves for the mutagenized viruses (open black symbols and dashed lines for TuMV-AS and open red symbols and dashed lines for TuMV-DV). Here we have compared mutagenized and non-mutagenized viruses. In the case of the non-adapted TuMV-AS isolate, the N_2O mutagenic treatment had no significant effect in the disease progression curve ($\chi^2 = 0.4097$, 1 df, $P = 0.5221$). The estimated median *AUDPS* for the mutagenized TuMV-AS was 0.3333 ±0.0196 (Fig. C3.1B, purple distributions). In sharp contrast, in the case of the arabidopsis-adapted TuMV-DV isolate, random mutagenesis had a strong negative effect on the progression curves ($\chi^2 = 10.9902$, 1 df, $P = 0.0009$), with the median *AUDPS* estimated for the mutagenized TuMV-DV being 1.3333 ±0.0529 (Fig. C3.1B, purple distributions), which means a reduction of 81.82% in disease progression efficiency.

The conclusion from this first experiment is that adaptation to arabidopsis was concomitant with a decrease in genetic robustness. This observation is consistent with the notion of TuMV-DV inhabits a high but narrow fitness peak while TuMV-AS occupied a flatter and more neutral region of the fitness landscape.

2.2. TuMV-AS and TuMV-DV differ in environmental robustness

Next, we sought to evaluate the environmental robustness of both viral isolates. First, we found that no significant differences exist among the disease progression curves observed for TuMV-AS across the four thermal environments (Fig. C3.2a, black lines and symbols: $\chi^2 = 0.3779$, 1 df, $P = 0.5387$). Again, in sharp contrast with this result, highly significant differences have been observed for the TuMV-DV across the four thermal environments (Fig. C3.2A, red lines and symbols: $\chi^2 = 8.7213$, 1 df, $P = 0.0031$).

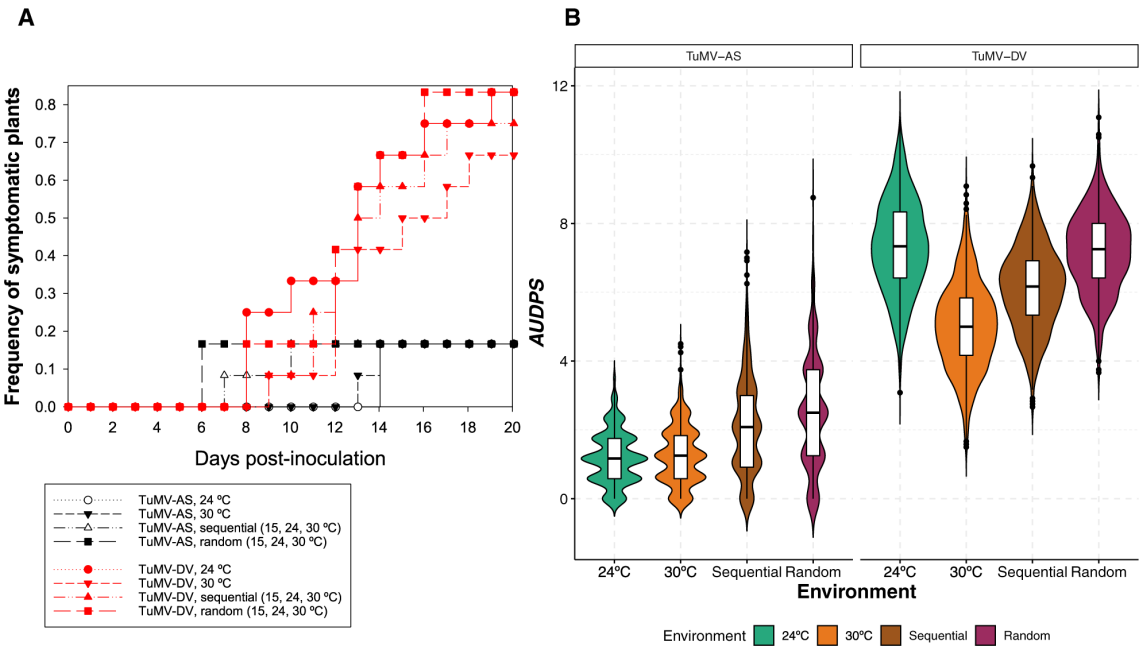


Figure C3.2. Evaluation of environmental robustness for TuMV-AS and TuMV-DV. (A) Disease progression curves for viruses growing under the four different thermal regimes (black symbols and lines for TuMV-AS and red symbols and lines for TuMV-DV). (B) Estimates of *AUDPS* for each experimental condition. The kernel distributions estimated using the bootstrap algorithm are over imposed to the box diagrams.

Interestingly, the variance component explained by differences among the four thermal environments was $\sigma_E^2 = 0.7636 \pm 0.0171$ (±1 SEM; maximum likelihood estimator of variance components in a one-way ANOVA) for TuMV-AS and $\sigma_E^2 = 1.3132 \pm 0.0294$ for

TuMV-DV (Fig. C3.2B); that is 71.97% more variance among thermal environments in the latter.

These two results together suggest that TuMV-AS generates more consistent disease progression curves across the four thermal environments than the arabidopsis-adapted TuMV-DV, which shows more variable responses across thermal environments. In other words, TuMV-AS is more environmentally robust (lower σ_E^2) than TuMV-DV.

3. Discussion

3.1. The tradeoff between robustness and evolvability in RNA viruses

The robustness of biological systems has several important implications. At the one side, it directly affects the probability of survival of organisms in the face of endogenous (*i.e.*, genetic and epigenetic mutations) and exogenous (*i.e.*, environmental uncertainties or developmental noise) perturbations (de Visser et al., 2003; Wagner, 2005; Bloom et al., 2006; Ciliberti et al., 2007; Wagner, 2008a), thus being a beneficial fitness trait. At the other side, however, robustness and evolvability represent the two faces of the same coin; genetic robustness may slow down the rate of adaptation by masking the effect of beneficial mutations as much as it buffers the effect of deleterious ones. Evidences showing this negative association between genetic robustness and evolvability have been somehow contradictory. Experimental results with digital organisms (Elena and Sanjuán, 2008) and vesicular stomatitis Indiana virus (VSV; Cuevas et al., 2009) have shown a negative association between short-term adaptability and genetic robustness. In contrast, experiments with bacteriophages have shown the opposite trend: genetic robustness promotes the evolution of thermal stability (McBride et al., 2008). Aligning with the bacteriophage results, Turner et al. (2010) have shown that environmentally robust (*i.e.*, generalists) populations of VSV were also more evolvable than highly specialized populations.

How to reconciliate all these apparently contradictory results? First, it has been suggested that genetic robustness can facilitate or jeopardize adaptation depending on population size, mutation rate, and the topography of the underlying fitness landscape (Krakauer and

Plotkin, 2002; Draghi et al., 2010). Second, the relationship between robustness and evolvability may be time-dependent. At the short-term genetic robustness will buffer the effect of potentially beneficial mutations, thus hampering adaptation. However, at the long-term genetic robustness will bolster evolvability by allowing populations to drift within neutral networks until reaching distant parts and switching to different neutral networks (Elena and Sanjuán, 2008; Wagner, 2008b). The epochal evolution of influenza A virus H3N2, alternating periods of phenotypic stasis punctuated by sudden changes in antigenic phenotypes (Koelle et al., 2006) fits well within this model of time-dependent effects of robustness: at the onset of an epochal evolution cycle, a H3N2 population is distributed over the neutral network of an antigenic cluster. Neutral mutations accumulate, allowing the virus to explore distant regions of the network. Later on, genotypes reach the edge of the network and create individuals that belong to a new antigenic cluster (Koelle et al., 2006; van Nimwegen, 2006).

3.2. The evolutionary origin of genetic robustness in RNA viruses

Still, the question of how genetic robustness evolves needs to be answered. An interesting proposal brought forward by Ancel and Fontana (2000) was the so-called plastogenetic congruence hypothesis. Under this hypothesis, genetic robustness evolves as a consequence of strong selection for mechanisms reducing the impact of environmental perturbations, that is, environmental robustness. Environmental perturbations along the life cycle of viruses occur constantly, thus imposing a strong selective advantage to any mechanism that may buffer them. How much evidence exists supporting the plastogenetic congruence hypothesis in the case of viruses? Domingo-Calap et al., (2010) directly tested the hypothesis by evolving bacteriophage Qβ under fluctuating temperatures to select for thermotolerant viruses. Then, these viruses were submitted to accumulation of random mutations in the same way we have used in this study. Their results provided support to the hypothesis, as the more thermotolerant viruses were also more robust against the deleterious effect of accumulated mutations. Here, we have also found an association between genetic and environmental robustness for two TuMV strains that differed in their degree of adaption to arabidopsis: the ancestral TuMV-AS shows more environmental robustness than its arabidopsis-adapted descendant TuMV-DV, echoing the observed differences in mutational robustness. Together these studies provide evidences supporting the link between genetic and environmental robustness, though a mechanistic explanation for such link is still missing.

3.3. Virus specialization limits evolvability

Here, we have observed that TuMV adaptation to a particular arabidopsis genotype (Col-0) and temperature conditions may be hampering its capacity to quickly respond to future changes in temperature. This observation mirrors the results of Turner et al. (2010) mentioned in Section 3.1, in which specialist populations of VSV where less evolvable to new cell types than generalist populations. A similar observation was done by Buckling et al., (2003) when exploring the evolvability of *Pseudomonas fluorescens* into different ecological niches. These authors argued that by climbing an adaptive peak, a population reduces standing genetic variability that would be beneficial in alternative environments, thus specializing into this particular niche. In contrast, a generalist population would exist outside of any particular fitness peak, gaining access to all of them (Buckling et al., 2003; Elena and Sanjuán, 2003). In this sense, by specializing to complete its infection cycle at 24 °C day temperature, TuMV-DV has limited its own evolvability.

In conclusion, we have shown results suggesting an association between environmental and genetic robustness in a natural pathosystem constituted by a plant virus and its natural multicellular eukaryotic host. This observation represents one small step forward in our understanding of the evolution of genetic robustness and adds generality to previous in vitro studies with bacteriophages and VSV. However, we still need to dig into the molecular and physiological mechanisms of such association between genetic and environmental robustness and the degree of adaptation to the host and growth conditions. An intriguing question is how much of the observed pattern is due to genomic changes in the virus versus the virus taking advantage from the host responses to thermal stress. For instance, it is well known that viruses take advantage of heat shock proteins (Elena et al., 2006; Geller et al. 2007) from the host, and thus an overexpression of such proteins by plant cells upon thermal stress may indirectly benefit the virus replication. This and similar questions will be explored in future works.

Final conclusions

Viruses that are naïve or preadapted to the host or have different host niches cause distinct symptomatology in the host and are more or less virulent. The exact host genes involved in this different response to infection are not yet known. Here we try to answer this question with the help of GWAS. Even though, GWAS in arabidopsis is not without controversy due to the high linkage disequilibrium and the polygenic nature of the genome, the use of suitable statistical methods that address these problems and a well described pathosystem made these studies valuable for the scientific community. Another important question is how the environment affects the viruses with different adaptation histories and their ability to maintain their phenotype. We tried answering that by looking at how temperature changes and mutagens affect the robustness of the naïve and preadapted virus. Here are the conclusions:

1. The generalist virus associated with more general host defense response genes compared to the specialist virus. Genes specific to infection with the generalist virus were ubiquitin protease and telomere repeat-binding protein 2. While genes specific to the specialist virus were the Nudix hydrolase, NADH dehydrogenase and DNAJ heat shock proteins. This highlights how the evolutionary history and host range of the viral strain dictates the way it interacts with the host. The genes specific to each virus strain could be targets of differential selection for evolution of specialist or generalist viruses. One gene was mapped for both viruses, *AT2G14080* a disease resistance TIR-NBS-LRR class protein, and it is a potential resistance factor.

2. Six host genes were shared between the naïve and preadapted strain (TIR-NBS-LRR and DRP3B proteins, a second LRR protein, *ALDH6B2*, *SNX4*, and a mitotic checkpoint protein) while many genes were strain-specific. These differentially expressed genes are potential drivers/targets of viral adaptation. The number of strain-specific genes found for the preadapted virus was significantly larger (47) than for the naïve virus (25). Two genes have been characterized as related to TuMV infection cycle for the first time: the TIR-NBS-LRR family disease resistance protein and gene *DRP3B*.

3. The highly adapted strain is strongly sensitive to the effect of random mutations and to changes in temperature conditions, while the naïve strain shows more robustness against both the accumulation of random mutations and drastic changes in temperature conditions. These results are consistent with the predictions of the plastogenetic congruence theory, suggesting that genetic and environmental robustness may be two sides of the same coin for TuMV.

Resumen

Introducción

Los virus de RNA tienen altas tasas de mutación, grandes tamaños de población y tiempos de generación cortos, lo que les otorga un gran potencial evolutivo y los hace responsables de muchas enfermedades emergentes nuevas. Los virus se enfrentan a especies de huéspedes con diferente respuesta a la infección o, en muchos casos, a distintos individuos dentro de una misma especie hospedadora que también pueden diferir en su respuesta a la infección. Cuando un virus entra en un huésped, necesita evadir las defensas del este el tiempo suficiente para replicarse y establecer una infección exitosa. Las plantas generan diferentes respuestas contra la infección viral. Estas respuestas son la resistencia gen a gen, las respuestas de lesiones locales necróticas hipersensibles, la reprogramación activa de la expresión génica, las explosiones oxidativas y el silenciamiento del RNA. Todas estas respuestas están orquestadas por la homeostasis hormonal, en la que varias hormonas activan genes de resistencia de la planta. Sin embargo, los virus tienen sus propios mecanismos para superar las defensas de las plantas. Un ejemplo es la utilización proteínas supresoras, las cuales interfieren en diferentes niveles de la vía de silenciamiento del RNA. Muchos de los genes del huésped relacionados con la infección son de interés ya que pueden ser responsables de una reducción de la susceptibilidad de las plantas o de una mayor tolerancia a la infección. Estos genes podrían tener posibles papeles como dianas de la adaptación del virus a la planta, determinando si la interacción virus-planta será más o menos virulenta.

Algunos virus se adaptan a una especie o genotipo particular del huésped en el que completan de manera eficiente su ciclo reproductivo. Estos virus se denominan especialistas. Los virus especializados suponen una gran amenaza para los monocultivos, ya que los virus bien adaptados generalmente muestran tasas de replicación mejoradas dentro del hospedador que a menudo se asocian con síntomas más fuertes. Cada estrategia de la gama de huéspedes tiene ventajas y desventajas. Al especializarse en un solo huésped, un virus puede limitar la competencia interespecífica y acceder mejor a los limitados recursos de su huésped. La ventaja del generalismo es la infección exitosa de múltiples huéspedes. Sin embargo, existe una limitación obvia para el generalismo: al ser capaz de infectar a múltiples huéspedes, un virus no maximiza la eficacia en ninguno en

particular, confirmando la hipótesis del *aprendiz de mucho, maestro de nada*. Se ha propuesto que la selección favorece a los virus especializados porque existe un compromiso que limita la eficacia de un virus generalista en cualquiera de los huéspedes alternativos y la evolución avanza más rápido en nichos más limitados. La pleiotropía antagonista consiste en que la que las adaptaciones beneficiosas a un huésped en particular podrían ser desventajosas en otro. Este es el mecanismo más comúnmente usado para explicar esta compensación. Para infectar múltiples huéspedes, los virus pueden necesitar codificar información genética adicional que ralentizaría su replicación y aumentaría su fragilidad mutacional. Además, las mutaciones que se fijan para compensar la pleiotropía antagonista limitan el acceso a rutas evolutivas alternativas a los máximos globales en el paisaje de eficacia, lo que reduce la capacidad de evolución. Todas estas características hacen que los especialistas sean capaces de una evolución y adaptación más rápida que los generalistas ante perturbaciones o nuevos entornos. Aunque los especialistas tienden a adaptarse más rápido a un solo huésped, los generalistas suelen superarlos en ambientes fluctuantes al estar más preparados para sobrevivir y reproducirse como consecuencia de tener una eficacia similar en los diferentes huéspedes. Esto permite que los virus generalistas tengan una mayor eficacia inicial en comparación con los virus especializados cuando infectan nuevas especies hospedadoras. Esto hace que sean patógenos emergentes y reemergentes con mayor probabilidad. De hecho, la evidencia generalizada apoya la teoría según la cual los linajes virales que se exponen secuencialmente a diferentes hospedadores durante largos períodos de tiempo maximizan su eficacia en todos los huéspedes en la misma medida que el especialista correspondiente, superando así los costos esperados del generalismo.

Un método que permite la identificación de genes importantes involucrados en la defensa o resistencia de las plantas es el estudio de asociación de genomas completos (GWAS en sus siglas inglesas). Estos estudios se basan en la hipótesis de que los alelos comunes que interactúan en múltiples loci que predisponen a enfermedades subyacen a las enfermedades más comunes. Esta hipótesis justifica el uso de GWAS para identificar los alelos, conectando fenotipos con genotipos. Esto nos permite identificar factores genéticos de riesgo de enfermedad, así como rasgos agronómicos importantes, tales como como susceptibilidad y resistencia a infecciones virales. *Arabidopsis thaliana* (L.) Heynh es uno de los organismos más útiles para el análisis de GWAS. Tiene más de 1000 accesiones naturales genotipadas y descritas, provenientes de Eurasia, América del Norte

y África del Norte (1001 Genomes Consortium, 2016). Los genotipos pueden mantenerse mediante autofecundación durante un número ilimitado de generaciones y, por lo tanto, hacen que los fenotipos sean altamente reproducibles, lo que facilita aún más el realizar GWAS. La disponibilidad de datos del genoma y la herramienta de análisis lo convierten en un organismo muy atractivo para estudiar varios rasgos diferentes. Por lo tanto, este estudio se propuso identificar genes de arabidopsis con un papel en la respuesta a infecciones virales y cómo estos genes pueden cambiar como consecuencia de la evolución del virus. En este trabajo se utiliza el patosistema formado por el virus del mosaico del nabo (TuMV; especie *Turnip mosaic potyvirus*, género *Potyvirus*, Familia *Potyviridae*) y arabidopsis. El TuMV es transmitido por más de 40 especies diferentes de pulgones y es capaz de infectar una amplia gama de plantas hospedantes, aunque en la mayoría pertenecen a la familia *Brassicaceae*. Estos virus causan diversos síntomas, que van desde moteado, clorosis, enanismo, mosaico, necrosis, esterilización y muerte de las plantas (Guerret et al., 2017). El TuMV no solo es uno de los virus más importantes que afectan a hortalizas y cultivos de importancia económica (Tomlinson, 1987), sino que también tiene una alta incidencia en las poblaciones silvestres de arabidopsis en la Península Ibérica (Pagán et al., 2010).

La identificación de los factores del huésped responsables de la resistencia o la permisividad a la infección es el objetivo final al estudiar las interacciones huésped-patógeno. Este conocimiento ayudará a un mejor manejo de las enfermedades. En este trabajo hemos caracterizado la infección de cepas generalistas y especializadas del TuMV en 450 accesiones naturales de arabidospsis. Las cepas virales utilizadas en este estudio se obtuvieron después de la evolución experimental de un aislado ancestral de TuMV. Este TuMV ancestral se evolucionó en genotipos de plantas deficientes en diferentes vías de señalización de enfermedades o en presencia de genes de susceptibilidad recesivos, lo que resultó en dos cepas particulares que diferían en gran medida en su rango de hospedadores experimentales. Además, se estudiaron otras dos cepas de TuMV: una primera *naïve* para arabidopsis y originalmente aislada de cala (*Zantedeschia sp*), TuMV-AS; y una segunda, TuMV-DV, adaptada a arabidopsis tras 12 pases de evolución experimental de TuMV-AS en plantas de la accesión Col-0. Ambas cepas difieren en gran medida en la intensidad de los síntomas que inducen, así como en la carga viral y en la magnitud de la perturbación inducida en el transcriptoma del hospedador y sus perfiles de metilación.

En resumen, la respuesta a la infección de accesiones naturales de arabidopsis de diferentes regiones geográficas fue fenotipada. Estas accesiones fueron inoculadas con cuatro cepas de TuMV que difieren en sus características e historia evolutiva. Los datos de infección se analizaron utilizando GWAS, buscando específicamente los SNPs asociados diferencialmente con la infección con las distintas cepas de TuMV. La arquitectura genética de los fenotipos relacionados con la infección también se estudió utilizando el modelo mixto lineal disperso Bayesiano (BSLMM). Aquí, nuestro objetivo es explorar si los virus con diferentes historias evolutivas o estrategias de rango de huéspedes afectan la fisiología de la planta y la progresión de la enfermedad de diferentes maneras, identificando genes del huésped candidatos que responden de manera diferente a cada virus. Finalmente, las asociaciones genéticas observadas se confirmaron mediante el estudio de la infección de plantas mutantes nulos de pérdida de función.

El tercer capítulo se centra en el estudio de la robustez en diferentes virus. La robustez genética se refiere a la constancia del fenotipo ante perturbaciones hereditarias (genéticas o epigenéticas). Sin embargo, el origen evolutivo y el mantenimiento de la robustez genética sigue siendo una cuestión sin resolver. Las altas tasas de mutación imponen un límite superior a la longitud del genoma que se puede mantener sin aumentar la carga mutacional, lo que da como resultado genomas altamente simplificados y compactados. En segundo lugar, la mayoría de las mutaciones ejercen un efecto negativo sobre la eficacia, siendo una gran fracción de ellas incluso letales. Esto pone en peligro la supervivencia de las poblaciones virales. Varios estudios han demostrado experimentalmente que dicha presión mutacional favorece los mecanismos que promueven la robustez mutacional en los virus de RNA. Cualquier mutación que aumente la robustez genética difícilmente aumentará en frecuencia porque no tiene otro efecto fenotípico que amortigüe el efecto de otras mutaciones. Esto significa que: (1) aumentarán en frecuencia solo a tasas de mutación deletéreas muy altas porque los genotipos sin estas mutaciones que confieren robustez simplemente sufrirán cargas mutacionales más fuertes. (2) Reducirán la velocidad de adaptación amortiguando el efecto de otras mutaciones beneficiosas vinculadas. En conclusión, a bajas tasas de mutaciones perjudiciales (que puede no ser el caso de los virus de RNA), la robustez genética no se seleccionará fácilmente. En teoría, los genotipos que producen mutaciones más neutrales (es decir, que habitan en una red neutra dentro del paisaje genotípico)

podrían seleccionarse directamente. Sin embargo, muchos estudios de acumulación de mutaciones realizados con diferentes virus de RNA sugieren que la fracción de mutaciones neutrales debería ser relativamente pequeña en comparación con las que tienen efectos deletéreos.

La acumulación de mutaciones en poblaciones pequeñas también puede seleccionar la robustez genética, aunque un tamaño de población bajo también reduciría la eficacia de la selección. Las fluctuaciones ambientales rápidas y la imprevisibilidad ambiental son presiones selectivas bastante comunes y, por lo tanto, cualquier mutación que confiera robustez ambiental será necesariamente seleccionada de manera eficiente. En un sentido amplio, la robustez ambiental se refiere a cualquier tipo de amortiguación contra perturbaciones no hereditarias (incluidas las alteraciones externas y el ruido del desarrollo causado por fluctuaciones en la concentración de morfógenos). La teoría de la congruencia plastogenética postula que la robustez genética surgirá como un rasgo correlacionado de una fuerte selección para la robustez ambiental. Los virus se enfrentan a entornos fuertemente impredecibles durante sus ciclos de vida: heterogeneidad en las especies hospedadoras susceptibles, diferencias en los tipos de células e incluso en las etapas fisiológicas de las células susceptibles dentro de una especie huésped, la presencia de respuestas inmunológicas y farmacológicas antivirales, y otros factores ambientales, como la temperatura, un conocido impulsor de la adaptación de virus. Las poblaciones evolucionadas de bacteriófago Qβ bajo pulsos de temperatura periódicos para seleccionar virus termotolerantes (es decir, ambientalmente robustos) demostraron ser también genéticamente más robustos que los virus de control. Estos resultados aportan apoyo experimental a la hipótesis de congruencia plastogenética en virus. En nuestro estudio probamos la hipótesis de la congruencia plastogenética utilizando TuMV en su hospedador natural *A. thaliana*. En concreto, las cepas TuMV-AS y TuMV-DV mencionadas con anterioridad. En nuestro estudio, hemos evaluado la robustez mutacional y ambiental (estabilidad térmica) de ambas cepas. Los resultados mostraron que TuMV-DV era muy frágil a la acumulación de mutaciones aleatorias y mostraba muy poca termoestabilidad. En contraste, TuMV-AS fue más robusto tanto a nivel mutacional como ambiental. Estos resultados pueden ser interpretados en el contexto de la hipótesis de la congruencia plastogenética, mostrando cómo la adaptación a un entorno limita la capacidad de evolución en otros alternativos.

Objetivos, metodología y resultados

El objetivo principal de esta tesis es caracterizar las diferencias entre las historias de adaptación viral y su interacción con el anfitrión o el medio ambiente:

1. Identificar diferentes genes de respuesta del huésped a un virus generalista o especializado.

2. Identificar genes específicos involucrados en la respuesta a cepas de virus adaptadas y no adaptadas.

3. Evaluar la robustez genética y ambiental de un virus de ARN con diferentes historias de adaptación.

Capítulo 1: Genes de arabidopsis que contribuyen a las diferencias en el resultado de la infección con cepas generalistas y especializadas del TuMV identificadas mediante estudios de asociación de todo el genoma

Los patógenos tendrán diferente virulencia e inducirán diferentes respuestas en sus hospedadores dependiendo de su historial de adaptación.

Los genes cartografiados en el GWAS pertenecían a categorías tales como: proteínas con cajas F, quinasas, hidrolasas, proteínas de la familia LRR, proteínas de resistencia a enfermedades, factores de transcripción, lectinas, helicasas, ubiquitina proteasas, proteínas involucradas en el metabolismo del hierro, péptidos con repeticiones de tipo pentatrico, GTPasas, y berberinas. Todas estas funciones están involucradas en la respuesta de la planta a la infección, el ciclo viral o el metabolismo del RNA. El locus *AT2G14080* se identificó como significativo para ambas cepas virales en el análisis de la gravedad de los síntomas. *AT2G14080* pertenece a los genes NBS-LRR, sus dominios LRR de reconocimiento efector reconocen patógenos específicos y pueden conducir a una respuesta inmune hipersensible (HR) o a una resistencia extrema frente a la infección

viral. También hubo algunos resultados específicos de cepas que se caracterizaron previamente como involucrados en la defensa de las plantas o en alguna parte importante del ciclo viral. De hecho, los genes que difieren entre los dos virus podrían ser objetivos de selección diferencial en la evolución de virus especializados o generalistas. Por ejemplo, la ubiquitina proteasa y la proteína 2 de unión a la repetición de telómeros fueron respuestas específicas de la planta a la infección por TuMV-G. Sin embargo, las proteínas de choque térmico Nudix hidrolasa, NADH deshidrogenasa y DNAJ fueron específicas para plantas infectadas con la cepa TuMV-S.

AT2G14080 tuvo un efecto significativo en el mutante inoculado con las dos cepas, por lo que parece estar involucrado en la defensa de la planta. Dos de los diez genes seleccionados para el análisis de mutantes provienen del análisis de TuMV-G y siete provienen del análisis de TuMV-S. Ocho de los genes seleccionados tuvieron un efecto significativo sobre la progresión y/o los síntomas de la enfermedad viral. *MATERNAL EFFECT EMBRYO ARREST 55* (*MEE55*, codifica para una proteína de biosíntesis de serina y esfingolípidos) y *AT1G57570* (codifica para un miembro de la proteína de la superfamilia de lectina de unión a manosa) aparentemente no tuvieron un efecto significativo en ninguna de las cepas virales en nuestras condiciones experimentales. Había cinco genes que tenían efecto en ambas cepas virales: *AT2G14080, AT4G02580, cpLEPA, NUDX5* y *NUDX6*. *AT2G14080* es un gen de resistencia de tipo NBS-LRR. Estas proteínas controlan el estado de las proteínas vegetales a las que se dirigen los patógenos y activan una serie de respuestas de defensa. Al eliminar este gen, los virus lograron inducir síntomas más fuertes.

