MIND THE GAP

Analysis of Tsw resistance against tospoviruses

Irene L. van Grinsven

Propositions

1. The transfer of the RNA silencing suppressor NSs to the nucleus indicates a role for NSs in nuclear transcription.
   (this thesis)

2. All 3D protein prediction modeling is inherently incorrect.
   (this thesis)

3. Lecturer positions at university benefit research in the long run.

4. The temperature dependency of dominant resistance genes is biochemically irrelevant.

5. Embedding queer people in STEM (Science, Technology, Engineering and Math) workplaces enhances overall productivity.

6. Moving helps to obtain new perspectives when you encounter a problem.

7. The dominant group in society is comparable to a control group in scientific experiments.

8. Justice and happiness are intertwined.

Propositions belonging to the thesis, entitled

Mind the gap – Analysis of Tsw resistance against tospoviruses

Irene Louise van Grinsven
Wageningen, 10 October 2022

MIND THE GAP

Analysis of Tsw resistance against tospoviruses

Irene L. van Grinsven

Thesis committee

Promotors
Prof. Dr M.M. van Oers
Professor of Virology
Wageningen University & Research

Dr R.J.M. Kormelink
Associate professor, Laboratory of Virology
Wageningen University & Research

Other members
Prof. Dr F.P.M. Govers, Wageningen University & Research
Prof. Dr D.C. Baulcombe, University of Cambridge, United Kingdom
Dr E. van der Vossen, Solynta, Wageningen
Prof. Dr E. Schnettler, Bernhard-Nocht-Institute for Tropical Medicine, Germany

MIND THE GAP

Analysis of Tsw resistance against tospoviruses

Irene L. van Grinsven

Table of contents

Chapter 1

General introduction

I.L. van Grinsven

Introduction

Harmful pathogens and insect pests are everywhere. Humans, with their well-researched innate and adaptive immune systems, can attempt to stay out of dangerous situations, while most plants are sessile and must rely on other strategies to counteract pathogens. Over the past decades, scientists have started to uncover the multilayered plant defense system that provides protection against a multitude of pathogens and insect pests. Insects such as white flies, thrips, and aphids, are often vectors for plant pathogens and upon feeding on plant tissues may provide access for viruses, bacteria, oomycetes, and fungi into these plants. The cellular damage created by both pests and pathogens triggers a general defense mechanism to protect against invaders. This innate immune response is advantageous for the plant. Some plants also have a pathogen-specific response, often triggered by the activation of a host resistance gene, allowing for targeted defense responses. Resistance genes do not always provide absolute resistance against a pathogen, in which case oftentimes low levels of virus can still be detected. Certain mechanisms can allow for tolerance of pathogens, in which the pathogen's harmful effect is reduced, therefore shows little to no symptoms, but does not affect pathogen levels.

Plant innate immune responses

The first line of defense is known as non-host resistance (NHR), a generic mechanism against all plant pathogens comprised of two types of resistance as reviewed by Mysore and Ryu (2004). The first type of NHR is always present as it prevents invasion by pathogens via a structural and/or chemical defense, e.g. by cell wall thickening and ever-present secondary metabolites. Once a pathogen has managed to overcome this type 1 NHR and gained access, type 2 NHR is able to detect its presence. Pathogens often contain shared conserved structures, such as β-glucans found in fungal cell walls, and flagellin in flagella of gram-negative bacteria (Vidal et al., 1998; Felix et al., 1999; Che et al., 2000). These microbe- or pathogen-associated molecular patterns (MAMPs, PAMPs) can be recognized by pattern recognition receptors (PRRs) present on the exterior of plant membranes. The recognition of such elicitors by PRRs will trigger the activation of a downstream signaling pathway; the PAMP-triggered immune (PTI) response (Figure 1.1). This is associated with clearly visible local necrosis around the infection site on invaded plant leaves or stem, which will (partly) contain the pathogen (Jones and Dangl, 2006). This in contrast to the symptomless type 1 NHR.