Al observar el análisis de la arquitectura genética subyacente de cada rasgo fenotipado, fue evidente que algunos fenotipos relacionados con la enfermedad se explicaban por unos pocos SNP (la infectividad y la gravedad de los síntomas a 14 días post inoculación (dpi), para los virus y progresión de la enfermedad (*AUDPS*) y la infectividad a 21 dpi para ambos virus bien), mientras que algunos rasgos eran altamente poligénicos y se explicaban por una gran cantidad de SNP (*AUDPS* para TuMV-S a 14 dpi). Los SNPs que pasaron el umbral de probabilidad de inclusión posterior se mapearon dentro del locus *AT2G04440* (proteína de la familia MutT/Nudix) para *AUDPS* de TuMV-S a 21 dpi y la posición 6.685.977 en una región intergénica en el cromosoma 3 junto con *AT3G19350*,

que corresponde al gen MPC, para la infectividad de TuMV-G a 21 dpi (Fig. complementaria S2). Todos tenían posibles roles en la infección viral.

En conjunto, este trabajo (1) describe las diferencias entre un patógeno generalista y un especialista, (2) identifica y caracteriza genes involucrados en una infección viral de un generalista y un especializada y (3) ilustra la variabilidad de los elementos genéticos involucrados en una infección viral dependiendo de sobre la historia evolutiva de la cepa viral.

Capítulo 2: Genes de arabidopsis implicados en respuestas diferenciales a cepas adaptadas y no-adaptadas del TuMV

Se realizó GWAS utilizando 1050 accesiones con el fin de responder arabidopsis a una infección viral identificando nuevos genes de arabidopsis implicados en la respuesta a la infección viral.

La necrosis y la resistencia fueron los únicos rasgos para los que se han detectado SNPs de gran efecto. Los otros rasgos medidos fueron altamente poligénicos, con valores bajos de heredabilidad. Los SNPs significativos para ambos virus se localizaron dentro de diversos genes que estaban involucrados principalmente en funciones ya descritas en relación con el ciclo de infección viral. Sin embargo, dos genes se han caracterizado como relacionados con el ciclo de infección de TuMV por primera vez: la proteína de resistencia a enfermedades de la familia TIR-NBS-LRR, codificada por el locus *AT2G14080*, y el gen *DRP3B*. Las proteínas de clase TIR-NBS-LRR pueden provocar una respuesta HR o una resistencia extrema contra la infección por virus. Las proteínas relacionadas con la dinamina son GTPasa autoensambladas que participan en la fisión y fusión de membranas. Estos dos genes, que no se han descrito antes como relacionados con la infección por TuMV, tienen en nuestro estudio funciones claras relacionadas con la respuesta del virus de la planta. Estos genes tienen un rol potencial importante en la resistencia o susceptibilidad de las plantas.

Además, se validó el papel de un subconjunto de los genes candidatos identificados en la infección por TuMV. Para ello, se estudiaron un total de nueve mutantes nulos. Siete mutantes tuvieron un efecto negativo significativo sobre las tasas de progresión de la

enfermedad y el desarrollo y la gravedad de los síntomas para la cepa adaptada. Sin embargo, solo dos también tuvieron un efecto significativo en la cepa no-adaptada. Esta diferencia resalta la dependencia más fuerte de TuMV-DV a los factores del huésped. Con respecto a los dos genes que contienen el dominio LRR que afectan a ambas cepas, el KO del locus *AT2G14080* muestra resultados particularmente intricados. Cabría esperar que el mutante nulo de un gen inductor de resistencia resulte en una mayor susceptibilidad de las plantas a la infección. Esto es consistente con la respuesta a la infección por TuMV-AS, con plantas mutantes *at2g14080* que muestran síntomas más intensos. Sin embargo, las plantas *at2g14080* infectadas con TuMV-DV mostraron una reducción en la intensidad de los síntomas, en desacuerdo con las expectativas. De manera similar, el mutante nulo del gen *AT2G15320* dio como resultado la atenuación de los síntomas y enfermedades más leves para ambas cepas virales. Finalmente, solo dos mutantes nulos no tuvieron un efecto significativo sobre ninguna de las cepas: *rbb1* y *rlp55*.

El objetivo de nuestro estudio fue doble, primero identificar nuevos genes de arabidopsis que podrían estar involucrados en la interacción con su patógeno natural TuMV y, segundo, resaltar entre esos genes cuáles pueden responder de manera diferente a dos cepas del virus que difieren en su grado de adaptación al hospedador, por lo que son posibles impulsores de la evolución viral. Sólo seis de los genes cartografiados fueron compartidos entre las cepas TuMV-AS y TuMV-DV. Los seis genes comunes se han descrito previamente como implicados en las respuestas de las plantas a las infecciones virales (*ALDH6B2*, *SNX4*, *DRP3B*, TIR-NBS-LRR, otra proteína LRR, y una proteína de punto de control mitótico). El número de genes específicos de la cepa de TuMV encontrados fue significativamente mayor para TuMV-DV (47) que para TuMV-AS (25). Entre los loci que tenían un efecto pleiotrópico en más de un rasgo relacionado con la enfermedad y que eran específicos de TuMV-DV (es decir, aquellos cuya implicación en el ciclo de infección del virus se adquirió como consecuencia de la adaptación), cabe mencionar las metiltransferasas, helicasas, ubiquitina proteasas, proteínas que contienen dominio NAC y proteínas implicadas en el metabolismo del hierro.

En este trabajo se caracterizó la infección de una amplia variedad de accesiones de arabidopsis (1.050) con una cepa de TuMV adaptada y otra no adaptada. Estos datos permitieron la identificación de loci involucrados en la infección por potyvirus,

informando sobre la naturaleza genética de la respuesta del huésped y señalando los genes involucrados en la respuesta a ambos virus junto con genes específicos para cada cepa. Los resultados del presente GWAS convergen hasta cierto punto con descripciones transcriptómicas previas de la infección por arabidopsis con las mismas cepas de TuMV. Combinando ambos resultados, fue posible identificar genes esenciales que regulan la interacción entre los dos componentes de este patosistema, así como los genes que pueden impulsar la adaptación del virus.

La importancia del loci *AT2G14080* también se confirmó con varios análisis y se caracterizó la complejidad de la región. Debido a la falta de patrones en la distribución geográfica de las accesiones necróticas, fue difícil obtener más conclusiones sobre las señales de selección.

Capítulo 3: La adaptación del TuMV a un nicho específico reduce su robustez genética y ambiental

La robustez de los sistemas biológicos tiene varias implicaciones importantes. Por un lado, afecta directamente la probabilidad de supervivencia de los organismos frente a perturbaciones endógenas (es decir, mutaciones genéticas y epigenéticas) y exógenas (es decir, incertidumbres ambientales o ruido del desarrollo). Por lo tanto, es un rasgo beneficioso. Por otro lado, la robustez y la capacidad de evolución representan las dos caras de la misma moneda: la robustez genética puede ralentizar la tasa de adaptación enmascarando el efecto de las mutaciones beneficiosas tanto como amortigua el efecto de las deletéreas. En primer lugar, se ha sugerido que la robustez genética puede facilitar o poner en peligro la adaptación según el tamaño de la población, la tasa de mutación y la topografía del paisaje de eficacia subyacente. En segundo lugar, la relación entre la robustez y la capacidad de evolución puede depender del tiempo. A corto plazo, la robustez genética amortiguará el efecto de mutaciones potencialmente beneficiosas, lo que dificultará la adaptación. Sin embargo, a largo plazo, la robustez genética reforzará la capacidad de evolución al permitir que las poblaciones se desplacen dentro de redes neutrales hasta llegar a partes distantes y cambiar a diferentes redes neutrales.

Aún así, la pregunta de cómo evoluciona la robustez genética debe ser investigada. Una propuesta para responder la pregunta es la llamada hipótesis de congruencia

plastogenética. Bajo esta hipótesis, la robustez genética evoluciona como consecuencia de una fuerte selección de mecanismos que reducen el impacto de las perturbaciones ambientales, es decir, la robustez ambiental. Las perturbaciones ambientales a lo largo del ciclo de vida de los virus ocurren constantemente, lo que impone una fuerte ventaja selectiva a cualquier mecanismo que pueda amortiguarlas. Aquí, también hemos encontrado una asociación entre la robustez genética y ambiental para dos cepas de TuMV que diferían en su grado de adaptación a arabidopsis: TuMV-AS y TuMV-DV descritos anteriormente. Esta observación está en la misma dirección que las observadas para la robustez mutacional. Juntos, estos estudios proporcionan evidencias que apoyan el vínculo entre la robustez genética y ambiental, aunque aún falta una explicación mecanicista para tal vínculo. Hemos observado que la adaptación de TuMV al accesión Col-0 y las condiciones de temperatura pueden estar obstaculizando su capacidad para responder rápidamente a futuros cambios de temperatura. Al escalar un pico adaptativo, una población reduce la variabilidad genética de fondo que sería beneficiosa en entornos alternativos, especializándose así en este nicho particular. Por el contrario, una población generalista existiría fuera de cualquier pico de eficacia en particular, obteniendo acceso a todos ellos. En este sentido, al especializarse para completar su ciclo de infección a una temperatura diurna de 24 °C, TuMV-DV ha limitado su propia evolucionabilidad.

En conclusión, hemos mostrado resultados que sugieren una asociación entre la robustez ambiental y genética en un patosistema natural constituido por un virus de plantas y su huésped natural. Esta observación representa un paso adelante en nuestra comprensión de la evolución de la robustez genética y agrega generalidad a los estudios *in vitro* previos con bacteriófagos y el virus Indiana de la estomatitis vesicular. Por ejemplo, es bien sabido que los virus se aprovechan de las proteínas de choque térmico del huésped y, por tanto, una sobreexpresión de dichas proteínas por las células vegetales tras el estrés térmico puede beneficiar indirectamente la replicación del virus. Esta y otras cuestiones similares se explorarán en trabajos futuros.

Conclusiones

Los virus con distinto grado de adaptación a su huésped o que tienen diferentes gamas de huésped causan una sintomatología distinta en cada uno, siendo más o menos virulentos. Aún no se sabe bien cuáles son los genes del huésped exactamente involucrados en esta

respuesta diferente a la infección. En esta tesis se intenta responder a esta pregunta con la ayuda de GWAS. Otra cuestión importante es cómo afecta el medio ambiente a los virus con diferentes historias de adaptación y su capacidad para mantener su fenotipo. Intentamos responder a eso observando cómo los cambios de temperatura y los mutágenos afectan la robustez de virus con distinto grado de adaptación. Aquí están las conclusiones:

1. Se identificaron seis genes del hospedador compartidos entre las cepas no adaptada y adaptada (proteínas TIR-NBS-LRR, *DRP3B*, *ALDH6B2*, *SNX4*, una proteína de punto de control mitótico y una segunda proteína LRR) mientras que muchos genes eran específicos de cada cepa. Estos genes expresados diferencialmente son impulsores u objetivos potenciales de la adaptación viral. El número de genes específicos de la cepa encontrados para el virus adaptado fue significativamente mayor (47) que para el virus no-adaptado (25). Se han caracterizado por primera vez dos genes relacionados con el ciclo de infección de TuMV: la proteína de resistencia a enfermedades de la familia TIR-NBS-LRR y la proteína DRP3B relacionada con la dinamina.

2. El virus generalista muestra una preferencia por genes de respuesta de defensa del huésped más generales en comparación con las dianas del virus especializado. Los genes específicos de la infección con el virus generalista fueron la ubiquitina proteasa y la proteína 2 de unión repetida de los telómeros. Los genes específicos del virus especializado fueron las proteínas de choque térmico hidrolasa Nudix, una deshidrogenasa de NADH y DNAJ. Esto resalta cómo la historia evolutiva y la gama de huéspedes de la cepa viral dicta la forma en que interactúa con el huésped. Los genes específicos de cada cepa de virus podrían ser objetivos de selección diferencial para la evolución de virus especializados o generalistas. Se identificó un gen para ambos virus, *AT2G14080*, una proteína de clase TIR-NBS-LRR de resistencia a enfermedades. Esta proteína es un potencial factor de resistencia.

3. La cepa bien adaptada a arabidopsis es muy sensible al efecto de mutaciones aleatorias y a los cambios en las condiciones de temperatura. Sin embargo, la cepa no adaptada muestra mayor robustez tanto frente a la acumulación de mutaciones aleatorias como a los cambios drásticos en las condiciones de temperatura. Estos resultados son consistentes

con las predicciones de la teoría de la congruencia plastogenética, lo que sugiere que la robustez genética y ambiental pueden ser dos caras de la misma moneda para TuMV.

References

1001 Genomes Consortium. (2016) 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. Cell 166, 481–491.

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.

Alcaide-Loridan, C., Jupin, I. (2012) Ubiquitin and plant viruses, let's play together! Plant Physiol. 160, 72–82.

Alonso-Blanco, C., et al. (2016) 1,135 Genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. Cell 166, 481–491.

Ancel, L. and Fontana, W. (2000) Plasticity, evolvability, and modularity in RNA. J. Exp. Zool. 288, 242–83.

Anderson, P.K., et al. (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol. Evol. 19, 535–544.

Ascencio-Ibáñez, J.T., et al. (2008) Global analysis of arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. Plant Physiol. 148, 436–454.

Baebler, Š., et al. (2014) Salicylic acid is an indispensable component of the *Ny-1* resistance-gene-mediated response against potato virus Y infection in potato. J. of Exp. Bot. 65, 1095–1109.

Baliji, S., Lacatus, G., Sunter, G. (2010) The Interaction between geminivirus pathogenicity proteins and adenosine kinase leads to increased expression of primary cytokinin-responsive genes. Virology 402, 238–247.

Bari, R., Jones, J. D.G. (2009) Role of plant hormones in plant defence responses. Plant Mol. Biol. 69, 473–488.

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

Bedhomme, S., Lafforgue, G., Elena, S. F. (2012) Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations. Mol. Biol. Evol. 29, 1481–1492.

Belshaw, R., Pybus, O. G., Rambaut, A. (2007) The Evolutionof genome compression and genomic novelty in RNA viruses. Genome Res. 17, 1496–504.

Benitez-Alfonso, Y., Faulkner, C., Ritzenthaler, C., Maule, A.J. (2010) Plasmodesmata: gateways to local and systemic virus infection. Mol. Plant Microbe Interact. 23, 1403–1412.

Bentley, K., Evans, D.J. (2018) Mechanisms and consequences of positive-strand RNA virus recombination. J. Gen. Virol. 99, 1345–1356.

Bloom, J. D. et al. (2006) Protein stability promotes evolvability. Proc. Natl. Acad. Sci. USA 103, 5869–5874.

Blüthgen, N., Menzel, F., Blüthgen, N. (2006) Measuring specialization in species interaction networks. BMC Ecol. 6, 9.

Bono, L. M., Draghi, J. A., Turner, P.E. (2020) Evolvability costs of niche expansion. Trends Genet. 36, 14–23.

Bordería, A.V., Stapleford, K.A., Vignuzzi, M. (2011) RNA virus population diversity: implications for inter-species transmission. Curr. Opin. Virol. 1, 643–648.

Boyes, D. C. et al. (2001) Growth stage-based phenotypic analysis of arabidopsis: a model for high throughput functional genomics in plants. Plant Cell 13, 1449–1510.

Buckling, A., Wills, M. A., Colegrave, N. (2003) Adaptation limits diversification of experimental bacterial populations. Science 302, 2107–2109.

Bush, W. S., Moore, J. H. (2012) Genome-wide association studies. PLoS Comput. Biol. 8: e1002822.

Calil, I. P., Fontes, E. P. B. (2016) Plant immunity against viruses: antiviral immune receptors in focus. Ann. Bot. 119, 711–723.

Campos, L., et al. (2014) Salicylic acid and gentisic acid induce RNA silencing-related genes and plant resistance to RNA pathogens. Plant Phys. Biochem. 77, 35–43.

Cantor, R. M., Lange, K., and Sinsheimer. J. S. (2010) Prioritizing GWAS results: a review of statistical methods and recommendations for their application. Am. J. Human Genet. 86, 6–22.

Carr, J.P., Lewsey, M.G., Palukaitis, P. (2010) Signaling in induced resistance. Adv. Virus Res. 76, 57–121.

Carrasco-Hernandez, R., Jácome, R., López Vidal, Y., Ponce de León, S. (2017) Are RNA viruses candidate agents for the next global pandemic? a review. ILAR Journal 58, 343–358.

Casteel, C.L., et al. (2015) Disruption of ethylene responses by turnip mosaic virus mediates suppression of plant defense against the green peach aphid vector. Plant Physiol. 169, 209–218.

Cervera, H., Lalić, J., Elena, S. F. (2016) Efficient escape from local optima in a highly rugged fitness landscape by evolving RNA virus populations. Proc. R. Soc. B: Biol. Sci. 283, 20160984.

Chang, H. X., et al. (2016) Genome-wide association and genomic prediction identifies associated loci and predicts the sensitivity of tobacco ringspot virus in soybean plant introductions. BMC Genomics 17, 153.

Chen, C. C., et al. (2003) Identification of turnip mosaic virus isolates causing yellow stripe and spot on calla lily. Plant Disease, 87, 901–905.

Chen, G., et al. (2015) Genome-wide association implicates candidate genes conferring resistance to maize rough dwarf disease in maize. PLoS ONE 10: e0142001.

Chen, L., et al. (2013) WRKY8 transcription factor functions in the TMV-cg defense response by mediating both abscisic acid and ethylene signaling in arabidopsis. Proc. Natl. Acad. Sci. USA 110, E1963–E1971.

Cheng, X., Wang, A. (2017) The *Potyvirus* silencing suppressor protein VPg mediates degradation of SGS3 via ubiquitination and autophagy pathways. J. Virol. 91, e01478-16.

Chenon, M., Camborde, L., Cheminant, S., Jupin, I. (2012) A viral deubiquitylating enzyme targets viral RNA-dependent RNA polymerase and affects viral infectivity: TYMV DUB targets viral RdRp. EMBO J. 31, 741–753.

Choudhary, D. K., Prakash, A., Johri, B. N. (2007) Induced systemic resistance (ISR) in plants: mechanism of action. Indian J. Microbiol. 47, 289–297.

Choudhury, S., et al. (2019) Genome wide association study reveals novel QTL for barley yellow dwarf virus resistance in wheat. BMC Genomics 20:,891.

Chung, B. N., et al. (2016) The effects of high temperature on infection by potato virus Y, potato virus A, and potato leafroll viru*s*. Plant Pathol. J. 32, 321–328.

Chung, B. Y. W., Miller, W. A., Atkins, J. F., Firth, A. E. (2008) An overlapping essential gene in the *Potyviridae*. Proc. Natl. Acad. Sci. USA 105, 5897–5902.

Ciliberti, S., Martin, O. C., Wagner, A. (2007) Innovation and robustness in complex regulatory gene networks. Proc. Natl. Acad. Sci. USA 104, 13591–13596.

Cleaveland, S., Haydon, D. T., Taylor, L. (2007) Overviews of pathogen emergence: which pathogens emerge, when and why? Curr. Top. Microbiol. Immunol. 315, 86–111.

Codoñer, F.M. et al. (2006) The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. PLoS Pathog. 2, e136.

Corrêa, R. L. et al. (2020) Viral fitness determines the magnitude of transcriptomic and epigenomic reprograming of defense responses in plants. Mol. Biol. Evol. 37, 1866–1881.

Corrêa, R. L., et al. (2013) The role of F-Box proteins during viral infection. Int. J. Mol.r Sci. 14, 4030–4049.

Cuevas, J. M., Moya, A., Sanjuán, R. (2009) A genetic background with low mutational robustness is associated with increased adaptability to a novel host in an RNA virus. J. Evol. Biol. 22, 2041–2048.

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

De Visser, J. A. G. M. et al. (2003) Evolution and detection of genetic robustness. Evolution 57, 1959–1972.

Deng, X. G., et al. (2016) Role of brassinosteroid signaling in modulating tobacco mosaic virus resistance in *Nicotiana benthamiana*. Sci. Rep. 6, 20579.

Dennehy, J. J. et al. (2013) Frequent coinfection reduces RNA virus population genetic diversity. J. Heredity 104, 704–712.

Dielen, A. S., et al. (2011) The 20S proteasome α5 subunit of *Arabidopsis thaliana* carries an RNase activity and interacts in planta with the lettuce mosaic potyvirus HcPro protein: α5 RNase activity and HcPro interaction. Mol. Plant Pathol. 12, 137–150.

Domingo-Calap, P., Pereira-Gomez, M., Sanjuán, R. (2010) Selection for thermostability can lead to the emergence of mutational robustness in an RNA virus: thermostability and mutational robustness. J. Evol. Biol. 23, 2453–2460.

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

Domingo, E., Holland, J. J. (1997) RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. 51, 151–178.

Draghi, J. A. et al. (2010) Mutational robustness can facilitate adaptation. Nature 463, 353–355.

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

Dufresne, P. J. et al. (2008) *Arabidopsis thaliana* class II poly(A)-binding proteins are required for efficient multiplication of turnip mosaic virus. J. Gen. Virol. 89, 2339–2348.

Elena, S. F., Sanjuán, R. (2008) The effect of genetic robustness on evolvability in digital organisms. BMC Evol. Biol. 8, 284.

Elena, S. F. (2012) RNA virus genetic robustness: possible causes and some consequences. Curr. Opin. Virol. 2, 525–530.

Elena, S. F., Sanjuán, R. (2003) Climb every mountain? Science 302, 2074–2075.

Elena, S. F., Sanjuán, R. (2005) Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J. Virol. 79, 11555–11558.

Elena, S. F. et al. (2006) Mechanisms of genetic robustness in RNA viruses. EMBO Rep. 7, 168–173.

Elena, S. F. et al. (2007) Effect of population size and mutation rate on the evolution of mutational robustness. Evolution 61: 666–674.

Elena, S. F., Agudelo-Romero, P., Lalić, J. (2009) The evolution of viruses in multi-host fitness landscapes. Open Virol. J. 3, 1–6.

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

Ertunc, F. (2020) Emerging plant viruses, in: Emerging and Reemerging Viral Pathogens. Elsevier, pp. 1041–1062.

Eu-ahsunthornwattana, J., et al. (2014) Comparison of methods to account for relatedness in genome-wide association studies with family-based data. PLoS Genet. 10, e1004445.

Fischer, U., Dröge-Laser, W. (2004) Overexpression of *NtERF5*, a new member of the tobacco ethylene response transcription factor family enhances resistance to tobacco mosaic virus. Mol. Plant Microbe Interact. 17, 1162–1171.

Flor, H. H. (1971) Current status of the gene-for-gene concept. Annu. Rev. Phytopathol. 9, 275–296.

Forster, R., Adami, C., Wilke, C. O. (2006) Selection for mutational robustness in finite populations. J. Theor. Biol. 243, 181–190.

Gan, X., et al. (2011) Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature 477, 419–423.

Ge, X., Xia, Y. (2008) The role of AtNUDT7, a Nudix hydrolase, in the plant defense response, Plant Signal. Behav. 3, 119–120.

Geller, R. et al. (2007) Evolutionary constraints on chaperone-mediated folding provide an antiviral approach refractory to development of drug resistance. Genes Develop. 21, 195–205.

Goh, L., Yap, V. B. (2009) Effects of normalization on quantitative traits in association test. BMC Bioinformatics 10, 415.

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 Evol. 5, vez024.

González, R., Butković, A., Elena, S. F. (2020) From foes to friends: viral infections expand the limits of host phenotypic plasticity. Adv. Virus Res. 106, 85–121.

González, R. et al. (2021) Plant virus evolution under strong drought conditions results in a transition from parasitism to mutualism. Proc. Natl. Acad. Sci. USA 118, e2020990118.

Graci, J. D., et al. (2012) Mutational robustness of an RNA virus influences sensitivity to lethal mutagenesis. J. Virol. 86, 2869–2873.

Grangeon, R., et al. (2012) Impact on the endoplasmic reticulum and Golgi apparatus of turnip mosaic virus infection. J. Virol. 86, 9255–9265.

Guerret, et al. (2017) Biological and molecular properties of a turnip mosaic virus (TuMV) strain that breaks TuMV resistances in *Brassica napus*. Plant Dis. 101, 674–683.

Guo, Q., et al. (2016) RNA silencing in plants: mechanisms, technologies and applications in horticultural crops. Curr. Genomics 17, 476–489.

Guo, W., et al. (2020) Berberine induces resistance against tobacco mosaic virus in tobacco. Pest Manag. Sci. 76, 1804–1813.

Han, S.W., Alonso J. M., Rojas-Pierce, M. (2015) Regulator of bulb biogenesis 1 (*RBB1*) is involved in vacuole bulb formation in arabidopsis. PLoS ONE 10, e0125621.

Hara-Nishimura, I., Hatsugai, N. (2011) The role of vacuole in plant cell death. Cell Death Differ. 18, 1298–1304.

Harrison, B. D. (1956) Studies on the effect of temperature on virus multiplication in inoculated leaves. Ann. Appl. Biol. 44, 215–226.

Herlihy, J. H., Long, T. A., McDowell, J. M. (2020) Iron homeostasis and plant immune responses: Recent insights and translational implications. J. Biol. Chem. 295, 13444–13457.

Hershko, A., Ciechanover, A. (1998) The ubiquitin system. Annu. Rev. Biochem. 67, 425–479.

Hillung, J. et al. (2016) The transcriptomics of an experimentally evolved plant-virus interaction. Sci. Rep. 6: 24901.

Hillung. J. et al. (2014) Experimental evolution of an emerging plant virus in host genotypes that differ in their susceptibility to infection. Evolution 68, 2467–2480.

Hily, J. M. et al. (2016) Environment and host genotype determine the outcome of a plant-virus interaction: from antagonism to mutualism. New Phytol. 209, 812–822.

Holmes, E. C. (2009) The evolutionary genetics of emerging viruses. Annu. Rev. Ecol. Evol. Syst. 40, 353–372.

Honjo, M. N., et al. (2020) Seasonality of interactions between a plant virus and its host during persistent infection in a natural environment. ISME J. 14, 506–518.

Huang, Y. P., et al. (2020) Dissecting the role of a plant-specific Rab5 small GTPase *NbRabF1* in bamboo mosaic virus infection. J. Exp. Bot. 71, 6932–6944.

Ivanov, K. I., Eskelin, K., Löhmus, A., Makinen, K. (2014) Molecular and cellular mechanisms underlying potyvirus infection. J. Gen. Virol. 95, 1415–1429.

Ji, D. L. et al. (2012) *CpLEPA* is critical for chloroplast protein synthesis under suboptimal conditions in *Arabidopsis thaliana*. PLoS ONE 7, e49746.

Jin, H., Li, S., Villegas, A. (2006) Down-regulation of the 26S proteasome subunit RPN9 inhibits viral systemic transport and alters plant vascular development. Plant Physiol. 142, 651–661.

Jin, Y., et al. (2007) HC-Pro protein of potato virus Y can interact with three arabidopsis 20S proteasome subunits in planta. J. Virol. 81, 12881–12888.

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

Kant, R., et al. (2019) Host alternative NADH: ubiquinone oxidoreductase serves as a susceptibility factor to promote pathogenesis of *Rhizoctonia solani* in plants. Phytopathology 109, 1741–1750.

Kasschau K. D., et al. (2003) P1/HC-Pro, a viral suppressor of RNA silencing, interferes with arabidopsis development and miRNA function. Dev. Cell 4, 205–217.

Kassanis, B. (1957) Effects of changing temperature on plant virus diseases. Adv. Virus Res. 4, 221–241.

Kassen, R. (2002) The experimental evolution of specialists, generalists, and the maintenance of diversity: Experimental evolution in variable environments. J. Evol. Biol. 15, 173–190.

Kim, S., et al. (2007) Recombination and linkage disequilibrium in *Arabidopsis thaliana*. Nat. Genet. 39, 1151–1155.

Koelle, K. et al. (2006) Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. Science 314, 1898–903.

Kone, N. et al. (2017) Influence of planting date on incidence and severity of viral disease on cucurbits under field condition. Ann. Agric. Sci. 62, 99–104.

Korte, A., Farlow, A. (2013) The advantages and limitations of trait analysis with GWAS: A review. Plant Meth. 9, 29.

Korte, A., et al. (2012) A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nat. Genet. 44, 1066–1071.

Krakauer, D. C., Plotkin, J. B. (2002) Redundancy, antiredundancy, and the robustness of genomes. Proc. Natl. Acad. Sci. USA 99, 1405–1409.

Kreuder Johnson, C., et al. (2015) Spillover and pandemic properties of zoonotic viruses with high host plasticity. Sci. Rep. 5, 14830.

Lacroix, C., et al. (2014) Non-random biodiversity loss underlies predictable increases in viral disease prevalence. J. R. Soc. Interface 11, 20130947.

Lalić, J., Cuevas, J. M., Elena, S. F. (2011) Effect of host species on the distribution of mutational fitness effects for an RNA virus. PLoS Genet. 11, e1002378.

Lauring, A. S., Frydman, J., Andino, R. (2013) The role of mutational robustness in RNA virus evolution. Nat. Rev. Microbiol.11, 327–336.

Lee, J. Y. (2008) Phosphorylation of movement proteins by the plasmodesmal-associated protein kinase. Plant Virology Protocols, Methods in Molecular Biology™. Humana Press, Totowa, NJ. 625–639.

Lefeuvre, P., et al. (2019) Evolution and ecology of plant viruses. Nat. Rev. Microbiol. 17, 632-644.

Li, D. et al. (2013) The unexpected roles of eukaryotic translation elongation factors in RNA virus replication and pathogenesis. Microbiol. Mol. Biol. Rev. 77, 253–266.

Li, Y., Xiong, R., Bernards, M., Wang, A. (2016) Recruitment of arabidopsis RNA helicase AtRH9 to the viral replication complex by viral replicase to promote turnip mosaic virus replication. Sci. Rep. 6, 30297.

Lippert, C. et al. (2014) LIMIX: Genetic analysis of multiple traits. bioRxiv, https://www.biorxiv.org/content/10.1101/003905v2.

Liu, J. Z., et al. (2011) Soybean homologs of *MPK4* negatively regulate defense responses and positively regulate growth and development. Plant Physiol. 157, 1363–1378.

Llamas-Saiz, A. L., et al. (1996) Structural analysis of a mutation in canine parvovirus which controls antigenicity and host range. Virology 225, 65–71.

Loebenstein, G. (2009) Local lesions and induced resistance. Adv. Virus Res. 75, 73–117.

Lozano-Durán, R. et al. (2011) Geminiviruses subvert ubiquitination by altering CSN-mediated derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in *Arabidopsis thaliana*. Plant Cell 23, 1014–1032.

Lv, Y., et al. (2016) New insights into receptor-like protein functions in arabidopsis. Plant Signal. Behav. 11, e1197469.

Makowski, D., Ben-Shachar, M., Lüdecke, D. (2019) bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. J. Open Source Softw. 4, 1541.

Mandadi, K .K., Scholthof, K. B. G. (2013) Plant immune responses against viruses: how does a virus cause disease? Plant Cell 25, 1489–1505.

Manna, S. (2015) An overview of pentatricopeptide repeat proteins and their applications. Biochimie 113, 93–99.

Manrubia, S., Lazaro, E. (2006) Viral evolution. Phys. Life Rev. 3, 65–92.

Marone, D., et al. (2013) Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. Int. J. Mol. Sci. 14, 7302–7326.

Mauch-Mani, B., Mauch, F. (2005) The role of abscisic acid in plant–pathogen interactions. Curr. Opin. Plant Biol. 8, 409–414.

McBride, R. C., Ogbunugafor, C. B., Turner, P. E. (2008) Robustness promotes evolvability of thermotolerance in an RNA virus. BMC Evol. Biol. 8, 231.

McHale, L., et al. (2006) Plant NBS-LRR proteins: Adaptable guards. Genome Biol. 7, 212.

McLeish, M. J., Fraile, A., García-Arenal, F. (2019) Evolution of plant–virus interactions: host range and virus emergence. Curr. Opin. Virol. 34, 50–55.

Meyers, B. C., et al. (2003) Genome-wide analysis of NBS-LRR–encoding genes in Arabidopsis. Plant Cell 15, 809–834.

Moffett, P. (2009) Mechanisms of recognition in dominant *R* gene mediated resistance. Adv. Virus Res. 75, 1–229.

Montes, N., Cobos, A., Gil-Valle, M., Caro, E., Pagán, I. (2021) *Arabidopsis thaliana* genes associated with cucumber mosaic virus virulence and their link to virus seed transmission. Microorganisms 9: 692.

Montville, R., et al. (2005) Evolution of mutational robustness in an RNA virus. PLoS Biol. 3, e381.

Nakashita, H., et al. (2003) Brassinosteroid functions in a broad range of disease resistance in tobacco and rice: steroid hormone-mediated disease resistance. Plant J. 33, 887–898.

Navarro, R. et al. (2020) Defects in plant immunity modulate the rates and patterns of RNA virus evolution. bioRxiv, https://www.biorxiv.org/content/10.1101/2020.10.13.337402v1.full.

Nuruzzaman, M., Sharoni, A.M., Kikuchi, S. (2013) Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Front. Microbiol. 4, 248.

Obrępalska-Stęplowska, A., et al. (2015) Effect of temperature on the pathogenesis, accumulation of viral and satellite RNAs and on plant proteome in peanut stunt virus and satellite RNA-infected plants. Front. Plant Sci. 6, 903.

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

Oka, K., Kobayashi, M., Mitsuhara, I., Seo, S. (2013) Jasmonic acid negatively regulates resistance to Tobacco mosaic virus in tobacco. Plant Cell Physiol. 54, 1999–2010.

Padmanabhan, M. S., et al. (2005) Interaction of the tobacco mosaic virus replicase protein with the Aux/IAA protein PAP1/IAA26 is associated with disease development. J. Virol. 79, 2549–2558.

Pagán, I. et al. (2010) *Arabidopsis thaliana* as a model for the study of plant–virus co-evolution. Phil. Trans. R. Soc. B: Biol. Sci. 365, 1983–1995.

Pagán, I., et al. (2012) Effect of biodiversity changes in disease risk: exploring disease emergence in a plant-virus system. PLoS Pathog. 8, e1002796.

Pieterse, C. M. J., et al. (2012) Hormonal modulation of plant immunity. Annu. Rev. Cell Dev. Biol. 28, 489–521.

Purcell, S., et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Human Genet. 81, 559–575.

Raja, P. et al. (2008) Viral genome methylation as an epigenetic defense against geminiviruses. J. Virol. 82, 8997–9007.

Rajamäki, M. L., Streng, J., Valkonen, J. P. T. (2014) Silencing suppressor protein VPg of a potyvirus interacts with the plant silencing-related protein SGS3. Mol. Plant Microbe Interact. 27, 1199–1210.

Ren, T., Qu, F., Morris, T. J. (2000) HRT gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. Plant Cell 12, 1917–1925.

Revers, F., García, J.A. (2015) Molecular biology of potyviruses. Adv. Virus Res. 92, 101–199.

Rodamilans, B., et al. (2018) An atypical RNA silencing suppression strategy provides a snapshot of the evolution of sweet potato-infecting potyviruses. Sci. Rep. 8, 15937.

Roossinck, M. J. (2010) Lifestyles of plant viruses. Phil. Trans. R. Soc. B: Biol. Sci. 365, 1899–1905.

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

Rubio, B. et al. (2019) Genome-wide association study reveals new loci involved in *Arabidopsis thaliana* and turnip mosaic virus (TuMV) interactions in the field. New Phytol. 221, 2026–2038.

Ruiz-Ferrer, V., Voinnet, O. (2009) Roles of plant small RNAs in biotic stress responses. Annu. Rev. Plant Biol. 60, 485–510.

Sahana, N., et al. (2012) Inhibition of the host proteasome facilitates papaya ringspot virus accumulation and proteosomal catalytic activity is modulated by viral factor HcPro. PLoS ONE 7, e52546.

Sánchez, F., et al. (2015) Viral strain-specific differential alterations in arabidopsis developmental patterns. Mol. Plant Microbe Interact. 28, 1304–1315.

Sanfaçon, H. (2015) Plant translation factors and virus resistance. Viruses 7, 3392–3419.

Sanjuán, R. (2010) Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. Phil. Trans. R. Soc. B: Biol. Sci. 365, 1975–82.

Sanjuán, R. et al. (2007) Selection for robustness in mutagenized RNA viruses. PLoS Genet. 3, e93.

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

Scholthof, K. B. G., et al. (2011) Top 10 plant viruses in molecular plant pathology: Top 10 plant viruses. Mol. Plant Pathol. 12, 938–954.

Segura, V., et al. (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet. 44, 825–830.

Selth, L. A., et al. (2005) A NAC domain protein interacts with tomato leaf curl virus replication accessory protein and enhances viral replication. Plant Cell 17, 311–325.

Seren, Ü. (2018) GWA-Portal: Genome-wide association studies made easy. In Ristova, D., and Barbez, E. (Eds.), Root Development. Pp. 303–319. New York: Springer.

Si, L., et al. (2018) Triterpenoids manipulate a broad range of virus-host fusion via wrapping the HR2 domain prevalent in viral envelopes. Sci. Adv. 4, eaau8408.

Simko, I., and Piepho, H. P. (2012) The area under the disease progress stairs: calculation, advantage, and application. Phytopathology 102, 381–389.

Soosaar, J. L. M., Burch-Smith, T. M., Dinesh-Kumar, S. P. (2005) Mechanisms of plant resistance to viruses. Nat. Rev. Microbiol. 3, 789-799.

Spoel, S. H., Dong, X. (2012) How do plants achieve immunity? Defence without specialized immune cells. Nat. Rev. Immunol. 12, 89–100.

Stern, A. et al. (2014) Costs and benefits of mutational robustness in RNA viruses. Cell Rep. 8, 1026–36.

Stobbe, A., and Roossinck, M. J. (2016) Plant virus diversity and evolution. In Wang, A., Zhou, X. (Eds.), Current Research Topics in Plant Virology. Pp. 197–215. Cham: Springer.

Suntio, T., Mäkinen, K. (2012) Abiotic stress responses promote potato virus A infection in *Nicotiana benthamiana*. Mol. Plant Pathol. 13, 775-784.

Tamisier, L., et al. (2020) Genome-wide association mapping of QTLs implied in potato virus Y population sizes in pepper: evidence for widespread resistance QTL pyramiding. Mol. Plant Pathol. 21, 3–16.

Takeuchi, Y., et al. (1991) Host range mutant of human immunodeficiency virus type 1: modification of cell tropism by a single point mutation at the neutralization epitope in the env gene. J. Virol. 65, 1710–1718.

Thyagarajan, B., and Bloom, J. (2014) The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. eLife 3, e03300.

Tomlinson, J. A. (1987) Epidemiology and control of virus diseases of vegetables. Ann. Appl. Biol. 110, 661–681.

Turner, P. E. et al. (2010) Role of evolved host breadth in the initial emergence of an RNA virus: predicting virus emergence. Evolution 64, 3273–3286.

Turner, P. E., Elena, S. F. (2000) Cost of host radiation in an RNA virus. Genetics 156, 1465–1470.

Turner, R. L., Groitl, P., Dobner, T., Ornelles, D. A. (2015) Adenovirus replaces mitotic checkpoint controls. J. Virol. 89, 5083–5096.

Valli, A. A., et al. (2018) The HCPro from the potyviridae family: an enviable multitasking helper component that every virus would like to have. Mol. Plant Pathol. 19, 744–763.

Van Nimwegen, E. (2006) Influenza escapes immunity along neutral networks. Science 314, 1884–903.

Van Der Biezen, E. A., Jones, J. D. G. (1998) Plant disease-resistance proteins and the gene-for-gene concept. Trends Biochem. Sci. 23, 454–456.

Van Munster, M. (2020) Impact of abiotic stresses on plant virus transmission by aphids. Viruses 12, 216.

Verchot, J. (2016) Plant virus infection and the ubiquitin proteasome machinery: arms race along the endoplasmic reticulum. Viruses 8, 314.

Vijayan, V., et al. (2017) Virulence evolution of a sterilizing plant virus: tuning multiplication and resource exploitation. Virus Evol. 3, vex033.

Visher, E. et al. (2016) The mutational robustness of influenza A virus. PLoS Pathog. 12, e1005856.

Visser, J. A. G. M., et al. (2003) Perspective: evolution and detection of genetic robustness. Evolution 57, 1959–1972.

Vlot, A. C., Dempsey, D. A., Klessig, D. F. (2009) Salicylic acid, a multifaceted hormone to combat disease. Annu. Rev. Phytopathol. 47, 177–206.

Wagner, A. (2005) Robustness, evolvability, and neutrality. FEBS Lett. 579, 1772–1778.

Wagner, A. (2008a) Robustness and evolvability: a paradox resolved. Proc. R. Soc. B: Biol. Sci. 275, 91–100.

Wagner, A. (2008b) Neutralism and selectionism: a network-based reconciliation. Nat. Rev. Genet. 9, 965–974.

Walsh, J. A., Jenner, C. E. (2006) Resistance to turnip mosaic virus in the *Brassicaceae*, in: Loebenstein, G., Carr, J.P. (Eds.), Natural Resistance Mechanisms of Plants to Viruses. Springer Netherlands, Dordrecht, pp. 415–430.

Wang, G., et al. (2008) A genome-wide functional investigation into the roles of receptor-like proteins in arabidopsis. Plant Physiol. 147, 503-517.

Wang, X., Goregaoker, S. P., Culver, J. N. (2009) Interaction of the tobacco mosaic virus replicase protein with a NAC domain transcription factor is associated with the suppression of systemic host defenses. J. Virol. 83, 9720–9730.

Whitlock, M. C. (1996) The Red Queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. Am. Nat. 148, S65–S77.

Whitlock, M. C., Phillips, P. C., Moore, F. B., Tonsor, S. J. (1995) Multiple fitness peaks and epistasis. Annu. Rev. Ecol. Syst. 26, 601–629.

Wieczorek, P., Obrępalska-Stęplowska, A. (2015) Suppress to survive - implication of plant viruses in PTGS. Plant Mol. Biol. Rep. 33, 335–346.

Wilke, C. O. (2001) Selection for fitness versus selection for robustness in RNA secondary structure folding. Evolution 55, 2412–2420.

Willemsen, A. et al. (2018) Going, going, gone: predicting the fate of genomic insertions in plant RNA viruses. Heredity 121, 499–509.

Wojtaszek, P. (1997) Oxidative burst: an early plant response to pathogen infection. Biochem. J. 322, 681–692.

Woo, J., et al. (2012) The response and recovery of the *Arabidopsis thaliana* transcriptome to phosphate starvation. BMC Plant Biol. 12, 62.

Woolhouse, M. E. J. (2001) Population biology of multihost pathogens. Science 292, 1109–1112.

Woolhouse, M. E. J., Gowtage-Sequeria, S. (2005) Host range and emerging and reemerging pathogens. Emerg. Infect. Dis. 11, 1842–1847.

Woolhouse, M. E. J. (2002) Population biology of emerging and re-emerging pathogens. Trends Microbiol. 10, s3–s7.

Wu, G., et al. (2018) Dynamin-like proteins of endocytosis in plants are coopted by potyviruses to enhance virus infection. J. Virol. 92, e01320-18.

Wu, X., Ye, J. (2020) Manipulation of jasmonate signaling by plant viruses and their insect vectors. Viruses 12, 148.

Xu, P., et al. (2008) Virus infection improves drought tolerance. New Phytol. 180, 911–921.

Yang, C., et al. (2007) Spatial analysis of *Arabidopsis thaliana* gene expression in response to turnip mosaic virus infection. Mol. Plant Microbe Interact. 20, 358–370.

Yang, H., et al. (2010) *BAK1* and *BKK1* in *Arabidopsis thaliana* confer reduced susceptibility to turnip crinkle virus. Eur. J. Plant. Pathol. 127, 149–156.

Yasaka, R. et al. (2017) The timescale of emergence and spread of turnip mosaic potyvirus. Sci. Rep. 7, 4240.

Yin, L. et al. (2020) rMVP: A Memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. bioRxiv, https://www.biorxiv.org/content/10.1101/2020.08.20.258491v1.

Yoshimura, K., Shigeoka, S. (2015) Versatile physiological functions of the Nudix hydrolase family in arabidopsis. Biosci. Biotech. Biochem. 79, 354–366.

Yuan, X., Wang, H., Cai, J., Li, D., Song, F. (2019) NAC transcription factors in plant immunity. Phytopathol. Res. 1, 3.

Zaitlen, N., Kraft, P. (2012) Heritability in the genome-wide association era. Human Genet. 131, 1655–1664.

Zappa, A., Amendola, A., Romanò, L., Zanetti, A. (2009) Emerging and re-emerging viruses in the era of globalisation. Blood Transfus. 7: 167-171.

Zhang, D. W., et al. (2015) Induction of plant virus defense response by brassinosteroids and brassinosteroid signaling in *Arabidopsis thaliana*. Planta 241 875–885.

Zhang, J., Zhou, J. M. (2010) Plant immunity triggered by microbial molecular signatures. Mol. Plant 3, 783–793.

Zhang, T. et al. (2019) Wheat yellow mosaic virus NIb interacting with host light induced protein (LIP) facilitates its infection through perturbing the abscisic acid pathway in wheat. Biology 8: 80.

Zhang, Y., et al. (2012) Genome-wide identification and analysis of grape aldehyde dehydrogenase (ALDH) gene superfamily. PLoS ONE 7, e32153.

Zhou, X., Stephens, M. (2012) Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44, 821–824.

Zhou, X., Carbonetto, P., Stephens, M. (2013) Polygenic modeling with Bayesian sparse linear mixed models. PLoS Genet. 9, e1003264.

Zhu, S., et al. (2005) The rice dwarf virus P2 protein interacts with ent -kaurene oxidases *in vivo*, leading to reduced biosynthesis of gibberellins and rice dwarf symptoms. Plant Physiol. 139, 1935–1945.

Zhu, S., et al. (2013) Double-stranded RNA-binding protein 4 is required for resistance signaling against viral and bacterial pathogens. Cell Rep. 4, 1168–1184.

Supplementary files

Supplementary Table S1. Selected accessions for Chapter 1 GWAS.

tg_ecotypeid	name	CS_number	country	latitude	longitude	collector	seq_by	block
430	Gr-1	CS76496	AUT	47	15.5	Albert Kranz	Salk	1
470	BRR4	CS78943	USA	408,313	-87,735	Diane Byers	MPI	1
476	BRR12	CS78944	USA	408,313	-87,735	Diane Byers	MPI	2
544	LI-WP-039	CS78949	USA	409,076	-732,089	Oliver Bossdorf	MPI	1
546	LI-WP-041	CS78950	USA	409,076	-732,089	Oliver Bossdorf	MPI	1
628	LI-OF-061	CS78951	USA	407,777	-729,069	Oliver Bossdorf	MPI	2
685	LI-EF-011	CS78954	USA	409,064	-731,493	Oliver Bossdorf	MPI	2
766	Dja-1	CS76473	KGZ	425,833	736,333	Olivier Loudet	Salk	1
772	Neo-6	CS76560	TJK	37.35	724,667	Olivier Loudet	Salk	1
801	KYC-33	CS76992	USA	379,169	-844,639	Kathleen Donohue	Monsanto	1
853	MIA-1	CS78958	USA	417,976	-866,691	Kathleen Donohue	MPI	2
854	MIA-5	CS78959	USA	417,976	-866,691	Kathleen Donohue	MPI	1
915	LIN S-5	CS77040	USA	418,972	-714,378	Kathleen Donohue	Monsanto	1
932	CHA-41	CS76765	USA	423,634	-711,445	Kathleen Donohue	Monsanto	2
997	Ale-Stenar-56-14	CS76653	SWE	553,833	14.05	Jon √Ögren	GMI	1
1006	Ale-Stenar-77-31	CS77636	SWE	553,833	14.05	Jon √Ögren	GMI	2
1061	Brvðsarp-11-135	CS76727	SWE	557,167	141,333	Jon √Ögren	GMI	1
1070	Brvðsarp-45-153	CS77643	SWE	557,167	141,333	Jon √Ögren	GMI	2
1158	Aledal-6-49	CS76656	SWE	56.7	165,167	Jon √Ögren	GMI	1
1166	Aledal-14-73	CS77651	SWE	56.7	165,167	Jon √Ögren	GMI	2
1254	Tos-82-387	CS77379	SWE	594,333	170,167	Jon √Ögren	GMI	1
1257	Tos-95-393	CS77380	SWE	594,333	170,167	Jon √Ögren	GMI	2
1622	Brn-24	CS78963	USA	41.9	-86,583	Diane Byers	MPI	2
1652	DuckLkSP40	CS78965	USA	433,431	-864,045	Diane Byers	MPI	2
1684	Haz-10	CS78967	USA	41,879	-86,607	Diane Byers	MPI	1
1756	Ker-4	CS78970	USA	42,184	-86,358	Diane Byers	MPI	2
1757	Ker-5	CS78971	USA	42,184	-86,358	Diane Byers	MPI	1
1797	L-R-10	CS78973	USA	41,847	-86.67	Diane Byers	MPI	1
1829	Mdn-1	CS77077	USA	42,051	-86,509	Diane Byers	Monsanto	2
1872	MNF-Pot-75	CS77100	USA	43,595	-862,657	Diane Byers	Monsanto	2
1925	MNF-Che-2	CS77096	USA	435,251	-861,843	Diane Byers	Monsanto	2
1942	MNF-Che-47	CS78980	USA	435,251	-861,843	Diane Byers	MPI	2

2057	Map-42	CS77732	USA	42,166	-86,412	Diane Byers	MPI	2
2091	MuskSP-83	CS78986	USA	432,483	-863,368	Diane Byers	MPI	1
2106	MSGA-10	CS78987	USA	432,749	-860,891	Diane Byers	MPI	1
2239	Riv-25	CS78994	USA	42,184	-86,382	Diane Byers	MPI	2
2286	SLSP-67	CS78997	USA	43,665	-86,496	Diane Byers	MPI	2
4807	UKSW06-207	CS78809	UK	50.4	-4.9	Eric Holub	Monsanto	2
4826	UKSW06-226	CS78810	UK	50.4	-4.9	Eric Holub	Monsanto	2
4939	UKSW06-341	CS79002	UK	50.4	-4.7	Eric Holub	MPI	2
4958	UKSW06-360	CS78814	UK	50.5	-4.5	Eric Holub	Monsanto	2
5023	UKSE06-118	CS78799	UK	51.3	0.5	Eric Holub	Monsanto	1
5104	UKSE06-252	CS78800	UK	51.3	0.5	Eric Holub	Monsanto	2
5165	UKSE06-362	CS78802	UK	51.3	0.4	Eric Holub	Monsanto	1
5395	UKNW06-102	CS79005	UK	54.4	-3	Eric Holub	MPI	2
5651	UKNW06-488	CS79006	UK	54.4	-2.9	Eric Holub	MPI	2
5720	Cal-2	CS78781	UK	53.3	-1.6	D. Ratcliffe	Monsanto	2
5741	For-2	CS78783	UK	56.6	-4.1	D. Ratcliffe	Monsanto	2
5748	Kil-0	CS78784	UK	56	-4.4	D. Ratcliffe	Monsanto	2
5776	UKID71	CS79009	UK	52.9	-1.3	Eric Holub	MPI	2
5779	UKID74	CS78789	UK	51	-3.1	Eric Holub	Monsanto	2
5784	Ty-1	CS78790	UK	56.4	-5.2	D. Ratcliffe	Monsanto	1
5798	UKID93	CS79010	UK	53.1	-3.3	Eric Holub	MPI	2
5837	Bor-1	CS76453	CZE	494,013	162,326	Jirina Relichov	Salk	2
5907	DraIV 2-9	CS76818	CZE	494,112	162,815	Jirina Relichov	Monsanto	2
5921	DraIV 3-7	CS76819	CZE	494,112	162,815	Jirina Relichov	Monsanto	2
5950	DraIV 5-12	CS76820	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5984	DraIV 6-13	CS76822	CZE	494,112	162,815	Jirina Relichov	Monsanto	2
6008	Duk	CS76824	CZE	49.1	16.2	Jirina Relichov	Salk	2
6012	Eden-7	CS76829	SWE	62,877	18,177	Magnus Nordborg	GMI	2
6016	Eds-1	CS76834	SWE	62.9	18.4	Magnus Nordborg	GMI	2
6021	FjV§2-4	CS76862	SWE	56.06	14.29	Magnus Nordborg	GMI	2
6023	Fly2-1	CS76863	SWE	557,509	133,712	Magnus Nordborg	GMI	2
6024	Fly2-2	CS76864	SWE	557,509	133,712	Magnus Nordborg	GMI	2
6036	Hov3-2	CS76934	SWE	56.1	13.74	Magnus Nordborg	GMI	2
6039	Hovdala-2	CS76937	SWE	56.1	13.74	Torbjorn Sall	GMI	1
6040	Kni-1	CS76970	SWE	55.66	13.4	Magnus Nordborg	GMI	2
6041	Lis-3	CS77044	SWE	560,328	14,775	Magnus Nordborg	GMI	2
6042	Lom1-1	CS77048	SWE	56.09	13.9	Magnus Nordborg	GMI	2
6070	Omn-1	CS77145	SWE	629,308	183,448	Magnus Nordborg	GMI	1
6071	Omn-5	CS77146	SWE	629,308	183,448	Magnus Nordborg	GMI	2

6077	Rev-3	CS78030	SWE	556,942	134,504	Magnus Nordborg	GMI	2
6086	Sr:3	CS77267	SWE	58.9	11.2	Ivo Cetl	GMI	2
6108	T480	CS77300	SWE	557,989	131,206	Mattias Jakobsson	GMI	2
6109	T510	CS77301	SWE	557,936	131,233	Mattias Jakobsson	GMI	2
6111	T530	CS77302	SWE	557,989	131,219	Mattias Jakobsson	GMI	2
6113	T550	CS77304	SWE	558,078	131,028	Mattias Jakobsson	GMI	2
6118	T610	CS77307	SWE	55.7	13.2	Mattias Jakobsson	GMI	2
6126	T720	CS77311	SWE	558,411	133,047	Mattias Jakobsson	GMI	2
6132	T790	CS77316	SWE	558,386	133,186	Mattias Jakobsson	GMI	2
6133	T800	CS77317	SWE	558,364	132,906	Mattias Jakobsson	GMI	2
6136	T840	CS77319	SWE	559,336	135,519	Mattias Jakobsson	GMI	2
6137	T850	CS77320	SWE	559,419	135,603	Mattias Jakobsson	GMI	2
6145	T930	CS77324	SWE	559,497	135,533	Mattias Jakobsson	GMI	2
6149	T970	CS77326	SWE	559,281	135,481	Mattias Jakobsson	GMI	2
6150	T980	CS77327	SWE	559,261	135,319	Mattias Jakobsson	GMI	2
6153	TAA 03	CS77329	SWE	626,425	177,422	Mattias Jakobsson	GMI	2
6163	TAA 14	CS77331	SWE	626,425	177,356	Mattias Jakobsson	GMI	2
6172	TvÖD 04	CS77335	SWE	628,717	183,436	Mattias Jakobsson	GMI	2
6195	TDr-9	CS77356	SWE	557,708	141,342	Mattias Jakobsson	GMI	2
6220	TGR 01	CS77365	SWE	62,806	181,896	Mattias Jakobsson	GMI	2
6235	TOM 01	CS77372	SWE	629,611	183,589	Mattias Jakobsson	GMI	2
6242	Tomegap-2	CS77377	SWE	55.7	13.2	Mattias Jakobsson	GMI	2
6749	FM-10	CS79013	USA	424,489	-765,072	Michael Nachman	MPI	1
6750	FM-11	CS79014	USA	424,489	-765,072	Michael Nachman	MPI	2
6805	HS-12	CS79015	USA	42,373	-710,627	Toby Kellogg	MPI	1
6830	Kz-13	CS76994	KAZ	49.5	73.1	Ihsan Al-Shehbaz	MPI	2
6920	Got-22	CS76884	GER	515,338	99,355	Gerhard RvÐbbelen	Salk	1
6929	Kondara	CS76532	TJK	38.48	68.49	Igor Vizir	Salk	2
6933	LL-0	CS77047	ESP	41.59	2.49	Albert Kranz	Salk	1
6938	Ms-0	CS76555	RUS	557,522	376,322	Albert Kranz	Salk	1
6944	NFA-8	CS78913	UK	514,083	-0.6383	Mick Crawley	Salk	2
6945	Nok-3	CS76562	NED	52.24	4.45	Albert Kranz	Salk	2
6974	Ull2-5	CS78818	SWE	560,648	139,707	Magnus Nordborg	GMI	1
6975	Uod-1	CS76621	AUT	48.3	14.45	Marcus Koch	Salk	2
6987	Ak-1	CS76431	GER	480,683	762,551	Albert Kranz	Salk	2