Once the generic NHR has been surpassed, pathogens can gain entry into cells. Bacteria and fungi excrete cell-wall damaging proteolytic enzymes to enter cells. Unlike plant viruses, which need assistance from insect or nematode vectors, or from mechanical inoculation, to gain access into the plant cell (Underwood, 2012; Stavolone and Lionetti, 2017). Most plant viruses contain RNA as genetic material. Upon entry, such plant viruses will start to replicate, creating double stranded RNA (dsRNA) either from secondary folding structures of RNA or replication intermediates (Figure 1.1). Viral dsRNA in the cytoplasm triggers the antiviral RNA silencing, or interference (RNAi) pathway. Dicer-
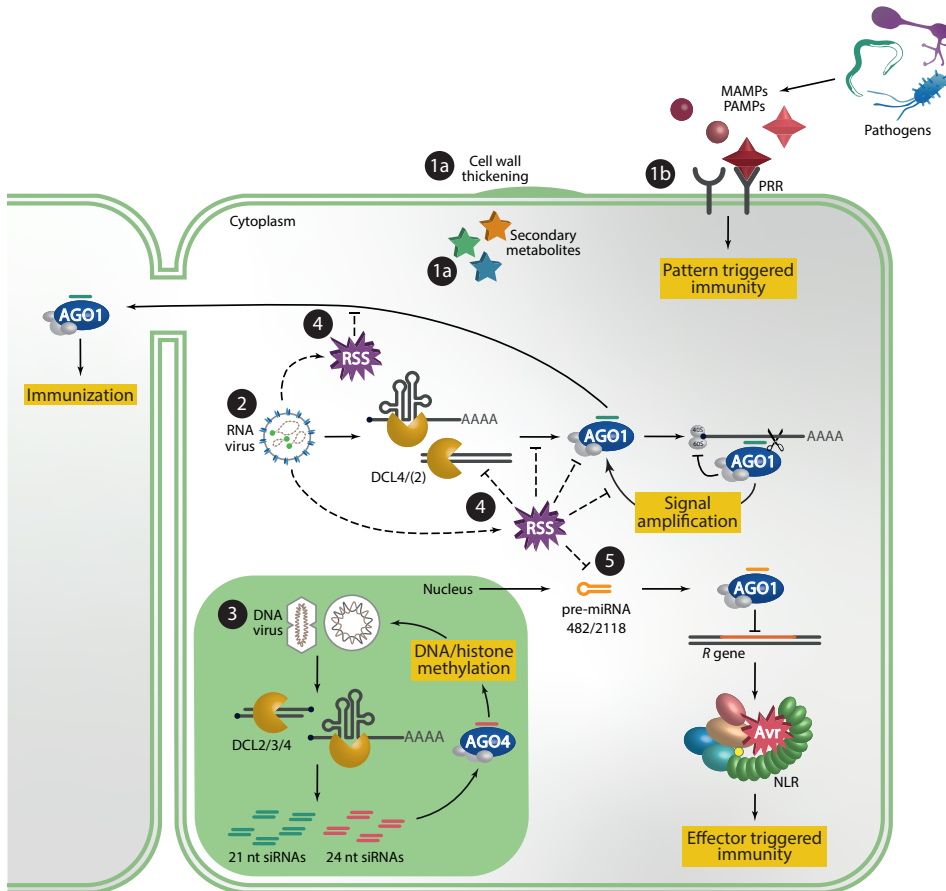
Figure 1.1 – Plant defenses and the viral counterattack. ❶ The first level of non-host resistance (NHR) consists of ever-present defenses such as increased cell-wall thickness and secondary metabolites in the cell cytoplasm. Pattern Recognition Receptors (PRRs) on the cell surface make up the second level of NHR. Conserved molecules from pathogens such as bacteria and nematodes, known as microbe- or pathogen-associated molecular patterns (MAMPs, PAMPs), are sensed by PRRs which activates the Pattern Triggered Immunity. **❷** Viruses gain access through damaged cell walls and membranes caused by other pathogens and insect pests. Viral double stranded RNA (dsRNA) originating from secondary folding structures of RNA or replication intermediates triggers the host RNA interference (RNAi) pathway to defend against invaders. dsRNA is cleaved into 21 24-nucleotide viral small interfering RNAs by Dicer-like (DCL) endoribonucleases. Incorporation of small interfering RNAs (siRNAs) into Argonaute (AGO) 1 activates the RNA-induced silencing complex (RISC) and enables the sensing and degradation of viral target RNAs. The RISC spreads to neighboring cells through the plasmodesmata to immunize the surrounding cells. The cleaved RNA is either degraded or used to amplify the RNAi response by creating more, secondary, siRNAs by RNA-dependent RNA polymerases (RDRs). **❸** dsRNA molecules derived from DNA viruses in the nucleus are also cleaved into siRNAs by DCLs and loaded into AGO4, which targets the complementary DNA for DNA/histone methylation to silence the affected gene. **❹** To counteract antiviral RNAi and successfully establish a viral infection, viruses produce RNA silencing suppressors (RSSs). Many of them sequester long dsRNA and siRNAs to prevent the biogenesis of siRNAs and their uploading into RISC, respectively, or prevent development of the RISC and/or cleavage of target RNA. **❺** As a side effect of the silencing suppression, the silencing of resistance (*R*) genes regulated by miRNAs is lifted. As a result, *R* genes are translated and able to detect effectors, such as the RSS protein, resulting in effector triggered immunity.

like (DCL) endoribonucleases detect dsRNA and cleave it into viral small interfering RNA (vsiRNA) of 21-24 nucleotides (nts) in size. The resulting double stranded vsiRNAs are unwound by a helicase and one strand, the so-called guide strand, is loaded into an RNA-induced silencing complex (RISC). This incorporated siRNA strand guides the RISC to complementary viral RNA, which is subsequently cleaved. For an extensive review on small RNA-based immunity, see Guo et al., 2019. The cleaved RNA is either further degraded or used by the host's RNA-dependent RNA polymerases (RDRs) to create secondary siRNAs, thereby amplifying the antiviral RNAi response (Yang and Li, 2018). Furthermore, the dsRNAs created by DNA plant viruses upon replication and transcription in the nucleus of plant cells are also cleaved into 24 nt siRNAs, which are loaded into Argonaute (AGO) 4. DNA with complementarity to the loaded siRNA strand is marked for *de novo* DNA/histone methylation by AGO4, which transcriptionally silences the targeted gene and as a consequence inhibits viral replication.

The antiviral RNAi defense system works against all plant viruses. However, the process is relatively slow and generally does not result in a total prevention and clearance of progeny virus particle formation. Furthermore, almost all plant-infecting viruses have developed proteins to suppress the host RNAi pathway. These RNA silencing suppressors (RSSs) are able to block one or more steps in the RNAi pathway, lifting the suppression of the virus and allowing its replication. RSSs suppress the RNAi pathway for instance by sequestering dsRNA or siRNA, respectively inhibiting the formation of siRNAs and the loading of guide RNA strands into the AGO proteins. For instance, the RSS of Cucumber mosaic virus (CMV), 2b, and NS3 of Rice Hoja Blanca Virus bind dsRNA (Goto et al., 2007; Hemmes et al., 2007). Other RSS proteins prevent the development of RISC or the cleavage of targeted RNA by RISC. For example, the RSS of Beet western yellow virus, P0, targets Argonaut proteins for degradation (Baumberger et al., 2007).

Effector-mediated resistance

Plants have acquired pathogen-specific defense genes to recognize particular pathogens and induce defense responses. The presence of a specific pathogen protein, known as elicitor, effector, or avirulence (Avr) determinant, is recognized by a resistance (*R*) gene that upon activation triggers the effector-triggered immune (ETI) response (Figure 1.1) (Caplan et al., 2008b; Jones et al., 2016). Regardless of the recognized pathogen, most *R* genes encode for intracellular receptors that contain a domain with a central nucleotide binding (NB) adaptor that shares homology with NB domains from several other apoptosis-regulating proteins: human Apaf1 and nematode CED-4. The domain is therefore also known as NB-ARC, where ARC comes from **A**paf1, **R** proteins and **C**ED-4 (van der Biezen and Jones, 1998). *R* genes also encode a C-terminal leucine-rich repeat (LRR) region (Figure 1.2). Upon infection, recognition by these nucleotide-binding leucine-rich repeat receptors (NLRs) leads to programmed cell death (PCD) and a hypersensitive response (HR) at the infection site, stopping the spread of the pathogen. NLRs are further

classified into Toll/interleukin-1 (TIR)-NLRs (TNLs) and coiled-coil (CC)-NLRs (CNLs) based on the N-terminal domain (Meyers et al., 2003).

The highly variable LRR domain is considered essential for pathogen recognition specificity (Dodds et al., 2001; Collier and Moffett, 2009a; Jones et al., 2016), although only a few R proteins are known to directly recognize the Avr. Most pathogen recognitions by R proteins occur indirectly through host proteins that function either as guards, decoys, or bait. Regardless of the model, interaction of these host proteins with the Avr triggers a change that activates the R protein (van der Biezen and Jones, 1998; Jones and Dangl, 2006; van der Hoorn and Kamoun, 2008; Collier and Moffett, 2009a).