6997	Appt-1	CS76440	NED	518,333	55,833	Maarten Koornneef	Salk	1
7025	Bl-1	CS76450	ITA	445,041	113,396	Albert Kranz	Salk	2
7033	Buckhorn Pass	CS76733	USA	413,599	-122,755	Angus Murphy	Salk,MPI	2
7058	Bur-0	CS76734	IRL	53.08	-907,555,556	Albert Kranz	Mott	2
7062	Ca-0	CS76459	GER	502,981	826,607	Albert Kranz	Salk	2
7072	Chi-0	CS76464	RUS	537,502	347,361	Albert Kranz	Salk	2
7081	Co	CS78895	POR	402,077	-842,639	George Redei	Salk	1
7096	Di-G	CS76472	FRA	473,239	504,278	Maarten Koornneef	Salk	2
7102	Do-0	CS76474	GER	507,224	82,372	Albert Kranz	Salk	2
7109	Ema-1	CS76480	UK	51.3	0.5	Eric Holub	Salk	1
7111	Edi-0	CS76831	UK	559,494	-316,028	Albert Kranz	Mott	1
7117	El-0	CS76479	GER	515,105	968,253	Albert Kranz	Salk	2
7248	Mv-0	CS76556	USA	413,923	-706,652	Albert Kranz	Salk	2
7287	Ove-0	CS76569	GER	533,422	842,255	Albert Kranz	Salk	2
7306	Pog-0	CS76576	CAN	492,655	-123,206	Albert Kranz	Salk	2
7342	Su-0	CS76606	UK	536,473	-300,733	Albert Kranz	Salk	2
7343	Sp-0	CS76603	GER	525,339	13,181	Albert Kranz	Salk	2
7353	Tha-1	CS76611	NED	52.08	4.3	Maarten Koornneef	Salk	2
7354	Ting-1	CS76612	SWE	56.5	14.9	Maarten Koornneef	Salk	1
7382	Utrecht	CS76622	NED	520,918	51,145	Viola Willemsen	Salk	1
7384	Ven-1	CS76624	NED	520,333	5.55	Maarten Koornneef	Salk	1
7387	Vind-1	CS76625	UK	549,902	-23,671	Maarten Koornneef	Salk	2
7396	Ws-0.2	CS78857	RUS	52.3	30	Albert Kranz	Mott	2
7413	Wil-2	CS78856	LTU	546,833	253,167	Albert Kranz	Mott	1
7417	Zu-0	CS78880	SUI	473,667	8.55	Albert Kranz	Mott	2
7419	Db-1	CS76471	GER	503,058	832,213	Albert Kranz	Salk	2
7521	Lp2-6	CS77052	CZE	49.38	16.81	Ivo Cetl	Salk	2
7523	Pna-17	CS76575	USA	420,945	-863,253	Joy Bergelson	Salk	2
7529	627RMX-1MN4	CS79022	USA	420,333	-865,128	Justin Borevitz	MPI	2
7917	PNA3.10	CS77183	USA	420,945	-863,253	Megan Dunning	Monsanto	2
7947	PNA3.40	CS77184	USA	420,945	-863,253	Megan Dunning	Monsanto	1
8037	PT1.52	CS79027	USA	413,423	-867,368	Megan Dunning	MPI	1
8057	PT1.85	CS79028	USA	413,423	-867,368	Megan Dunning	MPI	2
8077	PT2.21	CS77191	USA	413,423	-867,368	Megan Dunning	Monsanto	2
8227	THVñ 03	CS77367	SWE	627,989	179,103	Mattias Jakobsson	GMI	2

8234	Gul1-2	CS76896	SWE	564,606	158,127	Magnus Nordborg	GMI,Salk	2
8235	Hod	CS76924	CZE	48.8	17.1	Jirina Relichov	Salk	1
8236	HSm	CS76941	CZE	49.33	15.76	Jirina Relichov	Salk	2
8237	Kv§vlinge-1	CS76964	SWE	55.8	13.1	Torbjorn Sall	GMI	2
8238	Kent	CS76967	UK	51.15	0.4	Magnus Nordborg	Salk	2
8241	Liarum	CS77038	SWE	559,473	13,821	Torbjorn Sall	GMI	2
8246	NC-6	CS77124	USA	35	-79.18	Joy Bergelson	Salk	2
8247	San-2	CS77233	SWE	56.07	13.74	Magnus Nordborg	GMI	2
8285	DraIII-1	CS76815	CZE	494,112	162,815	Jirina Relichov	Salk	1
8343	Na-1	CS76558	FRA	47.5	1.5	Albert Kranz	Salk	1
8351	Ost-0	CS77154	SWE	60.25	18.37	Albert Kranz	GMI	2
8366	Rd-0	CS76584	GER	50.5	8.5	Albert Kranz	Salk	1
8369	Rev-1	CS77214	SWE	556,942	134,504	Magnus Nordborg	GMI	2
8419	Wil-1	CS78855	LTU	546,833	253,167	Albert Kranz	Salk	2
8422	Fjv§1-1	CS76859	SWE	56.06	14.29	Magnus Nordborg	GMI	1
8483	LP3413.53	CS79031	USA	416,862	-868,513	Justin Borevitz	MPI	1
8699	328PNA062	CS79032	USA	420,945	-863,253	Justin Borevitz	MPI	2
9079	Lerik2-1	CS77025	AZE	387,833	485,517	James Beck	Monsanto	2
9081	Lerik2-3	CS77026	AZE	387,833	485,517	James Beck	Monsanto	1
9084	Lerik2-6	CS77028	AZE	387,833	485,517	James Beck	Monsanto	2
9085	Lerik2-7	CS77029	AZE	387,833	485,517	James Beck	Monsanto	1
9111	Lag2-4	CS77005	GEO	418,296	462,831	James Beck	Monsanto	1
9113	Lag2-6	CS77006	GEO	418,296	462,831	James Beck	Monsanto	2
9133	Yeg-7	CS78867	ARM	398,692	453,622	James Beck	Monsanto	2
9314	Gol-2	CS76883	UK	579,672	-396,722	James Beck	Monsanto	1
9332	Bar 1	CS76688	SWE	628,698	18,381	Alison Anastasio	GMI	1
9343	Dju-1	CS78896	SWE	573,089	181,512	Alison Anastasio	GMI	1
9352	Dvðd 2	CS76797	SWE	572,608	163,675	Alison Anastasio	GMI	2
9353	Dvðd 3	CS76798	SWE	572,608	163,675	Alison Anastasio	GMI	1
9380	FlyA 3	CS76865	SWE	557,488	133,742	Alison Anastasio	GMI	1
9382	Fri 2	CS76869	SWE	558,106	142,091	Alison Anastasio	GMI	1
9391	Hadd-2	CS76905	SWE	573,263	158,979	Alison Anastasio	GMI	2
9394	Hag-2	CS76907	SWE	565,804	164,063	Alison Anastasio	GMI	2
9405	HolA-1 2	CS76927	SWE	557,491	13,399	Alison Anastasio	GMI	1

9407	HolA-2 2	CS76928	SWE	557,491	13,399	Alison Anastasio	GMI	1
9409	Kia 1	CS76968	SWE	560,573	14,302	Alison Anastasio	GMI	1
9412	Kor 3	CS76981	SWE	572,746	161,494	Alison Anastasio	GMI	2
9413	Kor 4	CS76982	SWE	572,746	161,494	Alison Anastasio	GMI	2
9416	Kru-3	CS76986	SWE	577,215	183,837	Alison Anastasio	GMI	1
9421	Lan 1	CS77009	SWE	559,745	143,997	Alison Anastasio	GMI	2
9451	Spro 2	CS77264	SWE	572,545	182,109	Alison Anastasio	GMI	2
9453	Stenk-2	CS77274	SWE	578,009	185,162	Alison Anastasio	GMI	1
9481	Yst-1	CS78869	SWE	554,242	138,484	Alison Anastasio	GMI	1
9507	IP-Coa-0	CS76775	POR	38.45	-7.5	Carlos Alonso-Blanco	Monsanto	2
9514	IP-Adm-0	CS76647	ESP	39.15	-4.54	Carlos Alonso-Blanco	Monsanto	1
9525	IP-Bis-0	CS76711	ESP	42.49	0.54	Xavier Picv≥	Monsanto	2
9526	IP-Cab-3	CS76738	ESP	41.54	2.39	Xavier Picv≥	Monsanto	1
9534	IP-Cmo-3	CS76774	ESP	40.05	-4.65	Carlos Alonso-Blanco	Monsanto	2
9536	IP-Cor-0	CS76782	ESP	40.83	-2	Carlos Alonso-Blanco	Monsanto	1
9545	IP-Her-12	CS76920	ESP	39.4	-5.78	Carlos Alonso-Blanco	Monsanto	1
9550	IP-Iso-4	CS76946	ESP	43.05	-5.37	Carlos Alonso-Blanco	Monsanto	1
9551	IP-Jim-1	CS76955	ESP	42.28	-5.92	Carlos Alonso-Blanco	Monsanto	2
9553	IP-Ldd-0	CS77012	ESP	41.58	-4.71	Carlos Alonso-Blanco	Monsanto	2
9555	IP-Mar-1	CS77068	ESP	39.58	-3.93	Carlos Alonso-Blanco	Monsanto	1
9557	IP-Moa-0	CS77102	ESP	42.46	0.7	Carlos Alonso-Blanco	Monsanto	2
9558	IP-Moc-11	CS77103	ESP	41.57	-5.64	Carlos Alonso-Blanco	Monsanto	1
9559	IP-Mon-5	CS77107	ESP	38.06	-4.38	Carlos Alonso-Blanco	Monsanto	2
9564	IP-Nog-17	CS77129	ESP	40.45	-1.6	Xavier Picv≥	Monsanto	1
9565	IP-Orb-10	CS77152	ESP	42.97	-1.23	Carlos Alonso-Blanco	Monsanto	2
9567	IP-Pal-0	CS77159	ESP	42.34	1.3	Xavier Picv≥	Monsanto	1
9571	IP-Pro-0	CS78914	ESP	43.28	-6.01	Carlos Alonso-Blanco	Monsanto	1
9573	IP-Rds-0	CS77206	ESP	41.86	2.99	Xavier Picv≥	Monsanto	2

9574	IP-Rel-0	CS77209	ESP	38.6	-2.7	Carlos Alonso-Blanco	Monsanto	2
9576	IP-Rev-0	CS77213	ESP	40.86	-4.11	Carlos Alonso-Blanco	Monsanto	1
9581	IP-Sdv-3	CS77242	ESP	42.84	-5.12	Carlos Alonso-Blanco	Monsanto	1
9582	IP-Ses-0	CS77244	ESP	41.48	-1.63	Carlos Alonso-Blanco	Monsanto	2
9590	IP-Trs-0	CS77387	ESP	43.37	-5.49	Carlos Alonso-Blanco	Monsanto	1
9593	IP-Vaz-0	CS78836	ESP	42.26	-2.99	Carlos Alonso-Blanco	Monsanto	1
9594	IP-Vdm-0	CS78837	ESP	42.04	1.01	Carlos Alonso-Blanco	Monsanto	2
9595	IP-Vdt-0	CS78838	ESP	40.89	-5.5	Carlos Alonso-Blanco	Monsanto	2
9598	IP-Vim-0	CS78844	ESP	41.88	-6.51	Carlos Alonso-Blanco	Monsanto	1
9599	IP-Vin-0	CS78846	ESP	42.8	-5.77	Carlos Alonso-Blanco	Monsanto	2
9601	IP-Voz-0	CS78849	ESP	41.85	-1.88	Carlos Alonso-Blanco	Monsanto	1
9602	IP-Vpa-1	CS78850	ESP	40.5	-3.96	Carlos Alonso-Blanco	Monsanto	2
9607	Panik-1	CS77161	RUS	53.05	52.15	0	Monsanto	2
9612	Lesno-2	CS77033	RUS	53.04	51.94	0	Monsanto	1
9616	Krazo-1	CS76984	RUS	53.06	51.96	0	Monsanto	1
9617	Karag-1	CS76960	RUS	51.37	59.44	0	Monsanto	2
9620	Basta-2	CS76692	RUS	51.82	79.48	0	Monsanto	2
9621	Basta-3	CS76693	RUS	51.84	79.46	0	Monsanto	2
9625	Kolyv-2	CS76977	RUS	51.31	82.59	0	Monsanto	1
9626	Kolyv-3	CS76978	RUS	51.36	82.59	0	Monsanto	2
9627	Kolyv-5	CS76979	RUS	51.32	82.55	0	Monsanto	2
9628	Kolyv-6	CS76980	RUS	51.33	82.54	0	Monsanto	2
9629	K-oze-1	CS76957	RUS	51.35	82.18	0	Monsanto	2
9630	K-oze-3	CS76958	RUS	51.34	82.16	0	Monsanto	2
9631	Lebja-1	CS77015	RUS	51.65	80.79	0	Monsanto	2
9632	Lebja-2	CS77016	RUS	51.67	80.82	0	Monsanto	2
9634	Masl-1	CS77073	RUS	54.13	81.31	0	Monsanto	2
9635	Nosov-1	CS77130	RUS	51.87	80.6	0	Monsanto	2
9636	Noveg-1	CS77131	RUS	51.75	80.82	0	Monsanto	2
9640	Rakit-1	CS77202	RUS	51.87	80.06	0	Monsanto	2
9641	Rakit-2	CS77203	RUS	51.9	80.06	0	Monsanto	2
9642	Rakit-3	CS77204	RUS	51.84	80.06	0	Monsanto	2
9643	Sever-1	CS77245	RUS	52.1	79.31	0	Monsanto	2
9644	Zupan-1	CS78882	CRO	45.07	18.72	0	Monsanto	2
9645	Gradi-1	CS76887	CRO	45.17	18.7	0	Monsanto	2
9649	Bivio-1	CS76713	ITA	39.13	16.17	0	Monsanto	2
9651	Filet-1	CS76858	ITA	40.68	14.87	0	Monsanto	2
9656	Marti-1	CS77072	ITA	40.64	17.31	0	Monsanto	2
9660	Sarno-1	CS77236	ITA	40.84	14.57	0	Monsanto	2
9664	Mitterberg-1-179	CS78907	ITA	46,366	112,832	Joerg Wunder	Monsanto	2
9669	Mitterberg-2-185	CS77086	ITA	463,668	112,837	Joerg Wunder	Monsanto	2
9670	Mitterberg-2-186	CS77087	ITA	463,668	112,837	Joerg Wunder	Monsanto	2

9671	Mitterberg-3-187	CS77088	ITA	463,668	112,837	Joerg Wunder	Monsanto	2
9672	Mitterberg-3-188	CS77089	ITA	463,668	112,837	Joerg Wunder	Monsanto	2
9673	Mitterberg-3-189	CS77090	ITA	463,668	112,837	Joerg Wunder	Monsanto	2
9676	Mitterberg-4-192	CS77093	ITA	463,718	112,866	Joerg Wunder	Monsanto	2
9678	Mitterberg-4-194	CS77095	ITA	463,718	112,866	Joerg Wunder	Monsanto	2
9679	Castelfed-1-195	CS76744	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9680	Castelfed-1-196	CS76745	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9681	Castelfed-1-197	CS76746	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9682	Castelfed-1-198	CS76747	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9685	Castelfed-2-201	CS76750	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9687	Castelfed-2-203	CS76752	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9689	Castelfed-3-205	CS76754	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9695	Castelfed-4-211	CS78892	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9696	Castelfed-4-214	CS78893	ITA	463,378	112,928	Joerg Wunder	Monsanto	2
9700	Dolna-1-10	CS76803	BUL	42.32	23.1	0	Monsanto	1
9701	Ivano-1	CS76954	BUL	43.7	25.91	0	Monsanto	2
9703	Melni-1	CS77079	BUL	41.53	23.39	0	Monsanto	1
9704	Melni-2	CS77080	BUL	41.53	23.39	0	Monsanto	2
9705	Choto-1	CS76769	BUL	41.5	23.33	0	Monsanto	2
9706	Dospa-1	CS76807	BUL	41.64	24.18	0	Monsanto	1
9710	Zerev-1-35	CS78879	BUL	41.85	23.13	0	Monsanto	2
9712	Dolna-1-40	CS76805	BUL	42.32	23.1	0	Monsanto	2
9713	Stara-1	CS77271	BUL	42.49	25.61	0	Monsanto	1
9717	Kardz-2	CS76963	BUL	41.66	25.47	0	Monsanto	1
9719	Koren-1	CS76983	BUL	41.83	25.69	0	Monsanto	2
9720	Malak-1	CS77064	BUL	41.77	25.68	0	Monsanto	1
9721	Schip-1	CS77239	BUL	42.72	25.33	0	Monsanto	2
9722	Groch-1	CS76890	BUL	41.71	24.41	0	Monsanto	2
9725	Epidauros-1	CS76844	GRC	37.6	23.08	0	Monsanto	2
9726	Faneronemi-3	CS76853	GRC	37.07	22.04	0	Monsanto	1
9728	Stiav-1	CS77279	SVK	48.46	18.9	0	Monsanto	2
9730	Bela-1	CS76696	SVK	48.47	18.94	0	Monsanto	1
9731	Stiav-3	CS77281	SVK	48.46	18.9	0	Monsanto	2
9732	Halca-1	CS76909	SVK	48.47	18.96	0	Monsanto	2
9736	Teiu-2	CS77361	ROU	44.69	25.17	0	Monsanto	1
9741	Orast-1	CS77151	ROU	45.84	23.16	0	Monsanto	1
9745	Sij 1/96	CS77249	UZB	41.45	70.05	0	Monsanto	2
9748	Zagub-1	CS78871	SRB	44.23	21.71	0	Monsanto	2
9749	Knjas-1	CS76971	SRB	43.54	22.29	0	Monsanto	1
9754	Sredn-1	CS77269	SRB	44.66	21.37	0	Monsanto	2
9755	Vajug-1	CS78828	SRB	44.56	22.56	0	Monsanto	2
9757	Staro-1	CS77272	SRB	44.3	21.08	0	Monsanto	2
9758	Altai-5	CS76433	CHN	47.75	88.4	0	Salk	1
9759	Anz-0	CS76439	IRN	37.47	49.47	0	Salk	2
9761	Bik-1	CS76449	LBN	33.92	35.7	0	Salk	2
9762	Etna-2	CS76487	ITA	37.69	14.98	0	Salk	1
9764	Qar-8a	CS76581	LBN	34.1	35.84	0	Salk	1
9769	HE-1	CS76916	GER	48.55	8.99	0	Monsanto	2
9770	KBG2-13	CS76966	GER	48.53	9.01	0	Monsanto	2
9772	Hof-1	CS76925	GER	48.41	8.85	0	Monsanto	2
9774	Alt-1	CS76663	GER	48.59	9.22	0	Monsanto	1
9775	Berg-1	CS76701	GER	48.41	8.79	0	Monsanto	2
9776	Fell3-7	CS76857	GER	48.43	8.79	0	Monsanto	1

9777	Gn-1	CS76880	GER	48.57	9.17	0	Monsanto	2
9778	Bach-7	CS76679	GER	48.41	8.84	0	Monsanto	1
9780	Fell2-4	CS76856	GER	48.43	8.79	0	Monsanto	2
9782	Lu3-30	CS77057	GER	48.53	9.09	0	Monsanto	2
9783	Tu-PK-7	CS77396	GER	48.52	9.05	0	Monsanto	1
9784	Erg2-6	CS76845	GER	48.5	8.8	0	Monsanto	2
9785	Ha-HBT1-2	CS76898	GER	48.54	9.02	0	Monsanto	2
9786	Ha-P-13	CS76901	GER	48.54	9.01	0	Monsanto	1
9788	KBG1-14	CS76965	GER	48.53	9.01	0	Monsanto	2
9790	Gn2-3	CS76881	GER	48.58	9.18	0	Monsanto	2
9791	Haes-1	CS76914	GER	48.6	9.2	0	Monsanto	2
9792	Lu4-2	CS77058	GER	48.54	9.09	0	Monsanto	1
9793	Ru-N2	CS77224	GER	48.57	9.16	0	Monsanto	2
9794	Tu-B1-2	CS77391	GER	48.52	9.08	0	Monsanto	2
9797	Ha-HBT2-10	CS76899	GER	48.54	9.02	0	Monsanto	2
9798	Ha-P2-1	CS76902	GER	48.54	9.01	0	Monsanto	1
9801	Ha-SP-2	CS76904	GER	48.54	9.01	0	Monsanto	2
9802	Kus3-1	CS76991	GER	48.51	9.11	0	Monsanto	1
9803	Muh-2	CS77113	GER	48.42	8.76	0	Monsanto	2
9804	Obe1-15	CS77139	GER	48.45	8.87	0	Monsanto	1
9806	Ru-2	CS77223	GER	48.56	9.16	0	Monsanto	2
9808	Tu-B2-3	CS77392	GER	48.52	9.08	0	Monsanto	1
9810	Tu-KS-7	CS77394	GER	48.53	9.07	0	Monsanto	1
9811	Tu-NK-12	CS77395	GER	48.52	9.05	0	Monsanto	2
9812	Tu-W1	CS77397	GER	48.52	9.03	0	Monsanto	2
9813	BI-4	CS76706	GER	48.4	8.77	0	Monsanto	1
9816	Tu-WH	CS77398	GER	48.55	9.06	0	Monsanto	2
9821	IP-Aru-0	CS76674	ESP	41.81	2.49	Xavier Picv≥	Monsanto	1
9822	IP-Aul-0	CS76675	ESP	40.52	-4.02	Carlos Alonso-Blanco	Monsanto	2
9823	IP-Bae-0	CS76681	ESP	43.34	-5.84	Carlos Alonso-Blanco	Monsanto	1
9825	IP-Boa-0	CS76714	ESP	40.4	-3.88	Carlos Alonso-Blanco	Monsanto	2
9826	IP-Bor-0	CS76717	ESP	42.49	-6.71	Carlos Alonso-Blanco	Monsanto	1
9827	IP-Bos-0	CS76719	ESP	42.78	0.69	Xavier Picv≥	Monsanto	2
9828	IP-Bra-0	CS76721	ESP	42.5	-6.15	Xavier Picv≥	Monsanto	1
9831	IP-Cas-0	CS76743	ESP	38.54	-3.39	Carlos Alonso-Blanco	Monsanto	2
9834	IP-Cho-0	CS76768	ESP	40.51	-3.9	Carlos Alonso-Blanco	Monsanto	2
9835	IP-Cir-0	CS76772	ESP	40.61	-6.57	Carlos Alonso-Blanco	Monsanto	1
9836	IP-Cod-0	CS76777	ESP	41.25	-1.32	Carlos Alonso-Blanco	Monsanto	2
9838	IP-Cot-0	CS76784	ESP	41.83	-5.38	Carlos Alonso-Blanco	Monsanto	1
9839	IP-Coy-0	CS76785	ESP	40.44	-4.27	Carlos Alonso-Blanco	Monsanto	2
9840	IP-Dar-0	CS76792	ESP	41.13	-1.43	Carlos Alonso-Blanco	Monsanto	1
9841	IP-Ees-0	CS76836	ESP	40.59	-4.15	Carlos Alonso-Blanco	Monsanto	2

9843	IP-Elp-0	CS76840	ESP	40.53	-3.92	Carlos Alonso-Blanco	Monsanto	1
9844	IP-Esn-2	CS76846	ESP	42.27	0.19	Carlos Alonso-Blanco	Monsanto	2
9845	IP-Evs-0	CS76848	ESP	40.48	-3.96	Carlos Alonso-Blanco	Monsanto	2
9846	IP-Ezc-2	CS76849	ESP	42.31	-3.02	Carlos Alonso-Blanco	Monsanto	2
9848	IP-Glo-1	CS76879	ESP	40.11	-5.77	Carlos Alonso-Blanco	Monsanto	1
9850	IP-Hec-0	CS76917	ESP	42.86	-0.7	Carlos Alonso-Blanco	Monsanto	2
9852	IP-Ini-0	CS76945	ESP	40.46	-3.75	Carlos Alonso-Blanco	Monsanto	1
9853	IP-Lac-0	CS76996	ESP	43.33	-5.91	Carlos Alonso-Blanco	Monsanto	2
9854	IP-Laf-1	CS76997	ESP	43.36	-5.88	Carlos Alonso-Blanco	Monsanto	2
9856	IP-Lch-0	CS77010	ESP	40.51	-4	Carlos Alonso-Blanco	Monsanto	2
9857	IP-Leg-0	CS77019	ESP	40.33	-3.8	Carlos Alonso-Blanco	Monsanto	1
9858	IP-Loz-0	CS77051	ESP	40.98	-3.8	Carlos Alonso-Blanco	Monsanto	2
9859	IP-Lro-0	CS77054	ESP	40.5	-3.88	Carlos Alonso-Blanco	Monsanto	2
9860	IP-Lum-0	CS77059	ESP	42.24	-2.62	Carlos Alonso-Blanco	Monsanto	1
9861	IP-Mac-0	CS77061	ESP	40.72	-3.21	Carlos Alonso-Blanco	Monsanto	2
9862	IP-Mad-0	CS77062	ESP	40.45	-3.67	Carlos Alonso-Blanco	Monsanto	2
9864	IP-Mat-0	CS77074	ESP	41.76	2.69	Xavier Picv≥	Monsanto	2
9866	IP-Mdd-0	CS77076	ESP	41.89	-2.79	Carlos Alonso-Blanco	Monsanto	2
9867	IP-Mie-1	CS77083	ESP	40.94	-3.22	Xavier Picv≥	Monsanto	2
9869	IP-Moj-0	CS77105	ESP	36.76	-5.28	Xavier Picv≥	Monsanto	2
9870	IP-Moz-0	CS77111	ESP	41.91	0.17	Carlos Alonso-Blanco	Monsanto	1
9871	IP-Nac-0	CS77117	ESP	40.75	-3.99	Carlos Alonso-Blanco	Monsanto	2
9875	IP-Ovi-1	CS77155	ESP	43.38	-5.87	Carlos Alonso-Blanco	Monsanto	2
9876	IP-Pad-0	CS77158	ESP	41.34	0.99	Carlos Alonso-Blanco	Monsanto	2