The central domain of NLRs can bind to adenosine 5′-diphosphate (ADP) or adenosine 5′-triphosphate (ATP). This binding functions as a molecular OFF (ADP)/ON (ATP) switch to change from an auto-inhibited "off" state to an activated "on" state (Figure 1.2) (van Ooijen et al., 2008; Takken and Tameling, 2009; Maekawa et al., 2011; Williams et al., 2011). The CC domains in CNLs and the LRR domains in both types of NLRs suppress the ADP/ATP exchange by intracellular domain interaction to keep the NLR in its auto-inhibited state
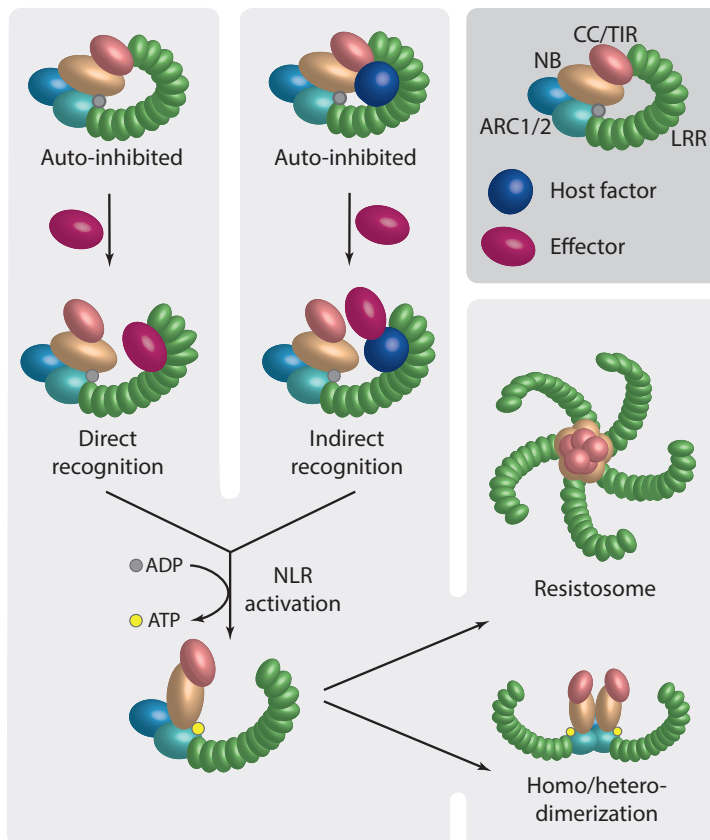
Figure 1.2 – Effector recognition and the subsequent activation of *R* genes. A pathogen effector is recognized directly, or indirectly by sensing the modification of a host factor by the effector. Exchange of ADP (grey dot) for ATP (yellow dot) switch activates the *R* gene. For some *R* genes, homo/heterodimerization or the formation of a resistosome is necessary to trigger downstream resistance pathways. *Abbreviations: ARC, Apaf-1, R-protein, and CED-4 domain; CC, coiled-coil; LRR, leucine-rich repeat domain; NB, nucleotide binding domain; NLR, nucleotide-binding leucine-rich repeat receptor.*

Table 1.1 – An overview of the cloned and characterized anti-viral NLRs and their corresponding effectors. *Abbreviations: CNL – coiled coil NLR; TNL – Toll/inter-leukin-1 NLR; TCV – Turnip crinkle virus; CMV – Cucumber mosaic virus; MYMIV – Mungbean yellow mosaic virus; TSWV – Tomato spotted wilt virus; PVX – Potato virus X; SMV – Soybean mosaic virus; GRSV – Groundnut ringspot virus; TCSV – Tomato chlorotic spot virus; CP- coat protein; RdRp – RNA dependent RNA polymerase; MP – Movement protein; HR – hypersensitive response; ER – extreme resistance; SN – systemic necrosis.*

R gene	Encoded protein	Plant species	Virus	Effector	Resistance response	Ref.
Hrt1	CNL	Arabidopsis	TCV	CP	HR	(Ren et al., 2005)
Rcy1	CNL	Arabidopsis	CMV	CP	HR	(Takahashi et al., 2002)
CYR1	CNL	Black gram	MYMIV	CP	-	(Maiti et al., 2012)
PvCMR1	TNL	French bean	CMV	2a	SN	(Seo et al., 2006)
L1, 2, 3, 4	CNL	Pepper	Tobamoviruses	CP	HR	(Tomita et al., 2011)
Pvr4	CNL	Pepper	Potyvirus	RdRp	HR	(Kim et al., 2017a)
Tsw	CNL	Pepper	TSWV	NSs	HR	(Kim et al., 2017a)
Rx1	CNL	Potato	PVX	CP	ER	(Tameling and Baulcombe, 2007)
Rx2	CNL	Potato	PVX	CP	ER	(Tameling and Baulcombe, 2007)
3gG2	CNL	Soybean	SMV	P3	ER	(Wen et al., 2013)
Rsv1	CNL	Soybean	SMV	HcPro proteins	ER, HR	(Wen et al., 2013)
N	TNL	Tobacco	Tobamovirus	Helicase/ p50 subunit	HR	(Caplan et al., 2008a)
N'	CNL	Tobacco	Tobamovirus	CP	HR	(Sekine et al., 2012)
Sw-5b	CNL	Tomato	TSWV, GRSV, TCSV	NSm	ER, HR	(Zhu et al., 2017)
Sl5R-1	CNL	Tomato	TSWV	Unknown	Unknown	(Qi et al., 2022)
Tm-2	CNL	Tomato	Tobamoviruses	30 kDa MP	ER	(Lanfermeijer et al., 2005)
Tm-22	CNL	Tomato	Tobamoviruses	30 kDa MP	ER	(Lanfermeijer et al., 2005)

(Sukarta et al., 2016). The inhibition by the CC and LRR is lifted only upon detection of the pathogen, allowing for the further activation of the NLR protein. While an HR response triggered upon *R* gene activation involves multiple cells, the PCD associated with the so-called extreme response (ER) only occurs in a single cell. ER is not associated with necrotic lesions and virus accumulation is not detectable. Both HR and ER involve PCD, rapidly killing the infected cells to prevent systemic spread, by activation and expression of for instance the phytohormones ethylene, salicylic acid (SA), and jasmonic acid (JA), as well as different reactive oxygen species (ROS) and Ca^{2+}. If the HR is not sufficient to contain the virus, the pathogen may escape from the initial infection site to infect other parts of the plant. The pathogen's presence will still trigger an HR along its route through the tissues, causing systemic necrosis (SN). However, as the virus is still able to replicate, its presence may ultimately lead to the death of the plant.

It may seem that there is a hard divide between extracellular PTI and the intracellular ETI. There are however indications that the split is more diffuse than assumed in the past (Hatsugai et al., 2017). For instance, the proteins of the EDS1 family are well-known ETI regulators, but are also used by surface receptors to trigger a defense response upon recognition of extracellular conserved pathogen molecules (Pruitt et al., 2021; Tian et al., 2021). In turn, the *A. thaliana* NLRs RRS1 and RPS4 require activation of PRRs usually involved in the PTI to trigger an ETI response (Ngou et al., 2021; Yuan et al., 2021).

Continuous high expression of *R* genes is often disadvantageous to the plant due to associated fitness costs (Burdon and Thrall, 2003; Korves and Bergelson, 2004; Huot et al., 2014). However, upon an infection the immune system should be able to respond quickly, which requires a good regulation of *R* gene transcription. One of the key pathways that down-regulates the transcription level of *R* genes is suppression by microRNAs (miRNAs). These short 21-22 nt RNAs are derived from hairpin structures in host-encoded RNA transcripts, which are used to produce mature single-stranded miRNAs in the end. Micro RNAs bind to AGO1 and specifically target messenger RNA (mRNA), which is then degraded or inhibited in its translation. Furthermore, RDRs can be recruited in this case as well, which triggers cleavage of the targeted mRNA into 21 nt phased secondary siRNAs (phasiRNAs) to amplify the silencing. Translation of NLRs has been shown to be suppressed by the miRNA superfamily miR482/2118, which targets a conserved motif in the RNA sequence encoding the NB-ARC domain (Shivaprasad et al., 2012). As pathogen invasion often leads to RNA silencing suppression by effectors acting as RSS, the degradation of the NLR mRNA by miRNAs is lifted, allowing for its translation (Figure 1.1). This negative regulation prevents autoimmunity in the cell under healthy circumstances, while also allowing for a rapid response of the immune system upon infection.

The gene-for-gene model postulated by Flor (1971) stated that the inheritance of both plant resistance and a pathogen's ability to lead to disease is caused by a pair of matching genes. In the fifty years that followed, many different *R* genes were cloned, some of which

indeed functioned as a single protein for recognition of a disease-causing protein of a pathogen and the activation of defense mechanisms. A handful of single dominant anti-viral NLRs has been cloned and characterized (Table 1.1).