9877	IP-Pdl-0	CS77165	ESP	43.02	-5.6	Carlos Alonso-Blanco	Monsanto	2
9878	IP-Pee-0	CS77167	ESP	40.78	-3.62	Carlos Alonso-Blanco	Monsanto	1
9880	IP-Pib-1	CS77175	ESP	42.72	-3.44	Carlos Alonso-Blanco	Monsanto	2
9882	IP-Pil-0	CS77178	ESP	40.46	-4.26	Carlos Alonso-Blanco	Monsanto	2
9883	IP-Piq-0	CS77179	ESP	42.1	-2.56	Carlos Alonso-Blanco	Monsanto	1
9885	IP-Prd-0	CS77189	ESP	41.14	-3.68	Carlos Alonso-Blanco	Monsanto	2
9887	IP-Pun-0	CS77196	ESP	40.4	-4.77	Carlos Alonso-Blanco	Monsanto	2
9890	IP-Rib-1	CS77217	ESP	43.16	-5.07	Carlos Alonso-Blanco	Monsanto	1
9891	IP-Sal-0	CS77230	ESP	41.93	2.92	Xavier Picv≥	Monsanto	2
9892	IP-Sam-0	CS77231	ESP	42.68	-6.96	Carlos Alonso-Blanco	Monsanto	2
9898	IP-Som-0	CS77259	ESP	41.14	-3.58	Carlos Alonso-Blanco	Monsanto	1
9899	IP-Tau-0	CS77342	ESP	42.54	0.84	Xavier Picv≥	Monsanto	2
9901	IP-Urd-1	CS78824	ESP	42.27	-2.98	Carlos Alonso-Blanco	Monsanto	2
9902	IP-Usa-0	CS78825	ESP	40.71	-3.24	Carlos Alonso-Blanco	Monsanto	2
9903	IP-Val-0	CS78829	ESP	42.31	-3.1	Carlos Alonso-Blanco	Monsanto	1
9908	ESP-1-11	CS76847	FRA	50.72	3.47	0	Monsanto	1
9911	ARGE-1-15	CS76672	FRA	47.16	4.28	0	Monsanto	2
9912	CIRY-13	CS76773	FRA	46.67	4.55	0	Monsanto	1
9918	SAUL-24	CS77237	FRA	47.43	5.21	0	Monsanto	1
9925	RUM-20	CS77226	FRA	48.91	4.52	0	Monsanto	2
9926	TRE-1	CS77385	FRA	48.86	4.1	0	Monsanto	2
9928	BEZ-9	CS76703	FRA	44.12	3.77	0	Monsanto	1
9937	CATS-6	CS76760	FRA	50.79	2.69	0	Monsanto	2
9938	WAV-8	CS78854	FRA	50.65	2.99	0	Monsanto	2
9939	Aitba-2	CS76347	MAR	31.48	-7.45		MPI	1
9944	Don-0	CS76411	ESP	36.83	-6.36		MPI	2
9946	Mer-6	CS76414	ESP	38.92	-6.34		MPI	2
9948	Pra-6	CS76416	ESP	41.05	-3.54		MPI	1
9949	Qui-0	CS76417	ESP	42.69	-6.93		MPI	2
9950	Vie-0	CS76418	ESP	42.63	0.76		MPI	1
9952	Kly-4	CS76384	RUS	51.32	82.55		MPI	2
9955	Stepn-2	CS76377	RUS	54.09	60.46		MPI	2
9956	Stepn-1	CS76378	RUS	54.06	60.48		MPI	1
9957	Borsk-2	CS76421	RUS	53.04	51.75		MPI	2
9958	Shigu-1	CS76375	RUS	53.33	49.48		MPI	2
9962	Galdo-1	CS76423	ITA	40.57	15.32		MPI	1
9963	Lago-1	CS76367	ITA	39.18	16.26		MPI	2
9964	Mammo-1	CS76365	ITA	38.36	16.23		MPI	2
9965	Mammo-2	CS76364	ITA	38.38	16.22		MPI	1
9966	Monte-1	CS76361	ITA	40.28	15.65		MPI	2
9968	Timpo-1	CS76424	ITA	39.27	16.27		MPI	2

9970	Altenb-2	CS76353	ITA	463,716	112,376	Joerg Wunder	MPI	1
9974	Castelfed-4-212	CS76355	ITA	463,378	112,928	Joerg Wunder	MPI	2
9975	Castelfed-4-213	CS76356	ITA	463,378	112,928	Joerg Wunder	MPI	2
9976	Rovero-1	CS76351	ITA	462,543	111,670	Joerg Wunder	MPI	2
9978	Vezzano-2.2	CS76350	ITA	466,297	108,170	Joerg Wunder	MPI	1
9979	Voeran-1	CS76352	ITA	46.36	11.23		MPI	1
9980	Angel-1	CS76362	ITA	38.62	16.17		MPI	2
9981	Angit-1	CS76366	ITA	38.76	16.24		MPI	1
9983	Ciste-1	CS76359	ITA	41.62	12.87		MPI	2
9991	Vash-1	CS76391	GEO	412,381	463,728	James Beck	MPI	1
9995	HKT2.4	CS76404	GER	48.14	9.4		MPI	2
9996	Nie1-2	CS76402	GER	48.52	8.8		MPI	2
9997	Rue3-1-31	CS76406	GER	48.56	9.16		MPI	1
9999	TueSB30-3	CS76403	GER	48.53	9.06		MPI	1
10001	TueV-13	CS76407	GER	48.52	9.05	Kirsten Bomblies	MPI	1
10002	TueWa1-2	CS76405	GER	48.53	9.04	Kirsten Bomblies	MPI	1
10004	Bolin-1	CS76373	ROU	44.46	25.74		MPI	2
10005	Copac-1	CS76420	ROU	46.11	21.95		MPI	2
10006	Kastel-1	CS76395	UKR	446,419	343,814	James Beck	MPI	1
10008	Sij-1	CS76379	UZB	41.45	70.05	Heike Schmuths	MPI	2
10009	Sij-2	CS76380	UZB	41.45	70.05		MPI	2
10010	Sij-4	CS76381	UZB	41.45	70.05		MPI	2
10011	Yeg-1	CS76394	ARM	398,692	453,622	James Beck	MPI	2
10012	Istisu-1	CS76389	AZE	389,786	485,594	James Beck	MPI	2
10013	Lerik1-3	CS76388	AZE	387,406	486,131	James Beck	MPI	2
10014	Xan-1	CS76387	AZE	386,536	487,992	James Beck	MPI	1
10015	Ara-1	CS76382	AFG	37.29	71.3		MPI	2
10017	Petro-1	CS76370	SRB	44.34	21.46		MPI	2
10018	Dobra-1	CS76369	SRB	44.84	20.16		MPI	2
10020	Jl-2	CS76956	CZE	49.17	16.5		MPI	2
10022	Uk-3	CS78777	GER	480,333	77,667	Albert Kranz	MPI	1
10023	Strand-1	CS77284	NOR	68.8	15.45		MPI	2
10027	Uk-6	CS78938	GER	4,802,838	7,765,567	Eunyoung Chae	MPI	2
14312	Kos-1	CS78923	RUS	62.02	34.12	Magnus Nordborg	GMI	2
14313	Kos-2	CS78924	RUS	62.02	34.12	Magnus Nordborg	GMI	2
14314	Radk-1	CS78927	RUS	61.59	35.11	Magnus Nordborg	GMI	2
14315	Radk-2	CS78928	RUS	61.59	35.11	Magnus Nordborg	GMI	1
14318	Shu-1	CS78930	RUS	61.94	34.24	Magnus Nordborg	GMI	1
15560	Valm	CS78932	RUS	61.37	61.37	Magnus Nordborg	GMI	2
15591	OOE1-1	CS78939	AUT	483,315,333	1,472,665	Wolfram Weckwert	GMI	2
15592	OOE3-1	CS78940	AUT	483,314,667	147,158,667	Wolfram Weckwert	GMI	2
15593	OOE3-2	CS78941	AUT	483,314,667	147,158,667	Wolfram Weckwert	GMI	2
18694	Pien	CS78926	RUS	60.4	32.09	Magnus Nordborg	GMI	1

19949	OOE2-1	CS79036	AUT	483,268,333	147,181,667	Wolfram Weckwert	GMI	2
19951	OOE23	CS79038	AUT	483,268,333	147,181,667	Wolfram Weckwert	GMI	1

Supplementary Table S2. Selected accessions for Chapter 2 GWAS.

tg_ecotypeid	name	CS_number	country	latitude	longitude	collector	seq_by	block
88	CYR	CS76790	FRA	47.4	0.683333	Valerie Le Corre	Monsanto	1
108	LDV-18	CS77013	FRA	485,167	-406,667	Valerie Le Corre	Monsanto	1
139	LDV-46	CS77014	FRA	485,167	-406,667	Valerie Le Corre	Monsanto	1
159	MAR2-3	CS77070	FRA	47.35	393,333	Valerie Le Corre	Monsanto	1
265	PYL-6	CS77198	FRA	44.65	-116,667	Valerie Le Corre	Monsanto	1
350	TOU-A1-88	CS77382	FRA	466,667	411,667	Fabrice Roux	Monsanto	1
351	TOU-A1-89	CS77383	FRA	466,667	411,667	Fabrice Roux	Monsanto	1
403	Zdarec3	CS78873	CZE	493,667	162,667	Marie-Theres Hauser	Monsanto	1
410	Doubravnik7	CS76808	CZE	494,211	163,497	Marie-Theres Hauser	Monsanto	1
424	Draha2	CS76812	CZE	494,112	162,815	Marie-Theres Hauser	Monsanto	1
428	Borky1	CS76718	CZE	49,403	16,232	Marie-Theres Hauser	Monsanto	1
430	Gr-1	CS76496	AUT	47	15.5	Albert Kranz	Salk	1
470	BRR4	CS78943	USA	408,313	-87,735	Diane Byers	MPI	1
476	BRR12	CS78944	USA	408,313	-87,735	Diane Byers	MPI	1
484	BRR23	CS78945	USA	408,313	-87,735	Diane Byers	MPI	1
504	BRR57	CS78946	USA	408,313	-87,735	Diane Byers	MPI	1
506	BRR60	CS78947	USA	408,313	-87,735	Diane Byers	MPI	1
531	BRR107	CS78948	USA	408,313	-87,735	Diane Byers	MPI	1
544	LI-WP-039	CS78949	USA	409,076	-732,089	Oliver Bossdorf	MPI	1
546	LI-WP-041	CS78950	USA	409,076	-732,089	Oliver Bossdorf	MPI	1
628	LI-OF-061	CS78951	USA	407,777	-729,069	Oliver Bossdorf	MPI	1
630	LI-OF-065	CS77036	USA	407,777	-729,069	Oliver Bossdorf	Monsanto	1
680	LI-RR-096	CS78952	USA	409,447	-728,615	Oliver Bossdorf	MPI	1
681	LI-RR-097	CS78953	USA	409,447	-728,615	Oliver Bossdorf	MPI	1
685	LI-EF-011	CS78954	USA	409,064	-731,493	Oliver Bossdorf	MPI	1
687	LI-EF-018	CS78955	USA	409,064	-731,493	Oliver Bossdorf	MPI	1
728	LI-SET-019	CS78956	USA	409,352	-73,114	Oliver Bossdorf	MPI	1

742	LI-SET-036	CS78957	USA	409,352	-73,114	Oliver Bossdorf	MPI	1
763	Kar-1	CS76522	KGZ	42.3	743,667	Olivier Loudet	Salk	1
765	Sus-1	CS76607	KGZ	421,833	73.4	Olivier Loudet	Salk	1
766	Dja-1	CS76473	KGZ	425,833	736,333	Olivier Loudet	Salk	1
768	Zal-1	CS76634	KGZ	42.8	76.35	Olivier Loudet	Salk	1
772	Neo-6	CS76560	TJK	37.35	724,667	Olivier Loudet	Salk	1
801	KYC-33	CS76992	USA	379,169	-844,639	Kathleen Donohue	Monsanto	1
853	MIA-1	CS78958	USA	417,976	-866,691	Kathleen Donohue	MPI	1
854	MIA-5	CS78959	USA	417,976	-866,691	Kathleen Donohue	MPI	1
867	MIC-20	CS78960	USA	418,266	-864,366	Kathleen Donohue	MPI	1
868	MIC-24	CS78961	USA	418,266	-864,366	Kathleen Donohue	MPI	1
870	MIC-31	CS77082	USA	418,266	-864,366	Kathleen Donohue	Monsanto	1
915	LIN S-5	CS77040	USA	418,972	-714,378	Kathleen Donohue	Monsanto	1
932	CHA-41	CS76765	USA	423,634	-711,445	Kathleen Donohue	Monsanto	1
991	Ale-Stenar-41-1	CS76651	SWE	553,833	14.05	Jon vÖgren	GMI	1
992	Ale-Stenar-44-4	CS76652	SWE	553,833	14.05	Jon vÖgren	GMI	1
997	Ale-Stenar-56-14	CS76653	SWE	553,833	14.05	Jon vÖgren	GMI	1
1002	Ale-Stenar-64-24	CS76654	SWE	553,833	14.05	Jon vÖgren	GMI	1
1006	Ale-Stenar-77-31	CS77636	SWE	553,833	14.05	Jon vÖgren	GMI	1
1061	Brvðsarp-11-135	CS76727	SWE	557,167	141,333	Jon vÖgren	GMI	1
1062	Brvðsarp-15-138	CS76728	SWE	557,167	141,333	Jon vÖgren	GMI	1
1063	Brvðsarp-21-140	CS76729	SWE	557,167	141,333	Jon vÖgren	GMI	1
1066	Brvðsarp-34-145	CS76730	SWE	557,167	141,333	Jon vÖgren	GMI	1
1070	Brvðsarp-45-153	CS77643	SWE	557,167	141,333	Jon vÖgren	GMI	1
1158	Aledal-6-49	CS76656	SWE	56.7	165,167	Jon vÖgren	GMI	1
1166	Aledal-14-73	CS77651	SWE	56.7	165,167	Jon vÖgren	GMI	1
1254	Tos-82-387	CS77379	SWE	594,333	170,167	Jon vÖgren	GMI	1
1257	Tos-95-393	CS77380	SWE	594,333	170,167	Jon vÖgren	GMI	1
1313	vÑngsvð-59-422	CS77658	SWE	595,667	168,667	Jon vÖgren	GMI	1
1317	vÑngsvð-74-430	CS76665	SWE	595,667	168,667	Jon vÖgren	GMI	1
1552	Sku-30	CS77251	SWE	630,833	183,667	Jon vÖgren	GMI	1
1612	Brn-10	CS78962	USA	41.9	-86,583	Diane Byers	MPI	1

1622	Brn-24	CS78963	USA	41.9	-86,583	Diane Byers	MPI	1
1651	DuckLkSP 38	CS78964	USA	433,431	-864,045	Diane Byers	MPI	1
1652	DuckLkSP 40	CS78965	USA	433,431	-864,045	Diane Byers	MPI	1
1676	Haz-2	CS78966	USA	41,879	-86,607	Diane Byers	MPI	1
1684	Haz-10	CS78967	USA	41,879	-86,607	Diane Byers	MPI	1
1739	KBS-Mac-68	CS78968	USA	42,405	-85,398	Diane Byers	MPI	1
1741	KBS-Mac-74	CS78969	USA	42,405	-85,398	Diane Byers	MPI	1
1756	Ker-4	CS78970	USA	42,184	-86,358	Diane Byers	MPI	1
1757	Ker-5	CS78971	USA	42,184	-86,358	Diane Byers	MPI	1
1793	L-R-5	CS78972	USA	41,847	-86.67	Diane Byers	MPI	1
1797	L-R-10	CS78973	USA	41,847	-86.67	Diane Byers	MPI	1
1820	Lak-13	CS78975	USA	41.8	-86.67	Diane Byers	MPI	1
1829	Mdn-1	CS77077	USA	42,051	-86,509	Diane Byers	Monsanto	1
1834	Mdn-8	CS78976	USA	42,051	-86,509	Diane Byers	MPI	1
1835	Mdn-10	CS78977	USA	42,051	-86,509	Diane Byers	MPI	1
1852	MNF-Pot-15	CS78979	USA	43,595	-862,657	Diane Byers	MPI	1
1853	MNF-Pot-21	CS77099	USA	43,595	-862,657	Diane Byers	Monsanto	1
1872	MNF-Pot-75	CS77100	USA	43,595	-862,657	Diane Byers	Monsanto	1
1890	MNF-Riv-21	CS77101	USA	435,139	-861,859	Diane Byers	Monsanto	1
1925	MNF-Che-2	CS77096	USA	435,251	-861,843	Diane Byers	Monsanto	1
1942	MNF-Che-47	CS78980	USA	435,251	-861,843	Diane Byers	MPI	1
1943	MNF-Che-49	CS78981	USA	435,251	-861,843	Diane Byers	MPI	1
1954	MNF-Jac-12	CS77097	USA	435,187	-861,739	Diane Byers	Monsanto	1
2016	MNF-Pin-39	CS77098	USA	435,356	-861,788	Diane Byers	Monsanto	1
2031	Map-8	CS78983	USA	42,166	-86,412	Diane Byers	MPI	1
2053	Map-35	CS78984	USA	42,166	-86,412	Diane Byers	MPI	1
2057	Map-42	CS77732	USA	42,166	-86,412	Diane Byers	MPI	1
2081	MuskSP-68	CS78985	USA	432,483	-863,368	Diane Byers	MPI	1
2091	MuskSP-83	CS78986	USA	432,483	-863,368	Diane Byers	MPI	1
2106	MSGA-10	CS78987	USA	432,749	-860,891	Diane Byers	MPI	1
2108	MSGA-12	CS78988	USA	432,749	-860,891	Diane Byers	MPI	1
2141	MSGA-61	CS78989	USA	432,749	-860,891	Diane Byers	MPI	1
2159	Paw-13	CS78990	USA	42,148	-86,431	Diane Byers	MPI	1
2166	Paw-20	CS78991	USA	42,148	-86,431	Diane Byers	MPI	1

2171	Paw-26	CS77164	USA	42,148	-86,431	Diane Byers	Monsanto	1
2191	Pent-7	CS78992	USA	437,623	-863,929	Diane Byers	MPI	1
2202	Pent-23	CS77168	USA	437,623	-863,929	Diane Byers	Monsanto	1
2212	Pent-46	CS78993	USA	437,623	-863,929	Diane Byers	MPI	1
2239	Riv-25	CS78994	USA	42,184	-86,382	Diane Byers	MPI	1
2240	Riv-26	CS78995	USA	42,184	-86,382	Diane Byers	MPI	1
2276	SLSP-31	CS77254	USA	43,665	-86,496	Diane Byers	Monsanto	1
2278	SLSP-35	CS77255	USA	43,665	-86,496	Diane Byers	Monsanto	1
2285	SLSP-69	CS78996	USA	43,665	-86,496	Diane Byers	MPI	1
2286	SLSP-67	CS78997	USA	43,665	-86,496	Diane Byers	MPI	1
2317	Ste-40	CS77278	USA	42.03	-86,514	Diane Byers	Monsanto	1
2370	Yng-4	CS78998	USA	41,865	-86,646	Diane Byers	MPI	1
2412	Yng-53	CS78999	USA	41,865	-86,646	Diane Byers	MPI	1
4779	UKSW06-179	CS78808	UK	50.4	-4.9	Eric Holub	Monsanto	1
4807	UKSW06-207	CS78809	UK	50.4	-4.9	Eric Holub	Monsanto	1
4826	UKSW06-226	CS78810	UK	50.4	-4.9	Eric Holub	Monsanto	1
4840	UKSW06-240	CS79000	UK	50.4	-4.9	Eric Holub	MPI	1
4857	UKSW06-257	CS79001	UK	50.3	-4.9	Eric Holub	MPI	1
4884	UKSW06-285	CS78811	UK	50.3	-4.9	Eric Holub	Monsanto	1
4900	UKSW06-302	CS78812	UK	50.3	-4.8	Eric Holub	Monsanto	1
4939	UKSW06-341	CS79002	UK	50.4	-4.7	Eric Holub	MPI	1
4958	UKSW06-360	CS78814	UK	50.5	-4.5	Eric Holub	Monsanto	1
5023	UKSE06-118	CS78799	UK	51.3	0.5	Eric Holub	Monsanto	1
5104	UKSE06-252	CS78800	UK	51.3	0.5	Eric Holub	Monsanto	1
5151	UKSE06-325	CS78801	UK	52.2	-1.7	Eric Holub	Monsanto	1
5165	UKSE06-362	CS78802	UK	51.3	0.4	Eric Holub	Monsanto	1
5210	UKSE06-432	CS78803	UK	51.2	0.3	Eric Holub	Monsanto	1
5236	UKSE06-470	CS78804	UK	51.2	0.4	Eric Holub	Monsanto	1
5249	UKSE06-491	CS79003	UK	51.2	0.3	Eric Holub	MPI	1
5253	UKSE06-500	CS78805	UK	51.1	0.6	Eric Holub	Monsanto	1
5276	UKSE06-533	CS78806	UK	51.3	1.1	Eric Holub	Monsanto	1
5279	UKSE06-541	CS79004	UK	51.3	1.1	Eric Holub	MPI	1
5349	UKSE06-639	CS78807	UK	51.1	0.4	Eric Holub	Monsanto	1
5353	UKNW06-003	CS78792	UK	54.5	-3	Eric Holub	Monsanto	1

5395	UKNW06-102	CS79005	UK	54.4	-3	Eric Holub	MPI	1
5486	UKNW06-233	CS78794	UK	54.6	-3.3	Eric Holub	Monsanto	1
5577	UKNW06-403	CS78797	UK	54.7	-3.4	Eric Holub	Monsanto	1
5644	UKNW06-481	CS78798	UK	54.4	-2.9	Eric Holub	Monsanto	1
5651	UKNW06-488	CS79006	UK	54.4	-2.9	Eric Holub	MPI	1
5717	Bra-1	CS79007	UK	54.6	-3.2	Eric Holub	MPI	1
5718	UKID11	CS79008	UK	57	-3.4	Eric Holub	MPI	1
5720	Cal-2	CS78781	UK	53.3	-1.6	D. Ratcliffe	Monsanto	1
5726	Cnt-1	CS78782	UK	51.3	1.1	Eric Holub	Monsanto	1
5741	For-2	CS78783	UK	56.6	-4.1	D. Ratcliffe	Monsanto	1
5748	Kil-0	CS78784	UK	56	-4.4	D. Ratcliffe	Monsanto	1
5757	Mc-1	CS78785	UK	54.6	-2.3	D. Ratcliffe	Monsanto	1
5768	UKID63	CS78786	UK	54.1	-1.5	Eric Holub	Monsanto, MPI	1
5772	Set-1	CS78787	UK	54.1	-2.3	D. Ratcliffe	Monsanto	1
5776	UKID71	CS79009	UK	52.9	-1.3	Eric Holub	MPI	1
5779	UKID74	CS78789	UK	51	-3.1	Eric Holub	Monsanto	1
5784	Ty-1	CS78790	UK	56.4	-5.2	D. Ratcliffe	Monsanto	1
5798	UKID93	CS79010	UK	53.1	-3.3	Eric Holub	MPI	1
5800	UKID96	CS78791	UK	57.4	-5.5	Eric Holub	Monsanto	1
5811	UKID107	CS78778	UK	52.9	-3.1	Eric Holub	Monsanto	1
5822	UKID116	CS78780	UK	567,333	-598,333	Eric Holub	Monsanto	1
5830	App1-12	CS76667	SWE	563,333	159,667	Magnus Nordborg	GMI	1
5831	App1-14	CS76668	SWE	563,333	159,667	Magnus Nordborg	GMI	1
5832	App1-16	CS76669	SWE	563,333	159,667	Magnus Nordborg	GMI	1
5836	Boo2-3	CS77906	SWE	55.86	13.51	Magnus Nordborg	GMI	1
5837	Bor-1	CS76453	CZE	494,013	162,326	Jirina Relichov	Salk	1
5856	Dvðr-10	CS76806	SWE	630,167	174,914	Magnus Nordborg	GMI	1
5860	Dra-3	CS77913	SWE	626,814	180,165	Magnus Nordborg	GMI	1
5865	Dra1-4	CS76809	SWE	55.76	14.12	Magnus Nordborg	GMI	1
5867	Dra2-1	CS76810	SWE	55.76	14.12	Magnus Nordborg	GMI	1
5874	DraII-6	CS76814	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5890	DraIV 1-8	CS76817	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5893	DraIV 1-11	CS76816	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5907	DraIV 2-9	CS76818	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5921	DraIV 3-7	CS76819	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5950	DraIV 5-12	CS76820	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5984	DraIV 6-13	CS76822	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
5993	DraIV 6-22	CS76823	CZE	494,112	162,815	Jirina Relichov	Monsanto	1
6008	Duk	CS76824	CZE	49.1	16.2	Jirina Relichov	Salk	1
6009	Eden-1	CS76826	SWE	62,877	18,177	Magnus Nordborg	GMI	1
6010	Eden-5	CS78000	SWE	62,877	18,177	Magnus Nordborg	GMI	1
6011	Eden-6	CS76828	SWE	62,877	18,177	Magnus Nordborg	GMI	1

6012	Eden-7	CS76829	SWE	62,877	18,177	Magnus Nordborg	GMI	1
6013	Eden-9	CS76830	SWE	62,877	18,177	Magnus Nordborg	GMI	1
6016	Eds-1	CS76834	SWE	62.9	18.4	Magnus Nordborg	GMI	1
6017	Eds-9	CS76835	SWE	62.9	18.4	Magnus Nordborg	GMI	1
6019	Fjv§1-2	CS76860	SWE	56.06	14.29	Magnus Nordborg	GMI	1
6020	Fjv§1-5	CS76861	SWE	56.06	14.29	Magnus Nordborg	GMI	1
6021	Fjv§2-4	CS76862	SWE	56.06	14.29	Magnus Nordborg	GMI	1
6022	Fjv§2-6	CS78009	SWE	56.06	14.29	Magnus Nordborg	GMI	1
6023	Fly2-1	CS76863	SWE	557,509	133,712	Magnus Nordborg	GMI	1
6024	Fly2-2	CS76864	SWE	557,509	133,712	Magnus Nordborg	GMI	1
6025	Gro-3	CS76889	SWE	626,437	177,339	Magnus Nordborg	GMI	1
6030	Grvðn-5	CS76893	SWE	62,806	181,896	Magnus Nordborg	GMI	1
6034	Hov1-7	CS76932	SWE	56.1	13.74	Magnus Nordborg	GMI	1
6035	Hov1-10	CS76931	SWE	56.1	13.74	Magnus Nordborg	GMI	1
6036	Hov3-2	CS76934	SWE	56.1	13.74	Magnus Nordborg	GMI	1
6038	Hov3-5	CS76935	SWE	56.1	13.74	Magnus Nordborg	GMI	1
6039	Hovdala-2	CS76937	SWE	56.1	13.74	Torbjorn Sall	GMI	1
6040	Kni-1	CS76970	SWE	55.66	13.4	Magnus Nordborg	GMI	1
6041	Lis-3	CS77044	SWE	560,328	14,775	Magnus Nordborg	GMI	1
6042	Lom1-1	CS77048	SWE	56.09	13.9	Magnus Nordborg	GMI	1
6043	Lvðv-1	CS77049	SWE	62,801	18,079	Magnus Nordborg	GMI	1
6046	Lvðv-5	CS77050	SWE	62,801	18,079	Magnus Nordborg	GMI	1
6064	Nyl-2	CS77136	SWE	629,513	182,763	Magnus Nordborg	GMI	1
6069	Nyl-7	CS77137	SWE	629,513	182,763	Magnus Nordborg	GMI	1
6070	Omn-1	CS77145	SWE	629,308	183,448	Magnus Nordborg	GMI	1
6071	Omn-5	CS77146	SWE	629,308	183,448	Magnus Nordborg	GMI	1
6073	vñMvð1-7	CS77147	SWE	561,481	158,155	Magnus Nordborg	GMI	1
6074	vñr-1	CS77150	SWE	564,573	161,308	Magnus Nordborg	GMI	1
6076	Rev-2	CS77215	SWE	556,942	134,504	Magnus Nordborg	GMI	1
6077	Rev-3	CS78030	SWE	556,942	134,504	Magnus Nordborg	GMI	1
6085	Sparta-1	CS77260	SWE	557,097	132,145	Magnus Nordborg	GMI	1
6086	Sr:3	CS77267	SWE	58.9	11.2	Ivo Cetl	GMI	1
6087	Stu-2	CS78033	SWE	564,666	161,284	Magnus Nordborg	GMI	1
6088	Stu1-1	CS77285	SWE	564,666	161,284	Magnus Nordborg	GMI	1
6090	T1000	CS77288	SWE	556,525	132,197	Mattias Jakobsson	GMI	1

6091	T1010	CS78035	SWE	556,525	13,215	Mattias Jakobsson	GMI	1
6092	T1020	CS77289	SWE	556,514	132,233	Mattias Jakobsson	GMI	1
6094	T1040	CS77290	SWE	556,494	132,147	Mattias Jakobsson	GMI	1
6095	T1050	CS78039	SWE	556,486	132,161	Mattias Jakobsson	GMI	1
6097	T1070	CS77291	SWE	556,481	132,264	Mattias Jakobsson	GMI	1
6098	T1080	CS77292	SWE	556,561	132,178	Mattias Jakobsson	GMI	1
6099	T1090	CS77293	SWE	556,575	132,386	Mattias Jakobsson	GMI	1
6100	T1110	CS77294	SWE	55.6	13.2	Mattias Jakobsson	GMI	1
6101	T1120	CS78045	SWE	55.6	13.2	Mattias Jakobsson	GMI	1
6102	T1130	CS77295	SWE	55.6	13.2	Mattias Jakobsson	GMI	1
6104	T1160	CS77296	SWE	55.7	13.2	Mattias Jakobsson	GMI	1
6105	T450	CS77297	SWE	557,967	131,211	Mattias Jakobsson	GMI	1
6106	T460	CS77298	SWE	557,931	131,186	Mattias Jakobsson	GMI	1
6107	T470	CS77299	SWE	557,942	131,222	Mattias Jakobsson	GMI	1
6108	T480	CS77300	SWE	557,989	131,206	Mattias Jakobsson	GMI	1
6109	T510	CS77301	SWE	557,936	131,233	Mattias Jakobsson	GMI	1
6111	T530	CS77302	SWE	557,989	131,219	Mattias Jakobsson	GMI	1
6112	T540	CS77303	SWE	557,967	131,044	Mattias Jakobsson	GMI	1
6113	T550	CS77304	SWE	558,078	131,028	Mattias Jakobsson	GMI	1
6114	T570	CS77305	SWE	558,097	131,342	Mattias Jakobsson	GMI	1
6115	T580	CS77306	SWE	55.8	131,367	Mattias Jakobsson	GMI	1
6118	T610	CS77307	SWE	55.7	13.2	Mattias Jakobsson	GMI	1
6119	T620	CS78060	SWE	55.7	13.2	Mattias Jakobsson	GMI	1
6122	T670	CS77308	SWE	558,364	133,075	Mattias Jakobsson	GMI	1
6123	T680	CS78064	SWE	558,369	133,033	Mattias Jakobsson	GMI	1
6124	T690	CS77309	SWE	558,378	133,092	Mattias Jakobsson	GMI	1
6125	T710	CS77310	SWE	558,403	133,106	Mattias Jakobsson	GMI	1
6126	T720	CS77311	SWE	558,411	133,047	Mattias Jakobsson	GMI	1
6128	T740	CS77313	SWE	558,397	132,881	Mattias Jakobsson	GMI	1
6131	T780	CS77315	SWE	558,369	133,181	Mattias Jakobsson	GMI	1
6132	T790	CS77316	SWE	558,386	133,186	Mattias Jakobsson	GMI	1
6133	T800	CS77317	SWE	558,364	132,906	Mattias Jakobsson	GMI	1
6134	T810	CS77318	SWE	558,383	132,906	Mattias Jakobsson	GMI	1
6136	T840	CS77319	SWE	559,336	135,519	Mattias Jakobsson	GMI	1

6137	T850	CS77320	SWE	559,419	135,603	Mattias Jakobsson	GMI	1
6138	T860	CS77321	SWE	559,403	135,511	Mattias Jakobsson	GMI	1
6140	T880	CS77322	SWE	559,392	135,539	Mattias Jakobsson	GMI	1
6141	T890	CS78079	SWE	559,414	135,542	Mattias Jakobsson	GMI	1
6142	T900	CS77323	SWE	559,428	135,558	Mattias Jakobsson	GMI	1
6145	T930	CS77324	SWE	559,497	135,533	Mattias Jakobsson	GMI	1
6148	T960	CS77325	SWE	559,319	135,508	Mattias Jakobsson	GMI	1
6149	T970	CS77326	SWE	559,281	135,481	Mattias Jakobsson	GMI	1
6150	T980	CS77327	SWE	559,261	135,319	Mattias Jakobsson	GMI	1
6151	T990	CS77328	SWE	556,528	132,244	Mattias Jakobsson	GMI	1
6153	TAA 03	CS77329	SWE	626,425	177,422	Mattias Jakobsson	GMI	1
6154	TAA 04	CS77330	SWE	626,422	177,406	Mattias Jakobsson	GMI	1
6163	TAA 14	CS77331	SWE	626,425	177,356	Mattias Jakobsson	GMI	1
6166	TAA 17	CS77332	SWE	626,425	177,372	Mattias Jakobsson	GMI	1
6169	TvÖD 01	CS77333	SWE	628,714	183,447	Mattias Jakobsson	GMI	1
6172	TvÖD 04	CS77335	SWE	628,717	183,436	Mattias Jakobsson	GMI	1
6173	TvÖD 05	CS77336	SWE	628,717	183,419	Mattias Jakobsson	GMI	1
6174	TvÖD 06	CS77337	SWE	628,719	183,422	Mattias Jakobsson	GMI	1
6177	TvÑL 03	CS77338	SWE	626,322	17.69	Mattias Jakobsson	GMI	1
6180	TvÑL 07	CS77339	SWE	626,322	176,906	Mattias Jakobsson	GMI	1
6184	TBvñ 01	CS77343	SWE	628,892	184,522	Mattias Jakobsson	GMI	1
6188	TDr-1	CS77345	SWE	557,683	141,386	Mattias Jakobsson	GMI	1
6189	TDr-2	CS77351	SWE	557,686	141,383	Mattias Jakobsson	GMI	1
6191	TDr-4	CS77352	SWE	557,689	141,375	Mattias Jakobsson	GMI	1
6192	TDr-5	CS77353	SWE	557,692	141,369	Mattias Jakobsson	GMI	1
6193	TDr-7	CS77354	SWE	557,694	141,347	Mattias Jakobsson	GMI	1
6194	TDr-8	CS77355	SWE	557,706	141,342	Mattias Jakobsson	GMI	1
6195	TDr-9	CS77356	SWE	557,708	141,342	Mattias Jakobsson	GMI	1
6201	TDr-16	CS77348	SWE	557,719	141,211	Mattias Jakobsson	GMI	1
6202	TDr-17	CS77349	SWE	557,717	141,206	Mattias Jakobsson	GMI	1
6203	TDr-18	CS77350	SWE	557,714	141,208	Mattias Jakobsson	GMI	1
6209	TEDEN 02	CS77358	SWE	628,836	181,842	Mattias Jakobsson	GMI	1
6210	TEDEN 03	CS77359	SWE	628,839	181,836	Mattias Jakobsson	GMI	1
6214	TFvÑ 04	CS78121	SWE	630,175	183,281	Mattias Jakobsson	GMI	1

6216	TFvÑ 06	CS77362	SWE	630,167	183,283	Mattias Jakobsson	GMI	1
6217	TFvÑ 07	CS77363	SWE	630,169	183,283	Mattias Jakobsson	GMI	1
6218	TFvÑ 08	CS77364	SWE	630,172	183,283	Mattias Jakobsson	GMI	1
6220	TGR 01	CS77365	SWE	62,806	181,896	Mattias Jakobsson	GMI	1
6221	TGR 02	CS77366	SWE	62,806	181,896	Mattias Jakobsson	GMI	1
6231	TNY 04	CS77368	SWE	62.96	182,844	Mattias Jakobsson	GMI	1
6235	TOM 01	CS77372	SWE	629,611	183,589	Mattias Jakobsson	GMI	1
6237	TOM 03	CS77373	SWE	629,619	18.35	Mattias Jakobsson	GMI	1
6238	TOM 04	CS77374	SWE	629,619	18.35	Mattias Jakobsson	GMI	1
6240	TOM 06	CS77375	SWE	629,622	18.35	Mattias Jakobsson	GMI	1
6241	TOM 07	CS77376	SWE	629,614	183,608	Mattias Jakobsson	GMI	1
6242	Tomegap -2	CS77377	SWE	55.7	13.2	Mattias Jakobsson	GMI	1
6243	Tottarp-2	CS77381	SWE	5,627,373	1,390,045	Magnus Nordborg	GMI	1
6244	TRvÑ 01	CS77384	SWE	629,169	184,728	Mattias Jakobsson	GMI	1
6252	TV-4	CS78771	SWE	555,796	143,336	Mattias Jakobsson	GMI	1
6255	TV-7	CS78772	SWE	555,796	14,334	Mattias Jakobsson	GMI	1
6258	TV-10	CS78767	SWE	555,796	143,336	Mattias Jakobsson	GMI	1
6268	TV-22	CS78768	SWE	555,796	143,336	Mattias Jakobsson	GMI	1
6276	TV-30	CS78769	SWE	555,796	143,336	Mattias Jakobsson	GMI	1
6284	TV-38	CS78770	SWE	555,796	143,336	Mattias Jakobsson	GMI	1
6296	Udul 1-11	CS78774	CZE	492,771	166,314	Jirina Relichov	Monsanto	1
6390	Udul 3-36	CS78775	CZE	492,771	166,314	Jirina Relichov	Monsanto	1
6396	Udul 4-9	CS78776	CZE	492,771	166,314	Jirina Relichov	Monsanto	1
6413	Ull3-4	CS78819	SWE	56.06	13.97	Magnus Nordborg	GMI	1
6424	Zdrl 1-23	CS78875	CZE	493,853	162,544	Jirina Relichov	Monsanto	1
6434	Zdrl 2-9	CS78877	CZE	493,853	162,544	Jirina Relichov	Monsanto	1
6445	Zdrl 2-21	CS78876	CZE	493,853	162,544	Jirina Relichov	Monsanto	1
6739	CSHL-15	CS79011	USA	408,585	-734,675	Catherine Weiss	MPI	1
6740	CSHL-17	CS79012	USA	408,585	-734,675	Catherine Weiss	MPI	2
6744	CSHL-5	CS76779	USA	408,585	-734,675	Catherine Weiss	Monsanto	2
6749	FM-10	CS79013	USA	424,489	-765,072	Michael Nachman	MPI	2
6750	FM-11	CS79014	USA	424,489	-765,072	Michael Nachman	MPI	2
6805	HS-12	CS79015	USA	42,373	-710,627	Toby Kellogg	MPI	2
6806	HS-17	CS79016	USA	42,373	-710,627	Toby Kellogg	MPI	2

6814	KNO-15	CS79017	USA	412,816	-86,621	Joy Bergelson	MPI	2
6830	Kz-13	CS76994	KAZ	49.5	73.1	Ihsan Al-Shehbaz	MPI	2
6897	Ag-0	CS76430	FRA	45	1.3	Albert Kranz	Salk	2
6898	An-1	CS76435	BEL	512,167	4.4	Albert Kranz	Salk	2
6900	Bil-5	CS76709	SWE	63,324	18,484	Magnus Nordborg	GMI	2
6901	Bil-7	CS76710	SWE	63,324	18,484	Magnus Nordborg	GMI	2
6903	Bor-4	CS76454	CZE	494,013	162,326	Jirina Relichov	Salk	2
6904	Br-0	CS76455	CZE	49.2	166,166	Albert Kranz	Salk	2
6907	CIBC-17	CS76770	UK	514,083	-0.6383	Mick Crawley	Salk	2
6908	CIBC-5	CS78894	UK	514,083	-0.6383	Mick Crawley	Salk	2
6911	Cvi-0	CS76789	CPV	151,111	-236,167	Albert Kranz	Salk	2
6913	Eden-2	CS76827	SWE	62,877	18,177	Magnus Nordborg	GMI	2
6915	Ei-2	CS76478	GER	50.3	6.3	Albert Kranz	Salk	2
6917	Fv§b-2	CS76850	SWE	630,165	183,174	Magnus Nordborg	GMI	2
6918	Fv§b-4	CS76851	SWE	630,165	183,174	Magnus Nordborg	GMI	2
6919	Ga-0	CS76490	GER	50.3	8	Albert Kranz	Salk	2
6920	Got-22	CS76884	GER	515,338	99,355	Gerhard Rvðbbelen	Salk	2
6922	Gu-0	CS76498	GER	50.3	8	Albert Kranz	Salk	2
6923	HR-10	CS76940	UK	514,083	-0.6383	Mick Crawley	Salk	2
6924	HR-5	CS76514	UK	514,083	-0.6383	Mick Crawley	Salk	2
6926	Kin-0	CS76527	USA	44.46	-85.37	Albert Kranz	Salk	2
6927	KNO-10	CS76973	USA	412,816	-86,621	Joy Bergelson	Salk,MPI	2
6929	Kondara	CS76532	TJK	38.48	68.49	Igor Vizir	Salk	2
6931	Kz-9	CS76537	KAZ	49.5	73.1	Ihsan Al-Shehbaz	Salk	2
6932	Ler-1	CS77021	GER	47,984	108,719	Eric Holub	MPI,GMI	2
6933	LL-0	CS77047	ESP	41.59	2.49	Albert Kranz	Salk	2
6938	Ms-0	CS76555	RUS	557,522	376,322	Albert Kranz	Salk	2
6940	Mz-0	CS76557	GER	50.3	8.3	Albert Kranz	Salk	2
6943	NFA-10	CS77126	UK	514,083	-0.6383	Mick Crawley	Salk	2
6944	NFA-8	CS78913	UK	514,083	-0.6383	Mick Crawley	Salk	2
6945	Nok-3	CS76562	NED	52.24	4.45	Albert Kranz	Salk	2
6951	Pu2-23	CS76579	CZE	49.42	16.36	Ivo Cetl	Salk	2
6956	Pu2-7	CS76580	CZE	49.42	16.36	Ivo Cetl	Salk	2
6957	Pu2-8	CS77192	CZE	49.42	16.36	Ivo Cetl	Salk	2
6958	Ra-0	CS76582	FRA	46	3.3	Albert Kranz	Salk	2
6959	Rennes-1	CS77210	FRA	48.5	-1.41	Gerhard Rvðbbelen	Salk	2
6960	Rennes-11	CS77211	FRA	48.5	-1.41	Gerhard Rvðbbelen	Salk	2

6961	Se-0	CS76597	ESP	383,333	-353,333	Albert Kranz	Salk	2
6963	Sorbo	CS78917	TJK	38.35	68.48	Igor Vizir	Salk	2
6966	Sq-1	CS77266	UK	514,083	-0.6383	Mick Crawley	Salk	2
6967	Sq-8	CS76604	UK	514,083	-0.6383	Mick Crawley	Salk	2
6968	Tamm-2	CS76610	FIN	60	23.5	Outi Savolainen	Salk,GMI	2
6970	Ts-1	CS76615	ESP	417,194	293,056	Albert Kranz	Salk	2
6971	Ts-5	CS77388	ESP	417,194	293,056	Albert Kranz	Salk	2
6973	Ull2-3	CS78817	SWE	560,648	139,707	Magnus Nordborg	GMI	2
6974	Ull2-5	CS78818	SWE	560,648	139,707	Magnus Nordborg	GMI	2
6975	Uod-1	CS76621	AUT	48.3	14.45	Marcus Koch	Salk	2
6976	Uod-7	CS78823	AUT	48.3	14.45	Marcus Koch	Salk	2
6979	Wei-0	CS76628	SUI	47.25	8.26	Alan Slusarenko	Salk	2
6981	Ws-2	CS78920	RUS	52.3	30	Kenneth Feldmann	MPI,Salk	2
6982	Wt-5	CS76632	GER	52.3	9.3	Albert Kranz	Salk	2
6984	Zdr-1	CS76635	CZE	493,853	162,544	Jirina Relichov	Salk	2
6986	Abd-0	CS76429	UK	571,539	-22,207	R. Mitchelso n	Salk	2
6987	Ak-1	CS76431	GER	480,683	762,551	Albert Kranz	Salk	2
6989	Alst-1	CS76432	UK	54.8	-24,333	Maarten Koornneef	Salk	2
6990	Amel-1	CS76434	NED	53,448	5.73	Maarten Koornneef	Salk	2
6992	Ang-0	CS76436	BEL	50.3	5.3	Albert Kranz	Salk	2
6997	Appt-1	CS76440	NED	518,333	55,833	Maarten Koornneef	Salk	2
7000	Aa-0	CS76428	GER	509,167	957,073	Albert Kranz	Salk	2
7002	Baa-1	CS76442	NED	513,333	6.1	Maarten Koornneef	Salk	2
7003	Bs-1	CS78888	SUI	47.5	7.5	Albert Kranz	Salk	2
7008	Benk-1	CS76447	NED	52	5,675	Maarten Koornneef	Salk	2
7013	Bd-0	CS76445	GER	524,584	13,287	Albert Kranz	Salk	2
7014	Ba-1	CS76441	UK	565,459	-479,821	Albert Kranz	Salk	2
7025	Bl-1	CS76450	ITA	445,041	113,396	Albert Kranz	Salk	2
7026	Boot-1	CS76452	UK	54.4	-32,667	Maarten Koornneef	Salk	2
7028	Bch-1	CS76444	GER	495,166	93,166	Albert Kranz	Salk	2
7031	Bsch-0	CS76457	GER	500,167	86,667	Albert Kranz	Salk	2
7033	Buckhorn Pass	CS76733	USA	413,599	-122,755	Angus Murphy	Salk,MPI	2
7036	Bu-0	CS78889	GER	50.5	9.5	Albert Kranz	MPI,Salk	2
7058	Bur-0	CS76734	IRL	53.08	- 907,555,55 6	Albert Kranz	Mott	2

7061	Cal-0	CS76460	UK	532,699	-164,293	Albert Kranz	Salk	2
7062	Ca-0	CS76459	GER	502,981	826,607	Albert Kranz	Salk	2
7064	Cnt-1	CS76467	UK	51.3	1.1	Eric Holub	Salk	2
7067	Ct-1	CS76786	ITA	37.3	15	Albert Kranz	Mott	2
7068	Cerv-1	CS76462	ITA	42	12.1	Maarten Koornneef	Salk	2
7071	Chat-1	CS76463	FRA	480,717	133,867	Maarten Koornneef	Salk	2
7072	Chi-0	CS76464	RUS	537,502	347,361	Albert Kranz	Salk	2
7077	Co-1	CS76468	POR	40.12	-8.25	Albert Kranz	Salk	2
7081	Co	CS78895	POR	402,077	-842,639	George Redei	Salk	2
7092	Com-1	CS76469	FRA	49,416	2,823	Maarten Koornneef	Salk	2
7094	Da-0	CS76791	GER	498,724	865,081	Albert Kranz	Monsanto	2
7096	Di-G	CS76472	FRA	473,239	504,278	Maarten Koornneef	Salk	2
7102	Do-0	CS76474	GER	507,224	82,372	Albert Kranz	Salk	2
7103	Dra-0	CS76476	CZE	494,167	162,667	Albert Kranz	Salk	2
7106	Dr-0	CS78897	GER	51,051	137,336	Albert Kranz	MPI,Salk	2
7107	Durh-1	CS76477	UK	547,761	-15,733	Maarten Koornneef	Salk	2
7109	Ema-1	CS76480	UK	51.3	0.5	Eric Holub	Salk	2
7111	Edi-0	CS76831	UK	559,494	-316,028	Albert Kranz	Mott	2
7117	El-0	CS76479	GER	515,105	968,253	Albert Kranz	Salk	2
7119	En-2	CS76481	GER	50	8.5	Albert Kranz	Salk	2
7120	En-D	CS76482	GER	50	8.5	Albert Kranz	Salk	2
7125	Er-0	CS78898	GER	495,955	110,087	Albert Kranz	MPI,Salk	2
7126	Es-0	CS76484	FIN	601,997	245,682	Albert Kranz	Salk	2
7127	Est	CS76485	EST	586,656	249,871	Brigitte Damm	Salk	2
7130	Et-0	CS76486	FRA	446,447	256,481	Albert Kranz	Salk	2
7133	Fr-2	CS76489	GER	501,102	86,822	Albert Kranz	Salk	2
7143	Gel-1	CS76492	NED	510,167	586,667	Maarten Koornneef	Salk	2
7147	Gie-0	CS76493	GER	50,584	867,825	Albert Kranz	Salk	2
7158	Gr-5	CS76885	AUT	47	15.5	Albert Kranz	Monsanto	2
7160	Gre-0	CS76497	USA	43,178	-852,532	Albert Kranz	Salk	2
7161	Gd-1	CS76491	GER	53.5	10.5	Albert Kranz	Salk	2
7162	Hs-0	CS76515	GER	52.24	9.44	Albert Kranz	Salk	2
7163	Ha-0	CS76500	GER	523,721	973,569	Albert Kranz	Salk	2
7165	Hn-0	CS76513	GER	513,472	828,844	Albert Kranz	Salk	2
7169	Hh-0	CS76512	GER	544,175	988,682	Albert Kranz	Salk	2

7177	Jm-0	CS76520	CZE	49	15	Albert Kranz	Salk	2
7181	Je-0	CS76518	GER	50,927	11,587	Albert Kranz	Salk	2
7183	Kas-1	CS79018	IND	35	77	Shauna Somerville	Salk	2
7186	Kn-0	CS76969	LTU	548,969	238,924	Albert Kranz	Mott	2
7192	Kil-0	CS76526	UK	556,395	-566,364	Albert Kranz	Salk	2
7199	Kl-5	CS76528	GER	50.95	69,666	Albert Kranz	Salk	2
7202	Kb-0	CS76524	GER	501,797	850,861	Albert Kranz	Salk	2
7203	Krot-0	CS76534	GER	49,631	115,722	Thomas Mitchell-Olds	Salk	2
7207	Kyoto	CS76535	JPN	350,085	135,752	Hirokazu Tsukaya	Salk	2
7208	Lan-0	CS76539	UK	556,739	-378,181	Albert Kranz	Salk	2
7209	La-0	CS76538	POL	527,333	152,333	Albert Kranz	Salk	2
7213	Ler-0	CS77020	GER	47,984	108,719	Maarten Koornneef	Mott	2
7217	Lm-2	CS76545	FRA	48	0.5	Albert Kranz	Salk	2
7218	Le-0	CS76540	NED	521,611	449,015	Albert Kranz	Salk	2
7223	Li-2:1	CS76541	GER	503,833	80,666	Albert Kranz	Salk	2
7231	Li-7	CS77035	GER	503,833	80,666	Albert Kranz	Monsanto	2
7236	Litva	CS76543	LTU			Igor Vizir	Salk	2
7244	Mnz-0	CS76552	GER	50,001	826,664	Albert Kranz	Salk	2
7248	Mv-0	CS76556	USA	413,923	-706,652	Albert Kranz	Salk	2
7250	Me-0	CS76549	GER	519,183	101,138	Albert Kranz	Salk	2
7255	Mh-0	CS76550	POL	50.95	20.5	Albert Kranz	Salk	2
7258	Nw-0	CS76564	GER	50.5	8.5	Albert Kranz	Salk	2
7268	Np-0	CS76563	GER	526,969	10,981	Albert Kranz	Salk	2
7273	No-0	CS77128	GER	510,581	132,995	Albert Kranz	Mott	2
7276	Ob-0	CS76566	GER	50.2	85,833	Albert Kranz	Salk	2
7280	Old-1	CS76567	GER	531,667	8.2	Albert Kranz	Salk	2
7282	Or-0	CS76568	GER	503,827	801,161	Albert Kranz	Salk	2
7287	Ove-0	CS76569	GER	533,422	842,255	Albert Kranz	Salk	2
7288	Oy-0	CS77156	NOR	60,385,543	6,193,019	Albert Kranz	Mott	2
7296	Petergof	CS77170	RUS	59	29	Igor Vizir	Monsanto	2
7298	Pi-0	CS76572	AUT	47.04	10.51	Albert Kranz	Salk	2
7305	Pt-0	CS78915	GER	53,476	106,065	Albert Kranz	Salk	2
7306	Pog-0	CS76576	CAN	492,655	-123,206	Albert Kranz	Salk	2
7307	Pn-0	CS77182	FRA	480,653	-296,591	Albert Kranz	Monsanto	2
7314	Ragl-1	CS76583	UK	543,512	-341,697	Maarten Koornneef	Salk	2

7316	Rhen-1	CS78916	NED	519,667	556,667	Maarten Koornneef	Salk	2
7319	Rome-1	CS76590	ITA	42	12.1	Maarten Koornneef	Salk	2
7320	Rou-0	CS76591	FRA	494,424	109,849	Albert Kranz	Salk	2
7322	Rsch-4	CS77222	RUS	56.3	34	Albert Kranz	Mott	2
7323	Rubezhnoe-1	CS76594	UKR	49	38.28	Igor Vizir	Salk	2
7327	Sf-1	CS77246	ESP	417,833	303,333	F. Laibach	Salk	2
7328	Sf-2	CS77247	ESP	417,833	303,333	Albert Kranz	Mott	2
7332	Seattle-0	CS76598	USA	47	-122.2	Rick Amasino	Salk	2
7333	Sei-0	CS76599	ITA	465,438	115,614	Albert Kranz	Salk	2
7337	Si-0	CS76601	GER	508,738	802,341	Albert Kranz	Salk	2
7342	Su-0	CS76606	UK	536,473	-300,733	Albert Kranz	Salk	2
7343	Sp-0	CS76603	GER	525,339	13,181	Albert Kranz	Salk	2
7344	Sg-1	CS76600	GER	476,667	9.5	Albert Kranz	Salk	2
7346	Sten-0	CS77277	GER	526,058	118,558	Albert Kranz	Monsanto	2
7347	Stw-0	CS76605	RUS	52	36	Albert Kranz	Salk	2
7349	Ta-0	CS76608	CZE	49.5	14.5	Albert Kranz	Salk	2
7350	Tac-0	CS76609	USA	472,413	-122,459	Thomas Mitchell-Olds	Salk	2
7353	Tha-1	CS76611	NED	52.08	4.3	Maarten Koornneef	Salk	2
7354	Ting-1	CS76612	SWE	56.5	14.9	Maarten Koornneef	Salk	2
7356	Tol-0	CS76614	USA	416,639	-835,553	Scott Leisner	Salk,MPI	2
7358	Tol-2	CS79019	USA	416,639	-835,553	Scott Leisner	MPI	2
7359	Tol-3	CS79020	USA	416,639	-835,553	Scott Leisner	MPI	2
7372	Tscha-1	CS76616	AUT	470,748	99,042	Maarten Koornneef	Salk	2
7373	Tsu-0	CS77389	JPN	34.43	136.31	Albert Kranz	Mott	2
7378	Uk-1	CS76620	GER	480,333	77,667	Albert Kranz	Salk	2
7382	Utrecht	CS76622	NED	520,918	51,145	Viola Willemsen	Salk	2
7383	Van-0	CS76623	CAN	492,655	-123,206	Albert Kranz	Salk	2
7384	Ven-1	CS76624	NED	520,333	5.55	Maarten Koornneef	Salk	2
7387	Vind-1	CS76625	UK	549,902	-23,671	Maarten Koornneef	Salk	2
7394	Wa-1	CS76626	POL	52.3	21	Albert Kranz	Salk	2
7396	Ws-0.2	CS78857	RUS	52.3	30	Albert Kranz	Mott	2
7404	Wc-1	CS76627	GER	52.6	100,667	Albert Kranz	Salk	2
7411	Wl-0	CS76630	GER	479,299	108,134	Albert Kranz	Salk	2
7413	Wil-2	CS78856	LTU	546,833	253,167	Albert Kranz	Mott	2

7416	Yo-0	CS76633	USA	37.45	-119.35	Albert Kranz	Salk	2
7417	Zu-0	CS78880	SUI	473,667	8.55	Albert Kranz	Mott	2
7418	Zu-1	CS78881	SUI	473,667	8.55	Albert Kranz	Monsanto	2
7419	Db-1	CS76471	GER	503,058	832,213	Albert Kranz	Salk	2
7424	JI-3	CS76519	CZE	49.2	166,166	Albert Kranz	Salk	2
7430	Nc-1	CS76559	FRA	486,167	6.25	Albert Kranz	Salk	2
7460	Da(1)-12	CS76470	CZE			Igor Vizir	Salk	2
7461	H55	CS76897	CZE	49	15	Igor Vizir	Salk	2
7471	RLD-1	CS76588	UNK			Maarten Koornneef	Salk	2
7475	KEN	CS79021	USA	41,767	-72,677	Massimo Pigliucci	MPI	2
7477	WAR	CS78853	USA	417,302	-712,825	Massimo Pigliucci	Monsanto	2
7514	RRS-7	CS76593	USA	415,609	-864,251	Joy Bergelson	Salk	2
7515	RRS-10	CS76592	USA	415,609	-864,251	Joy Bergelson	MPI,Salk	2
7516	Vv•r2-1	CS78830	SWE	55.58	14,334	Magnus Nordborg	GMI	2
7517	Vv•r2-6	CS78831	SWE	55.58	14,334	Magnus Nordborg	GMI	2
7520	Lp2-2	CS76546	CZE	49.38	16.81	Ivo Cetl	Salk	2
7521	Lp2-6	CS77052	CZE	49.38	16.81	Ivo Cetl	Salk	2
7523	Pna-17	CS76575	USA	420,945	-863,253	Joy Bergelson	Salk	2
7525	Rmx-A180	CS77218	USA	42,036	-86,511	Joy Bergelson	Salk	2
7529	627RMX-1MN4	CS79022	USA	420,333	-865,128	Justin Borevitz	MPI	2
7530	627RMX-1MN5	CS79023	USA	420,333	-865,128	Justin Borevitz	MPI	2
7566	627ME-13Y1	CS78367	USA	42,093	-86,359	Justin Borevitz	MPI	2
7568	627ME-1MI1	CS79024	USA	42,093	-86,359	Justin Borevitz	MPI	2
7717	KNO1.37	CS76972	USA	41,273	-86,625	Megan Dunning	Monsanto	2
7757	KNO2.41	CS79025	USA	41,273	-86,625	Megan Dunning	MPI	2
7767	KNO2.54	CS79026	USA	41,273	-86,625	Megan Dunning	MPI	2
7917	PNA3.10	CS77183	USA	420,945	-863,253	Megan Dunning	Monsanto	2
7947	PNA3.40	CS77184	USA	420,945	-863,253	Megan Dunning	Monsanto	2
8037	PT1.52	CS79027	USA	413,423	-867,368	Megan Dunning	MPI	2
8057	PT1.85	CS79028	USA	413,423	-867,368	Megan Dunning	MPI	2
8077	PT2.21	CS77191	USA	413,423	-867,368	Megan Dunning	Monsanto	2
8132	RMX3.22	CS77219	USA	42,036	-86,511	Megan Dunning	Monsanto	2
8171	RMX4.118	CS79029	USA	42,036	-86,511	Megan Dunning	MPI	2
8214	Gy-0	CS78901	FRA	49	2	Albert Kranz	Salk	2
8222	Lis-2	CS77043	SWE	560,328	14,775	Magnus Nordborg	GMI	2
8227	THVñ 03	CS77367	SWE	627,989	179,103	Mattias Jakobsson	GMI	2

8230	Algutsrum	CS76657	SWE	56.68	16.5	Magnus Nordborg	GMI	2
8231	BrVð1-6	CS76726	SWE	56.3	16	Magnus Nordborg	GMI	2
8233	Dem-4	CS76794	USA	411,876	-871,923	Joy Bergelson	Salk	2
8234	Gul1-2	CS76896	SWE	564,606	158,127	Magnus Nordborg	GMI,Salk	2
8235	Hod	CS76924	CZE	48.8	17.1	Jirina Relichov	Salk	2
8236	HSm	CS76941	CZE	49.33	15.76	Jirina Relichov	Salk	2
8237	KV§vlinge-1	CS76964	SWE	55.8	13.1	Torbjorn Sall	GMI	2
8238	Kent	CS76967	UK	51.15	0.4	Magnus Nordborg	Salk	2
8239	PHW-3	CS76976	GER	51	7	Paul Williams	Salk	2
8240	Kulturen-1	CS76987	SWE	55,705	13,196	Torbjorn Sall	GMI	2
8241	Liarum	CS77038	SWE	559,473	13,821	Torbjorn Sall	GMI	2
8242	LillVð-1	CS77039	SWE	561,494	157,884	Magnus Nordborg	GMI	2
8243	PHW-2	CS77173	ITA	437,703	112,547	Paul Williams	Salk	2
8244	PHW-34	CS77174	FRA	486,103	23,086	Paul Williams	Salk	2
8246	NC-6	CS77124	USA	35	-79.18	Joy Bergelson	Salk	2
8247	San-2	CS77233	SWE	56.07	13.74	Magnus Nordborg	GMI	2
8249	Vimmerby	CS78845	SWE	57.7	15.8	Magnus Nordborg	GMI	2
8256	BV•1-2	CS76676	SWE	56.4	12.9	Magnus Nordborg	GMI	2
8259	BV•5-1	CS76678	SWE	56.4	12.9	Magnus Nordborg	GMI	2
8264	Bla-1	CS76451	ESP	416,833	2.8	Albert Kranz	Salk	2
8283	Dra3-1	CS76811	SWE	55.76	14.12	Magnus Nordborg	GMI	2
8284	DraII-1	CS76813	CZE	494,112	162,815	Jirina Relichov	Salk	2
8285	DraIII-1	CS76815	CZE	494,112	162,815	Jirina Relichov	Salk	2
8290	En-1	CS76841	GER	50	8.5	Albert Kranz	Salk	2
8297	Ge-0	CS76875	SUI	46.5	6.08	Albert Kranz	Salk	2
8306	Hov4-1	CS76936	SWE	56.1	13.74	Magnus Nordborg	GMI	2
8307	Hovdala-6	CS76938	SWE	56.1	13.74	Torbjorn Sall	GMI	2
8311	In-0	CS78903	AUT	47.5	11.5	Albert Kranz	Salk	2
8312	Is-0	CS78904	GER	50.5	7.5	Albert Kranz	MPI,Salk	2
8326	Lis-1	CS77042	SWE	560,328	14,775	Magnus Nordborg	GMI	2
8334	Lu-1	CS77056	SWE	55.71	13.2	Albert Kranz	GMI	2
8335	Lund	CS77060	SWE	55.71	13.2	Magnus Nordborg	GMI	2
8337	Mir-0	CS76551	ITA	44	12.37	Albert Kranz	Salk	2
8343	Na-1	CS76558	FRA	47.5	1.5	Albert Kranz	Salk	2

8351	Ost-0	CS77154	SWE	60.25	18.37	Albert Kranz	GMI	2
8354	Per-1	CS76571	RUS	58	563,167	Albert Kranz	Salk	2
8357	Pla-0	CS76573	ESP	41.5	2.25	Albert Kranz	Salk	2
8365	Rak-2	CS77201	CZE	49	16	Albert Kranz	Salk	2
8366	Rd-0	CS76584	GER	50.5	8.5	Albert Kranz	Salk	2
8369	Rev-1	CS77214	SWE	556,942	134,504	Magnus Nordborg	GMI	2
8376	Sanna-2	CS77234	SWE	62.69	18	Magnus Nordborg	GMI	2
8386	Sr:5	CS77268	SWE	58.9	11.2	Ivo Cetl	GMI	2
8419	Wil-1	CS78855	LTU	546,833	253,167	Albert Kranz	Salk	2
8420	Kelsterbach-4	CS76525	GER	500,667	85,333	Paul Williams	Salk	2
8422	FjV§1-1	CS76859	SWE	56.06	14.29	Magnus Nordborg	GMI	2
8424	Kas-2	CS78905	IND	35	77	Shauna Somerville	Salk	2
8426	Ull1-1	CS78816	SWE	56.06	13.97	Magnus Nordborg	GMI	2
8427	Ull2-13	CS78552	SWE	560,648	139,707	Magnus Nordborg	GMI	2
8483	LP3413.53	CS79031	USA	416,862	-868,513	Justin Borevitz	MPI	2
8699	328PNA062	CS79032	USA	420,945	-863,253	Justin Borevitz	MPI	2
9027	RMX413.85	CS79034	USA	42,036	-86,511	Justin Borevitz	MPI	2
9057	VinslVðv	CS78847	SWE	56.1	139,167	Torbjorn Sall	GMI	2
9058	VV§stervik	CS78834	SWE	57.75	166,333	Magnus Nordborg	GMI	2
9067	Xan-3	CS78860	AZE	386,536	487,992	James Beck	Monsanto	2
9069	Xan-5	CS78861	AZE	386,536	487,992	James Beck	Monsanto	2
9070	Xan-6	CS78862	AZE	386,536	487,992	James Beck	Monsanto	2
9075	Lerik1-4	CS77023	AZE	387,406	486,131	James Beck	Monsanto	2
9078	Lerik1-7	CS77024	AZE	387,406	486,131	James Beck	Monsanto	2
9079	Lerik2-1	CS77025	AZE	387,833	485,517	James Beck	Monsanto	2
9081	Lerik2-3	CS77026	AZE	387,833	485,517	James Beck	Monsanto	2
9084	Lerik2-6	CS77028	AZE	387,833	485,517	James Beck	Monsanto	2
9085	Lerik2-7	CS77029	AZE	387,833	485,517	James Beck	Monsanto	2
9089	Nar-3	CS77119	AZE	389,522	48,925	James Beck	Monsanto	2
9091	Nar-5	CS77121	AZE	389,522	48,925	James Beck	Monsanto	2
9095	Istisu-5	CS76950	AZE	389,786	485,594	James Beck	Monsanto	2
9099	Istisu-9	CS76953	AZE	389,786	485,594	James Beck	Monsanto	2
9100	Lag1-2	CS76998	GEO	418,296	462,831	James Beck	Monsanto	2
9102	Lag1-4	CS76999	GEO	418,296	462,831	James Beck	Monsanto	2
9103	Lag1-5	CS77000	GEO	418,296	462,831	James Beck	Monsanto	2

9104	Lag1-6	CS77001	GEO	418,296	462,831	James Beck	Monsanto	2
9105	Lag1-7	CS77002	GEO	418,296	462,831	James Beck	Monsanto	2
9106	Lag1-8	CS77003	GEO	418,296	462,831	James Beck	Monsanto	2
9111	Lag2-4	CS77005	GEO	418,296	462,831	James Beck	Monsanto	2
9113	Lag2-6	CS77006	GEO	418,296	462,831	James Beck	Monsanto	2
9114	Lag2-7	CS77007	GEO	418,296	462,831	James Beck	Monsanto	2
9121	Bak-5	CS76685	GEO	417,942	434,767	James Beck	Monsanto	2
9125	Geg-14	CS76876	ARM	401,408	448,203	James Beck	Monsanto	2
9128	Yeg-2	CS78864	ARM	398,692	453,622	James Beck	Monsanto	2
9130	Yeg-4	CS78865	ARM	398,692	453,622	James Beck	Monsanto	2
9131	Yeg-5	CS78866	ARM	398,692	453,622	James Beck	Monsanto	2
9133	Yeg-7	CS78867	ARM	398,692	453,622	James Beck	Monsanto	2
9134	Yeg-8	CS78868	ARM	398,692	453,622	James Beck	Monsanto	2
9298	Edi-1	CS76832	UK	559,681	-321,833	James Beck	Monsanto	2
9314	Gol-2	CS76883	UK	579,672	-396,722	James Beck	Monsanto	3
9321	√Ödal 1	CS76643	SWE	628,622	18,336	Alison Anastasio	GMI	3
9323	√Ödal 3	CS76644	SWE	628,622	18,336	Alison Anastasio	GMI	3
9332	Bar 1	CS76688	SWE	628,698	18,381	Alison Anastasio	GMI	3
9336	Bv∂n 1	CS76715	SWE	628,794	184,473	Alison Anastasio	GMI	3
9339	Bv∂t 1	CS76720	SWE	577,133	150,689	Alison Anastasio	GMI	3
9343	Dju-1	CS78896	SWE	573,089	181,512	Alison Anastasio	GMI	3
9352	Dv∂d 2	CS76797	SWE	572,608	163,675	Alison Anastasio	GMI	3
9353	Dv∂d 3	CS76798	SWE	572,608	163,675	Alison Anastasio	GMI	3
9363	EdJ 2	CS76833	SWE	629,147	184,045	Alison Anastasio	GMI	3
9369	EkS 2	CS76837	SWE	576,781	149,986	Alison Anastasio	GMI	3
9370	EkS 3	CS79035	SWE	576,781	149,986	Alison Anastasio	GMI	3
9371	Fv§L 1	CS76852	SWE	63,016	183,175	Alison Anastasio	GMI	3
9380	FlyA 3	CS76865	SWE	557,488	133,742	Alison Anastasio	GMI	3
9381	Fri 1	CS76868	SWE	558,106	142,091	Alison Anastasio	GMI	3
9382	Fri 2	CS76869	SWE	558,106	142,091	Alison Anastasio	GMI	3
9383	Fri 3	CS76870	SWE	558,106	142,091	Alison Anastasio	GMI	3
9386	Grv∂n 12	CS76891	SWE	62,806	181,896	Alison Anastasio	GMI	3
9388	Grv∂n 14	CS76892	SWE	62,806	181,896	Alison Anastasio	GMI	3
9390	Hadd-1	CS78659	SWE	573,263	158,979	Alison Anastasio	GMI	3

9391	Hadd-2	CS76905	SWE	573,263	158,979	Alison Anastasio	GMI	3
9392	Hadd-3	CS76906	SWE	573,263	158,979	Alison Anastasio	GMI	3
9394	Hag-2	CS76907	SWE	565,804	164,063	Alison Anastasio	GMI	3
9395	Hal-1	CS76908	SWE	575,089	150,105	Alison Anastasio	GMI	3
9399	Hamm-1	CS76910	SWE	554,234	139,905	Alison Anastasio	GMI	3
9402	Hel-3	CS76918	SWE	578,765	148,549	Alison Anastasio	GMI	3
9404	HolA-1 1	CS76926	SWE	557,491	13,399	Alison Anastasio	GMI	3
9405	HolA-1 2	CS76927	SWE	557,491	13,399	Alison Anastasio	GMI	3
9407	HolA-2 2	CS76928	SWE	557,491	13,399	Alison Anastasio	GMI	3
9408	Kal 1	CS76959	SWE	56,047	139,519	Alison Anastasio	GMI	3
9409	Kia 1	CS76968	SWE	560,573	14,302	Alison Anastasio	GMI	3
9412	Kor 3	CS76981	SWE	572,746	161,494	Alison Anastasio	GMI	3
9413	Kor 4	CS76982	SWE	572,746	161,494	Alison Anastasio	GMI	3
9416	Kru-3	CS76986	SWE	577,215	183,837	Alison Anastasio	GMI	3
9421	Lan 1	CS77009	SWE	559,745	143,997	Alison Anastasio	GMI	3
9433	Nyl 13	CS77135	SWE	629,513	182,763	Alison Anastasio	GMI	3
9436	Puk-1	CS77194	SWE	561,633	146,806	Alison Anastasio	GMI	3
9437	Puk-2	CS77195	SWE	561,633	146,806	Alison Anastasio	GMI	3
9442	Sim-1	CS77250	SWE	555,678	143,398	Alison Anastasio	GMI	3
9450	Spro 1	CS77263	SWE	572,545	182,109	Alison Anastasio	GMI	3
9451	Spro 2	CS77264	SWE	572,545	182,109	Alison Anastasio	GMI	3
9452	Spro 3	CS77265	SWE	572,545	182,109	Alison Anastasio	GMI	3
9453	Stenk-2	CS77274	SWE	578,009	185,162	Alison Anastasio	GMI	3
9454	Stenk-3	CS77275	SWE	578,009	185,162	Alison Anastasio	GMI	3
9455	Stenk-4	CS77276	SWE	578,009	185,162	Alison Anastasio	GMI	3
9470	Tur-4	CS77399	SWE	576,511	148,043	Alison Anastasio	GMI	3
9471	Ull-A-1	CS78820	SWE	560,648	139,707	Alison Anastasio	GMI	3
9476	VV•rA 1	CS78832	SWE	555,796	143,336	Alison Anastasio	GMI	3
9481	Yst-1	CS78869	SWE	554,242	138,484	Alison Anastasio	GMI	3
9503	11C1	CS76640	UK	558,877	-321,072	Andrew Hudson	Monsanto	3
9506	IP-Alo-0	CS76662	POR	40.11	-7.47	Carlos Alonso-Blanco	Monsanto	3
9507	IP-Coa-0	CS76775	POR	38.45	-7.5	Carlos Alonso-Blanco	Monsanto	3
9508	IP-Mos-1	CS77108	POR	40.04	-7.11	Carlos Alonso-Blanco	Monsanto	3

9509	IP-Reg-0	CS77207	POR	39.29	-7.4	Carlos Alonso-Blanco	Monsanto	3
9510	IP-Rei-0	CS77208	POR	38.75	-7.59	Carlos Alonso-Blanco	Monsanto	3
9511	IP-Vav-0	CS78835	POR	38.53	-8.02	Carlos Alonso-Blanco	Monsanto	3
9512	IP-Vid-1	CS78842	POR	38.22	-7.84	Carlos Alonso-Blanco	Monsanto	3
9513	IP-Adc-5	CS76646	ESP	38.77	-4.07	Carlos Alonso-Blanco	Monsanto	3
9514	IP-Adm-0	CS76647	ESP	39.15	-4.54	Carlos Alonso-Blanco	Monsanto	3
9515	IP-Ala-0	CS76650	ESP	39.72	-6.89	Carlos Alonso-Blanco	Monsanto	3
9517	IP-All-0	CS76659	ESP	42.19	-7.8	Carlos Alonso-Blanco	Monsanto	3
9518	IP-Alm-0	CS76660	ESP	39.88	-0.36	Carlos Alonso-Blanco	Monsanto	3
9519	IP-Ang-0	CS78886	ESP	41.94	2.64	Xavier Picv≥	Monsanto	3
9521	IP-Bar-1	CS76689	ESP	41.43	2.13	Xavier Picv≥	Monsanto	3
9522	IP-Bea-0	CS76695	ESP	36.52	-5.27	Carlos Alonso-Blanco	Monsanto	3
9523	IP-Ben-0	CS76700	ESP	38.37	-2.66	Carlos Alonso-Blanco	Monsanto	3
9524	IP-Ber-0	CS78887	ESP	42.52	-0.56	Carlos Alonso-Blanco	Monsanto	3
9525	IP-Bis-0	CS76711	ESP	42.49	0.54	Xavier Picv≥	Monsanto	3
9526	IP-Cab-3	CS76738	ESP	41.54	2.39	Xavier Picv≥	Monsanto	3
9527	IP-Cad-0	CS76739	ESP	40.37	-5.74	Carlos Alonso-Blanco	Monsanto	3
9528	IP-Cal-0	CS78890	ESP	40.94	-1.37	Carlos Alonso-Blanco	Monsanto	3
9529	IP-Cap-1	CS76741	ESP	36.97	-3.36	Carlos Alonso-Blanco	Monsanto	3
9530	IP-Car-1	CS76742	ESP	38.25	-4.32	Carlos Alonso-Blanco	Monsanto	3
9531	IP-Cdc-3	CS76761	ESP	41.21	-4.54	Carlos Alonso-Blanco	Monsanto	3
9532	IP-Cdo-0	CS76762	ESP	42.23	-4.64	Carlos Alonso-Blanco	Monsanto	3
9533	IP-Cem-0	CS76763	ESP	41.15	-4.32	Carlos Alonso-Blanco	Monsanto	3
9534	IP-Cmo-3	CS76774	ESP	40.05	-4.65	Carlos Alonso-Blanco	Monsanto	3
9535	IP-Coc-1	CS76776	ESP	42.31	3.19	Xavier Picv≥	Monsanto	3

9536	IP-Cor-0	CS76782	ESP	40.83	-2	Carlos Alonso-Blanco	Monsanto	3
9537	IP-Cum-1	CS76787	ESP	38.07	-6.66	Carlos Alonso-Blanco	Monsanto	3
9539	IP-Deh-1	CS76793	ESP	40.29	-6.67	Carlos Alonso-Blanco	Monsanto	3
9540	IP-Elb-0	CS76838	ESP	41.81	2.34	Xavier Picv≥	Monsanto	3
9541	IP-Fue-2	CS76871	ESP	38.26	-5.42	Carlos Alonso-Blanco	Monsanto	3
9542	IP-Fun-0	CS76872	ESP	40.79	-4.05	Carlos Alonso-Blanco	Monsanto	3
9543	IP-Gra-0	CS76886	ESP	36.77	-5.39	Xavier Picv≥	Monsanto	3
9544	IP-Gua-1	CS76894	ESP	39.4	-5.33	Carlos Alonso-Blanco	Monsanto	3
9545	IP-Her-12	CS76920	ESP	39.4	-5.78	Carlos Alonso-Blanco	Monsanto	3
9546	IP-Hom-4	CS76929	ESP	40.82	-1.68	Carlos Alonso-Blanco	Monsanto	3
9547	IP-Hor-0	CS76930	ESP	41.67	2.62	Xavier Picv≥	Monsanto	3
9548	IP-Hoy-0	CS76939	ESP	40.4	-5	Carlos Alonso-Blanco	Monsanto	3
9549	IP-Hum-2	CS76943	ESP	42.23	-3.69	Carlos Alonso-Blanco	Monsanto	3
9550	IP-Iso-4	CS76946	ESP	43.05	-5.37	Carlos Alonso-Blanco	Monsanto	3
9551	IP-Jim-1	CS76955	ESP	42.28	-5.92	Carlos Alonso-Blanco	Monsanto	3
9552	IP-Lab-7	CS76995	ESP	40.87	-4.5	Carlos Alonso-Blanco	Monsanto	3
9553	IP-Ldd-0	CS77012	ESP	41.58	-4.71	Carlos Alonso-Blanco	Monsanto	3
9554	IP-Lso-0	CS77055	ESP	38.86	-3.16	Carlos Alonso-Blanco	Monsanto	3
9555	IP-Mar-1	CS77068	ESP	39.58	-3.93	Carlos Alonso-Blanco	Monsanto	3
9556	IP-Men-2	CS77081	ESP	39.66	-4.34	Xavier Picv≥	Monsanto	3
9557	IP-Moa-0	CS77102	ESP	42.46	0.7	Carlos Alonso-Blanco	Monsanto	3
9558	IP-Moc-11	CS77103	ESP	41.57	-5.64	Carlos Alonso-Blanco	Monsanto	3
9559	IP-Mon-5	CS77107	ESP	38.06	-4.38	Carlos Alonso-Blanco	Monsanto	3
9560	IP-Mot-0	CS77109	ESP	38.19	-6.24	Carlos Alonso-Blanco	Monsanto	3

9561	IP-Mun-0	CS77114	ESP	40.71	-5.04	Carlos Alonso-Blanco	Monsanto	3
9562	IP-Mur-0	CS77115	ESP	41.67	2	Xavier Picv≥	Monsanto	3
9564	IP-Nog-17	CS77129	ESP	40.45	-1.6	Xavier Picv≥	Monsanto	3
9565	IP-Orb-10	CS77152	ESP	42.97	-1.23	Carlos Alonso-Blanco	Monsanto	3
9567	IP-Pal-0	CS77159	ESP	42.34	1.3	Xavier Picv≥	Monsanto	3
9568	IP-Pan-0	CS77160	ESP	42.76	-0.23	Carlos Alonso-Blanco	Monsanto	3
9569	IP-Pds-1	CS77166	ESP	42.87	-6.45	Carlos Alonso-Blanco	Monsanto	3
9571	IP-Pro-0	CS78914	ESP	43.28	-6.01	Carlos Alonso-Blanco	Monsanto	3
9573	IP-Rds-0	CS77206	ESP	41.86	2.99	Xavier Picv≥	Monsanto	3
9574	IP-Rel-0	CS77209	ESP	38.6	-2.7	Carlos Alonso-Blanco	Monsanto	3
9576	IP-Rev-0	CS77213	ESP	40.86	-4.11	Carlos Alonso-Blanco	Monsanto	3
9577	IP-Ria-0	CS77216	ESP	42.34	2.17	Xavier Picv≥	Monsanto	3
9578	IP-Sac-0	CS77229	ESP	42.13	-6.7	Carlos Alonso-Blanco	Monsanto	3
9579	IP-San-10	CS77232	ESP	38.33	-3.51	Carlos Alonso-Blanco	Monsanto	3
9581	IP-Sdv-3	CS77242	ESP	42.84	-5.12	Carlos Alonso-Blanco	Monsanto	3
9582	IP-Ses-0	CS77244	ESP	41.48	-1.63	Carlos Alonso-Blanco	Monsanto	3
9583	IP-Sne-0	CS77258	ESP	37.09	-3.38	Carlos Alonso-Blanco	Monsanto	3
9584	IP-Stp-0	CS77283	ESP	41.19	-3.58	Carlos Alonso-Blanco	Monsanto	3
9585	IP-Svi-0	CS77287	ESP	43.4	-7.39	Carlos Alonso-Blanco	Monsanto	3
9586	IP-Tam-0	CS77340	ESP	41.03	-3.27	Carlos Alonso-Blanco	Monsanto	3
9587	IP-Tdc-0	CS77344	ESP	41.5	-1.88	Carlos Alonso-Blanco	Monsanto	3
9588	IP-Tol-7	CS77371	ESP	42.11	0.6	Xavier Picv≥	Monsanto	3
9589	IP-Tor-1	CS77378	ESP	41.6	-2.83	Carlos Alonso-Blanco	Monsanto	3
9590	IP-Trs-0	CS77387	ESP	43.37	-5.49	Carlos Alonso-Blanco	Monsanto	3
9591	IP-Vad-0	CS78826	ESP	42.86	-3.59	Carlos Alonso-Blanco	Monsanto	3

9593	IP-Vaz-0	CS78836	ESP	42.26	-2.99	Carlos Alonso-Blanco	Monsanto	3
9594	IP-Vdm-0	CS78837	ESP	42.04	1.01	Carlos Alonso-Blanco	Monsanto	3
9595	IP-Vdt-0	CS78838	ESP	40.89	-5.5	Carlos Alonso-Blanco	Monsanto	3
9596	IP-Ver-5	CS78841	ESP	41.95	-7.45	Carlos Alonso-Blanco	Monsanto	3
9597	IP-Vig-1	CS78843	ESP	42.31	-2.53	Carlos Alonso-Blanco	Monsanto	3
9598	IP-Vim-0	CS78844	ESP	41.88	-6.51	Carlos Alonso-Blanco	Monsanto	3
9599	IP-Vin-0	CS78846	ESP	42.8	-5.77	Carlos Alonso-Blanco	Monsanto	3
9601	IP-Voz-0	CS78849	ESP	41.85	-1.88	Carlos Alonso-Blanco	Monsanto	3
9602	IP-Vpa-1	CS78850	ESP	40.5	-3.96	Carlos Alonso-Blanco	Monsanto	3
9606	Aitba-1	CS76649	MAR	31.48	-7.45	0	Monsanto	3
9607	Panik-1	CS77161	RUS	53.05	52.15	0	Monsanto	3
9609	Adam-1	CS76645	RUS	51.41	59.98	0	Monsanto	3
9610	Lesno-4	CS77034	RUS	53.04	51.96	0	Monsanto	3
9611	Lesno-1	CS77032	RUS	53.04	51.9	0	Monsanto	3
9612	Lesno-2	CS77033	RUS	53.04	51.94	0	Monsanto	3
9613	Balan-1	CS76687	RUS	55.36	61.41	0	Monsanto	3
9615	Parti-1	CS77163	RUS	52.99	52.16	0	Monsanto	3
9616	Krazo-1	CS76984	RUS	53.06	51.96	0	Monsanto	3
9617	Karag-1	CS76960	RUS	51.37	59.44	0	Monsanto	3
9619	Basta-1	CS76691	RUS	51.84	79.48	0	Monsanto	3
9620	Basta-2	CS76692	RUS	51.82	79.48	0	Monsanto	3
9621	Basta-3	CS76693	RUS	51.84	79.46	0	Monsanto	3
9622	Bijisk-4	CS76707	RUS	52.52	85.27	0	Monsanto	3
9624	Chaba-2	CS76767	RUS	53.6	79.37	0	Monsanto	3
9625	Kolyv-2	CS76977	RUS	51.31	82.59	0	Monsanto	3
9626	Kolyv-3	CS76978	RUS	51.36	82.59	0	Monsanto	3
9627	Kolyv-5	CS76979	RUS	51.32	82.55	0	Monsanto	3
9628	Kolyv-6	CS76980	RUS	51.33	82.54	0	Monsanto	3
9629	K-oze-1	CS76957	RUS	51.35	82.18	0	Monsanto	3
9630	K-oze-3	CS76958	RUS	51.34	82.16	0	Monsanto	3
9631	Lebja-1	CS77015	RUS	51.65	80.79	0	Monsanto	3
9632	Lebja-2	CS77016	RUS	51.67	80.82	0	Monsanto	3
9633	Lebja-4	CS77017	RUS	51.63	80.83	0	Monsanto	3
9634	Masl-1	CS77073	RUS	54.13	81.31	0	Monsanto	3
9635	Nosov-1	CS77130	RUS	51.87	80.6	0	Monsanto	3
9636	Noveg-1	CS77131	RUS	51.75	80.82	0	Monsanto	3
9637	Noveg-2	CS77132	RUS	51.77	80.85	0	Monsanto	3
9638	Noveg-3	CS77133	RUS	51.73	80.86	0	Monsanto	3
9639	Panke-1	CS77162	RUS	53.82	80.31	0	Monsanto	3
9640	Rakit-1	CS77202	RUS	51.87	80.06	0	Monsanto	3
9641	Rakit-2	CS77203	RUS	51.9	80.06	0	Monsanto	3
9642	Rakit-3	CS77204	RUS	51.84	80.06	0	Monsanto	3
9643	Sever-1	CS77245	RUS	52.1	79.31	0	Monsanto	3
9644	Zupan-1	CS78882	CRO	45.07	18.72	0	Monsanto	3
9645	Gradi-1	CS76887	CRO	45.17	18.7	0	Monsanto	3
9646	Aiell-1	CS76648	ITA	39,126,899	16,170,188		Monsanto	3
9647	Basen-1	CS76690	ITA	40.37	16.77	0	Monsanto	3
9648	Bisig-1	CS76712	ITA	39.48	16.28	0	Monsanto	3
9649	Bivio-1	CS76713	ITA	39.13	16.17	0	Monsanto	3
9651	Filet-1	CS76858	ITA	40.68	14.87	0	Monsanto	3

9653	Giffo-1	CS76878	ITA	38.44	16.13	0	Monsanto	3
9655	Marce-1	CS77071	ITA	38.92	16.47	0	Monsanto	3
9656	Marti-1	CS77072	ITA	40.64	17.31	0	Monsanto	3
9657	Melic-1	CS77078	ITA	38.45	16.04	0	Monsanto	3
9658	Nicas-1	CS77127	ITA	38.97	16.34	0	Monsanto	3
9659	Pigna-1	CS77177	ITA	41.18	14.18	0	Monsanto	3
9660	Sarno-1	CS77236	ITA	40.84	14.57	0	Monsanto	3
9663	Teano-1	CS77357	ITA	41.33	14.09	0	Monsanto	3
9664	Mitterberg-1-179	CS78907	ITA	46,366	112,832	Joerg Wunder	Monsanto	3
9665	Mitterberg-1-180	CS78908	ITA	46,366	112,832	Joerg Wunder	Monsanto	3
9666	Mitterberg-1-182	CS78909	ITA	46,366	112,832	Joerg Wunder	Monsanto	3
9667	Mitterberg-1-183	CS78910	ITA	46,366	112,832	Joerg Wunder	Monsanto	3
9668	Mitterberg-2-184	CS77085	ITA	463,668	112,837	Joerg Wunder	Monsanto	3
9669	Mitterberg-2-185	CS77086	ITA	463,668	112,837	Joerg Wunder	Monsanto	3
9670	Mitterberg-2-186	CS77087	ITA	463,668	112,837	Joerg Wunder	Monsanto	3
9671	Mitterberg-3-187	CS77088	ITA	463,668	112,837	Joerg Wunder	Monsanto	3
9672	Mitterberg-3-188	CS77089	ITA	463,668	112,837	Joerg Wunder	Monsanto	3
9673	Mitterberg-3-189	CS77090	ITA	463,668	112,837	Joerg Wunder	Monsanto	3
9676	Mitterberg-4-192	CS77093	ITA	463,718	112,866	Joerg Wunder	Monsanto	3
9677	Mitterberg-4-193	CS77094	ITA	463,718	112,866	Joerg Wunder	Monsanto	3
9678	Mitterberg-4-194	CS77095	ITA	463,718	112,866	Joerg Wunder	Monsanto	3
9679	Castelfed-1-195	CS76744	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9680	Castelfed-1-196	CS76745	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9681	Castelfed-1-197	CS76746	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9682	Castelfed-1-198	CS76747	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9683	Castelfed-1-199	CS76748	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9684	Castelfed-2-200	CS76749	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9685	Castelfed-2-201	CS76750	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9686	Castelfed-2-202	CS76751	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9687	Castelfed-2-203	CS76752	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9689	Castelfed-3-205	CS76754	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9690	Castelfed-3-206	CS76755	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9691	Castelfed-3-207	CS76756	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9692	Castelfed-3-208	CS76757	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9693	Castelfed-3-209	CS76758	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9694	Castelfed-4-210	CS78891	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9695	Castelfed-4-211	CS78892	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9696	Castelfed-4-214	CS78893	ITA	463,378	112,928	Joerg Wunder	Monsanto	3
9697	Dolen-1	CS76802	BUL	41.62	23.94	0	Monsanto	3

9698	Goced-1	CS76882	BUL	41.57	23.85	0	Monsanto	3
9699	Kolar-1	CS76974	BUL	41.37	23.14	0	Monsanto	3
9700	Dolna-1-10	CS76803	BUL	42.32	23.1	0	Monsanto	3
9701	Ivano-1	CS76954	BUL	43.7	25.91	0	Monsanto	3
9703	Melni-1	CS77079	BUL	41.53	23.39	0	Monsanto	3
9704	Melni-2	CS77080	BUL	41.53	23.39	0	Monsanto	3
9705	Choto-1	CS76769	BUL	41.5	23.33	0	Monsanto	3
9706	Dospa-1	CS76807	BUL	41.64	24.18	0	Monsanto	3
9707	Podvi-1	CS77187	BUL	41.57	24.84	0	Monsanto	3
9708	Kardz-1	CS76962	BUL	41.62	25.35	0	Monsanto	3
9709	Zerev-1-34	CS78878	BUL	41.85	23.13	0	Monsanto	3
9710	Zerev-1-35	CS78879	BUL	41.85	23.13	0	Monsanto	3
9711	Dolna-1-39	CS76804	BUL	42.32	23.1	0	Monsanto	3
9712	Dolna-1-40	CS76805	BUL	42.32	23.1	0	Monsanto	3
9713	Stara-1	CS77271	BUL	42.49	25.61	0	Monsanto	3
9714	Grivo-1	CS76888	BUL	41.84	25.75	0	Monsanto	3
9716	Leska-1-44	CS77030	BUL	41.54	24.98	0	Monsanto	3
9717	Kardz-2	CS76963	BUL	41.66	25.47	0	Monsanto	3
9718	Smolj-1	CS77256	BUL	41.55	24.75	0	Monsanto	3
9719	Koren-1	CS76983	BUL	41.83	25.69	0	Monsanto	3
9720	Malak-1	CS77064	BUL	41.77	25.68	0	Monsanto	3
9721	Schip-1	CS77239	BUL	42.72	25.33	0	Monsanto	3
9722	Groch-1	CS76890	BUL	41.71	24.41	0	Monsanto	3
9723	Slavi-2	CS77252	BUL	41.42	23.67	0	Monsanto	3
9725	Epidauros-1	CS76844	GRC	37.6	23.08	0	Monsanto	3
9726	Fanerone mi-3	CS76853	GRC	37.07	22.04	0	Monsanto	3
9728	Stiav-1	CS77279	SVK	48.46	18.9	0	Monsanto	3
9730	Bela-1	CS76696	SVK	48.47	18.94	0	Monsanto	3
9731	Stiav-3	CS77281	SVK	48.46	18.9	0	Monsanto	3
9732	Halca-1	CS76909	SVK	48.47	18.96	0	Monsanto	3
9733	Bela-2	CS76697	SVK	48.47	18.94	0	Monsanto	3
9735	Bela-4	CS76699	SVK	48.47	18.94	0	Monsanto	3
9736	Teiu-2	CS77361	ROU	44.69	25.17	0	Monsanto	3
9737	Ulies-1	CS78815	ROU	45.95	22.62	0	Monsanto	3
9738	Bran-1	CS76722	ROU	45.57	25.42	0	Monsanto	3
9741	Orast-1	CS77151	ROU	45.84	23.16	0	Monsanto	3
9743	Furni-1	CS76873	ROU	45.14	25	0	Monsanto	3
9744	Iasi-1	CS76944	ROU	47.16	27.59	0	Monsanto	3
9745	Sij 1/96	CS77249	UZB	41.45	70.05	0	Monsanto	3
9747	Zabar-1	CS78870	SRB	44.38	21.22	0	Monsanto	3
9748	Zagub-1	CS78871	SRB	44.23	21.71	0	Monsanto	3
9749	Knjas-1	CS76971	SRB	43.54	22.29	0	Monsanto	3
9754	Sredn-1	CS77269	SRB	44.66	21.37	0	Monsanto	3
9755	Vajug-1	CS78828	SRB	44.56	22.56	0	Monsanto	3
9756	Staro-2	CS77273	SRB	44.3	21.08	0	Monsanto	3
9757	Staro-1	CS77272	SRB	44.3	21.08	0	Monsanto	3
9758	Altai-5	CS76433	CHN	47.75	88.4	0	Salk	3
9759	Anz-0	CS76439	IRN	37.47	49.47	0	Salk	3
9761	Bik-1	CS76449	LBN	33.92	35.7	0	Salk	3
9762	Etna-2	CS76487	ITA	37.69	14.98	0	Salk	3
9764	Qar-8a	CS76581	LBN	34.1	35.84	0	Salk	3
9766	Westkar-4	CS76629	KGZ	42.26	74.16	0	Salk	3
9768	Ru4-16	CS77225	GER	48.57	9.16	0	Monsanto	3
9769	HE-1	CS76916	GER	48.55	8.99	0	Monsanto	3
9770	KBG2-13	CS76966	GER	48.53	9.01	0	Monsanto	3
9771	Pfn-N2.2-6	CS77172	GER	48.56	9.11	0	Monsanto	3
9772	Hof-1	CS76925	GER	48.41	8.85	0	Monsanto	3
9774	Alt-1	CS76663	GER	48.59	9.22	0	Monsanto	3

9775	Berg-1	CS76701	GER	48.41	8.79	0	Monsanto	3
9776	Fell3-7	CS76857	GER	48.43	8.79	0	Monsanto	3
9777	Gn-1	CS76880	GER	48.57	9.17	0	Monsanto	3
9778	Bach-7	CS76679	GER	48.41	8.84	0	Monsanto	3
9779	Bai-10	CS76682	GER	48.5	8.78	0	Monsanto	3
9780	Fell2-4	CS76856	GER	48.43	8.79	0	Monsanto	3
9781	Kus2-2	CS76990	GER	48.52	9.11	0	Monsanto	3
9782	Lu3-30	CS77057	GER	48.53	9.09	0	Monsanto	3
9783	Tu-PK-7	CS77396	GER	48.52	9.05	0	Monsanto	3
9784	Erg2-6	CS76845	GER	48.5	8.8	0	Monsanto	3
9785	Ha-HBT1-2	CS76898	GER	48.54	9.02	0	Monsanto	3
9786	Ha-P-13	CS76901	GER	48.54	9.01	0	Monsanto	3
9787	HI-4	CS76922	GER	48.5	9	0	Monsanto	3
9788	KBG1-14	CS76965	GER	48.53	9.01	0	Monsanto	3
9789	Obh-13	CS77140	GER	48.39	8.96	0	Monsanto	3
9790	Gn2-3	CS76881	GER	48.58	9.18	0	Monsanto	3
9791	Haes-1	CS76914	GER	48.6	9.2	0	Monsanto	3
9792	Lu4-2	CS77058	GER	48.54	9.09	0	Monsanto	3
9793	Ru-N2	CS77224	GER	48.57	9.16	0	Monsanto	3
9794	Tu-B1-2	CS77391	GER	48.52	9.08	0	Monsanto	3
9795	Wank-2	CS78852	GER	48.5	9.11	0	Monsanto	3
9796	Bach2-1	CS76680	GER	48.41	8.84	0	Monsanto	3
9797	Ha-HBT2-10	CS76899	GER	48.54	9.02	0	Monsanto	3
9798	Ha-P2-1	CS76902	GER	48.54	9.01	0	Monsanto	3
9799	Hart-2	CS76913	GER	48.39	8.85	0	Monsanto	3
9800	Ha-S-B	CS76903	GER	48.54	9.01	0	Monsanto	3
9801	Ha-SP-2	CS76904	GER	48.54	9.01	0	Monsanto	4
9802	Kus3-1	CS76991	GER	48.51	9.11	0	Monsanto	4
9803	Muh-2	CS77113	GER	48.42	8.76	0	Monsanto	4
9804	Obe1-15	CS77139	GER	48.45	8.87	0	Monsanto	4
9805	Pfn-10	CS77171	GER	48.54	9.09	0	Monsanto	4
9806	Ru-2	CS77223	GER	48.56	9.16	0	Monsanto	4
9807	Schl-7	CS77240	GER	48.6	9.22	0	Monsanto	4
9808	Tu-B2-3	CS77392	GER	48.52	9.08	0	Monsanto	4
9809	Tu-KB-6	CS77393	GER	48.52	9.05	0	Monsanto	4
9810	Tu-KS-7	CS77394	GER	48.53	9.07	0	Monsanto	4
9812	Tu-W1	CS77397	GER	48.52	9.03	0	Monsanto	4
9815	Ha-HBT3-11	CS76900	GER	48.54	9.02	0	Monsanto	4
9816	Tu-WH	CS77398	GER	48.55	9.06	0	Monsanto	4
9822	IP-Aul-0	CS76675	ESP	40.52	-4.02	Carlos Alonso-Blanco	Monsanto	4
9823	IP-Bae-0	CS76681	ESP	43.34	-5.84	Carlos Alonso-Blanco	Monsanto	4
9824	IP-Bes-5	CS76702	ESP	42.91	-4.91	Carlos Alonso-Blanco	Monsanto	4
9825	IP-Boa-0	CS76714	ESP	40.4	-3.88	Carlos Alonso-Blanco	Monsanto	4
9826	IP-Bor-0	CS76717	ESP	42.49	-6.71	Carlos Alonso-Blanco	Monsanto	4
9827	IP-Bos-0	CS76719	ESP	42.78	0.69	Xavier Picv≥	Monsanto	4
9828	IP-Bra-0	CS76721	ESP	42.5	-6.15	Xavier Picv≥	Monsanto	4
9830	IP-Bus-0	CS76736	ESP	36.97	-3.28	Carlos Alonso-Blanco	Monsanto	4
9831	IP-Cas-0	CS76743	ESP	38.54	-3.39	Carlos Alonso-Blanco	Monsanto	4

9832	IP-Cat-0	CS76759	ESP	40.54	-3.69	Carlos Alonso-Blanco	Monsanto	4
9833	IP-Cha-0	CS76764	ESP	40.38	-4.21	Carlos Alonso-Blanco	Monsanto	4
9834	IP-Cho-0	CS76768	ESP	40.51	-3.9	Carlos Alonso-Blanco	Monsanto	4
9835	IP-Cir-0	CS76772	ESP	40.61	-6.57	Carlos Alonso-Blanco	Monsanto	4
9836	IP-Cod-0	CS76777	ESP	41.25	-1.32	Carlos Alonso-Blanco	Monsanto	4
9837	IP-Con-0	CS76780	ESP	37.94	-5.6	Xavier Picv≥	Monsanto	4
9839	IP-Coy-0	CS76785	ESP	40.44	-4.27	Carlos Alonso-Blanco	Monsanto	4
9843	IP-Elp-0	CS76840	ESP	40.53	-3.92	Carlos Alonso-Blanco	Monsanto	4
9846	IP-Ezc-2	CS76849	ESP	42.31	-3.02	Carlos Alonso-Blanco	Monsanto	4
9849	IP-Gud-3	CS76895	ESP	40.65	-4.11	Carlos Alonso-Blanco	Monsanto	4
9850	IP-Hec-0	CS76917	ESP	42.86	-0.7	Carlos Alonso-Blanco	Monsanto	4
9851	IP-Hue-3	CS76942	ESP	42.96	-6.1	Carlos Alonso-Blanco	Monsanto	4
9852	IP-Ini-0	CS76945	ESP	40.46	-3.75	Carlos Alonso-Blanco	Monsanto	4
9853	IP-Lac-0	CS76996	ESP	43.33	-5.91	Carlos Alonso-Blanco	Monsanto	4
9854	IP-Laf-1	CS76997	ESP	43.36	-5.88	Carlos Alonso-Blanco	Monsanto	4
9855	IP-Lam-0	CS77008	ESP	40.57	-3.89	Carlos Alonso-Blanco	Monsanto	4
9856	IP-Lch-0	CS77010	ESP	40.51	-4	Carlos Alonso-Blanco	Monsanto	4
9857	IP-Leg-0	CS77019	ESP	40.33	-3.8	Carlos Alonso-Blanco	Monsanto	4
9858	IP-Loz-0	CS77051	ESP	40.98	-3.8	Carlos Alonso-Blanco	Monsanto	4
9859	IP-Lro-0	CS77054	ESP	40.5	-3.88	Carlos Alonso-Blanco	Monsanto	4
9860	IP-Lum-0	CS77059	ESP	42.24	-2.62	Carlos Alonso-Blanco	Monsanto	4
9861	IP-Mac-0	CS77061	ESP	40.72	-3.21	Carlos Alonso-Blanco	Monsanto	4
9862	IP-Mad-0	CS77062	ESP	40.45	-3.67	Carlos Alonso-Blanco	Monsanto	4
9864	IP-Mat-0	CS77074	ESP	41.76	2.69	Xavier Picv≥	Monsanto	4

9867	IP-Mie-1	CS77083	ESP	40.94	-3.22	Xavier Picv≥	Monsanto	4
9868	IP-Moe-0	CS77104	ESP	41.78	2.37	Xavier Picv≥	Monsanto	4
9870	IP-Moz-0	CS77111	ESP	41.91	0.17	Carlos Alonso-Blanco	Monsanto	4
9877	IP-Pdl-0	CS77165	ESP	43.02	-5.6	Carlos Alonso-Blanco	Monsanto	4
9878	IP-Pee-0	CS77167	ESP	40.78	-3.62	Carlos Alonso-Blanco	Monsanto	4
9879	IP-Per-0	CS77169	ESP	37.6	-1.12	Carlos Alonso-Blanco	Monsanto	4
9880	IP-Pib-1	CS77175	ESP	42.72	-3.44	Carlos Alonso-Blanco	Monsanto	4
9881	IP-Pie-0	CS77176	ESP	40.46	-5.32	Carlos Alonso-Blanco	Monsanto	4
9882	IP-Pil-0	CS77178	ESP	40.46	-4.26	Carlos Alonso-Blanco	Monsanto	4
9883	IP-Piq-0	CS77179	ESP	42.1	-2.56	Carlos Alonso-Blanco	Monsanto	4
9885	IP-Prd-0	CS77189	ESP	41.14	-3.68	Carlos Alonso-Blanco	Monsanto	4
9886	IP-Pru-0	CS77190	ESP	42.38	1.73	Xavier Picv≥	Monsanto	4
9887	IP-Pun-0	CS77196	ESP	40.4	-4.77	Carlos Alonso-Blanco	Monsanto	4
9888	IP-Pva-1	CS77197	ESP	40.93	-3.31	Xavier Picv≥	Monsanto	4
9890	IP-Rib-1	CS77217	ESP	43.16	-5.07	Carlos Alonso-Blanco	Monsanto	4
9891	IP-Sal-0	CS77230	ESP	41.93	2.92	Xavier Picv≥	Monsanto	4
9892	IP-Sam-0	CS77231	ESP	42.68	-6.96	Carlos Alonso-Blanco	Monsanto	4
9894	IP-Sen-0	CS77243	ESP	42.59	0.76	Xavier Picv≥	Monsanto	4
9895	IP-Sfb-6	CS77248	ESP	41.78	2.57	Xavier Picv≥	Monsanto	4
9897	IP-Smt-1	CS77257	ESP	40.95	-5.63	Xavier Picv≥	Monsanto	4
9898	IP-Som-0	CS77259	ESP	41.14	-3.58	Carlos Alonso-Blanco	Monsanto	4
9899	IP-Tau-0	CS77342	ESP	42.54	0.84	Xavier Picv≥	Monsanto	4
9900	IP-Tri-0	CS77386	ESP	37.38	-6.01	Xavier Picv≥	Monsanto	4
9901	IP-Urd-1	CS78824	ESP	42.27	-2.98	Carlos Alonso-Blanco	Monsanto	4
9902	IP-Usa-0	CS78825	ESP	40.71	-3.24	Carlos Alonso-Blanco	Monsanto	4
9903	IP-Val-0	CS78829	ESP	42.31	-3.1	Carlos Alonso-Blanco	Monsanto	4

9904	IP-Vas-0	CS78833	ESP	40.95	-3.31	Carlos Alonso-Blanco	Monsanto	4
9905	IP-Ven-0	CS78840	ESP	40.76	-4.01	Carlos Alonso-Blanco	Monsanto	4
9906	IP-Mah-6	CS77063	ESP	40	4.25	Xavier Picv≥	Monsanto	4
9908	ESP-1-11	CS76847	FRA	50.72	3.47	0	Monsanto	4
9910	BRI-2	CS76725	FRA	50.68	3.52	0	Monsanto	4
9911	ARGE-1-15	CS76672	FRA	47.16	4.28	0	Monsanto	4
9912	CIRY-13	CS76773	FRA	46.67	4.55	0	Monsanto	4
9917	RAD-21	CS77200	FRA	46.69	4.34	0	Monsanto	4
9918	SAUL-24	CS77237	FRA	47.43	5.21	0	Monsanto	4
9920	DIR-9	CS76796	FRA	48.54	4.32	0	Monsanto	4
9925	RUM-20	CS77226	FRA	48.91	4.52	0	Monsanto	4
9926	TRE-1	CS77385	FRA	48.86	4.1	0	Monsanto	4
9927	ARR-17	CS76673	FRA	44.05	3.69	0	Monsanto	4
9933	VED-10	CS78839	FRA	43.74	3.89	0	Monsanto	4
9935	BAU-15	CS76694	FRA	50.6	2.93	0	Monsanto	4
9937	CATS-6	CS76760	FRA	50.79	2.69	0	Monsanto	4
9945	Leo-1	CS76413	ESP	41.8	-3.11		MPI	4
9948	Pra-6	CS76416	ESP	41.05	-3.54		MPI	4
9949	Qui-0	CS76417	ESP	42.69	-6.93		MPI	4
9950	Vie-0	CS76418	ESP	42.63	0.76		MPI	4
9952	Kly-4	CS76384	RUS	51.32	82.55		MPI	4
9953	Koz-2	CS76383	RUS	51.33	82.19		MPI	4
9955	Stepn-2	CS76377	RUS	54.09	60.46		MPI	4
9956	Stepn-1	CS76378	RUS	54.06	60.48		MPI	4
9957	Borsk-2	CS76421	RUS	53.04	51.75		MPI	4
9958	Shigu-1	CS76375	RUS	53.33	49.48		MPI	4
9959	Shigu-2	CS76374	RUS	53.33	49.48		MPI	4
9960	Kidr-1	CS76376	RUS	51.31	57.56		MPI	4
9962	Galdo-1	CS76423	ITA	40.57	15.32		MPI	4
9963	Lago-1	CS76367	ITA	39.18	16.26		MPI	4
9964	Mammo-1	CS76365	ITA	38.36	16.23		MPI	4
9965	Mammo-2	CS76364	ITA	38.38	16.22		MPI	4
9966	Monte-1	CS76361	ITA	40.28	15.65		MPI	4
9970	Altenb-2	CS76353	ITA	463,716	112,376	Joerg Wunder	MPI	4
9975	Castelfed-4-213	CS76356	ITA	463,378	112,928	Joerg Wunder	MPI	4
9978	Vezzano-2.2	CS76350	ITA	466,297	108,170	Joerg Wunder	MPI	4
9982	Apost-1	CS76368	ITA	39.01	16.47		MPI	4
9984	Ciste-2	CS76360	ITA	41.62	12.87		MPI	4
9986	Jablo-1	CS76372	BUL	41.59	25.2		MPI	4
9987	Lecho-1	CS76371	BUL	41.43	23.5		MPI	4
9988	Bak-2	CS76392	GEO	417,942	434,767	James Beck	MPI	4
9990	Lag2-2	CS76390	GEO	418,296	462,831	James Beck	MPI	4
9991	Vash-1	CS76391	GEO	412,381	463,728	James Beck	MPI	4
9993	Nemrut-1	CS76398	TUR	386,425	422,394	James Beck	MPI	4
9995	HKT2.4	CS76404	GER	48.14	9.4		MPI	4
9996	Nie1-2	CS76402	GER	48.52	8.8		MPI	4
9997	Rue3-1-31	CS76406	GER	48.56	9.16		MPI	4
9998	Star-8	CS76400	GER	48.43	8.82		MPI	4
9999	TueSB30-3	CS76403	GER	48.53	9.06		MPI	4
10001	TueV-13	CS76407	GER	48.52	9.05	Kirsten Bomblies	MPI	4

10002	TueWa1-2	CS76405	GER	48.53	9.04	Kirsten Bomblies	MPI	4
10004	Bolin-1	CS76373	ROU	44.46	25.74		MPI	4
10005	Copac-1	CS76420	ROU	46.11	21.95		MPI	4
10006	Kastel-1	CS76395	UKR	446,419	343,814	James Beck	MPI	4
10008	Sij-1	CS76379	UZB	41.45	70.05	Heike Schmuths	MPI	4
10009	Sij-2	CS76380	UZB	41.45	70.05		MPI	4
10010	Sij-4	CS76381	UZB	41.45	70.05		MPI	4
10011	Yeg-1	CS76394	ARM	398,692	453,622	James Beck	MPI	4
10012	Istisu-1	CS76389	AZE	389,786	485,594	James Beck	MPI	4
10013	Lerik1-3	CS76388	AZE	387,406	486,131	James Beck	MPI	4
10014	Xan-1	CS76387	AZE	386,536	487,992	James Beck	MPI	4
10015	Ara-1	CS76382	AFG	37.29	71.3		MPI	4
10017	Petro-1	CS76370	SRB	44.34	21.46		MPI	4
10018	Dobra-1	CS76369	SRB	44.84	20.16		MPI	4
10020	Jl-2	CS76956	CZE	49.17	16.5		MPI	4
10022	Uk-3	CS78777	GER	480,333	77,667	Albert Kranz	MPI	4
10023	Strand-1	CS77284	NOR	68.8	15.45		MPI	4
10027	Uk-6	CS78938	GER	4,802,838	7,765,567	Eunyoung Chae	MPI	4
14312	Kos-1	CS78923	RUS	62.02	34.12	Magnus Nordborg	GMI	4
14313	Kos-2	CS78924	RUS	62.02	34.12	Magnus Nordborg	GMI	4
14314	Radk-1	CS78927	RUS	61.59	35.11	Magnus Nordborg	GMI	4
14315	Radk-2	CS78928	RUS	61.59	35.11	Magnus Nordborg	GMI	4
14318	Shu-1	CS78930	RUS	61.94	34.24	Magnus Nordborg	GMI	4
14319	Shu-2	CS78931	RUS	61.94	34.24	Magnus Nordborg	GMI	4
15560	Valm	CS78932	RUS	61.37	61.37	Magnus Nordborg	GMI	4
15591	OOE1-1	CS78939	AUT	483,315,333	1,472,665	Wolfram Weckwert	GMI	4
15592	OOE3-1	CS78940	AUT	483,314,667	147,158,667	Wolfram Weckwert	GMI	4
15593	OOE3-2	CS78941	AUT	483,314,667	147,158,667	Wolfram Weckwert	GMI	4

Supplementary Table S3. The 118 accessions selected for the replica study of necrotic and non-necrotic accessions.

159	351	772	870	992	1063	1158	1257	1313	1317	1552	1622	1652	1684	1739	1741	1756
1820	1835	1853	1890	1925	2016	2031	2053	2057	2091	2106	2108	2141	2166	2285	2317	4857
4939	5023	5104	5151	5236	5353	5577	5860	6131	6191	6192	6220	6221	6231	6235	6237	6240
6252	6434	6926	6931	6940	6944	6957	6958	6959	6960	6961	6967	6970	6971	6981	7067	7103
7111	7127	7199	7202	7203	7231	7250	7268	7296	7346	7382	7394	7396	7417	7529	7530	8132
8171	8239	8264	8290	8366	8420	8426	8427	9078	9100	9102	9105	9106	9113	9555	9589	9597
9622	9681	9732	9747	9831	9890	9891	9892	9903	9927	9933	9963	9982	9995	9996		