An extensive review by de Ronde et al. covered all then known dominant *R* genes (de Ronde et al., 2014a). Recent findings have uncovered a large group of NLRs that do not function as a single unit, but rather as a pair consisting of one sensor NLR and one helper NLR (Cesari et al., 2014; Wu et al., 2016; Jubic et al., 2019). The two proteins in the pair fulfill complementary functions; the sensor NLR recognizes the pathogen, while the helper NLR is able to trigger the downstream immune pathway. Recently, a large network of NLRs was discovered in plants of the *Solanaceae* family containing multiple sensor and helper NLRs (Wu et al., 2017). The presence of several helper NLRs, known as NLR-required for cell death (NRC) helpers, was required to trigger an immune response. These NRCs could be activated by various sensor NLRs that recognized a diverse array of pathogens, from viruses to insects (Wu et al., 2017).

While many different plant NB-LRR genes and their corresponding Avrs have been studied, often the resistance inducing pathway has not been elucidated. Meanwhile, the research into their animal counterparts has advanced further. Human Apaf1 and some animal NLRs are known to oligomerize into large protein complexes (Zhang et al., 2015; Zhou et al., 2015), respectively named apoptosomes and inflammasomes, that trigger apoptosis and inflammation. It was already shown that upon pathogen recognition, some plant NLRs use their N-terminal domain to transactivate other receptors. These N-terminal domains homo- and heterodimerize, which has been found to be essential for cell death triggering (Maekawa et al., 2011; Williams et al., 2014). However, only recently the oligomerization of a CC-NLR from *Arabidopsis thaliana*, HOPZ-ACTIVATED RESISTANCE 1 (ZAR1), into a large protein complex was shown (Wang et al., 2019a). This pentameric complex, to which the name resistosome was given, is essential to induce resistance against the bacterium *Xanthomonas campestris* and resembles the wheel-like form of animal inflammasomes and apoptosomes (Figure 1.2). Following the discovery of the ZAR1 resistosome, two TNLs in *A. thaliana*, *ROQ1* and *RPP1* (Ma et al., 2020; Martin et al., 2020b), were both found to form a tetrameric resistosome. The full resistance mechanism of all resistosome NLRs remains to be elucidated, but the finding of the resistosome provides a steppingstone for future research.

Durability of resistance genes

A pathogen can develop ways to suppress the PTI and thereby block *R* gene activation, or become unrecognizable for the *R* gene product by mutating the effector. Cases have been reported in which a single amino acid mutation was sufficient to break resistance (Karasawa et al., 1999; Almási et al., 2017). The resulting resistance breaking (RB) strain will no longer trigger an immune response and can again infect its hosts. This "arms race" between plant immune system and pathogen is commonly illustrated by Jones and

Dangle's Zig Zag model (2006). This model does not take into account that mutations may lead to a loss of pathogen fitness (Janzac et al., 2010; Fraile et al., 2011). Especially for viruses, whose genes due to their small genome often have multiple functions, most of which are crucial for viral fitness. Mutations in effector genes therefore can lead to fitness losses, which may hinder the evolution of RB strains. Nevertheless, many virulent RB strains can be found, which can spread quickly as they are not halted by *R* genes, and reduce the durability of these dominantly inherited *R* genes.

While resistance (*R*) genes confer resistance to pathogens, host susceptibility (*S*) genes do the opposite and facilitate the infection. Mutation or loss of such *S* genes will limit the pathogen efficacy and provide (partial) resistance. *S* genes for example aid in the translation of a virus, as is the case for potyvirus infection in *A. thaliana*, where translation initiation factors were found to be recruited to help with viral translation (Nicaise et al., 2007). Mutation/loss of an *S* gene can offer another pathway for (non-host) resistance breeding aside from *R* genes as long as it leads to sufficient and durable resistance, that is also not easily broken or surpassed by the pathogen and does not affect cellular function. By side-stepping the *R* gene product, pathogens can surpass *R* gene resistance. Pathogens attempting to circumvent *S* gene resistance must overcome their dependence on a host factor and therefore must acquire a completely new function. Resistance based on removal or mutation of *S* genes is thus predicted to be more durable than *R* gene resistance (Garcia-Ruiz, 2018).

Tomato spotted wilt virus

One of the most economically important plant viruses is Tomato spotted wilt virus (TSWV) (Scholthof et al., 2011). *Tomato spotted wilt virus* is the type species of the genus *Orthotospovirus*, in the family *Tospoviridae* within the order *Bunyavirales* (Maes et al., 2018)*;* an order in which most other members are human- or animal-infecting viruses. TSWV has a very broad host range and is able to infect over 1000 different plant species, including crops as tomato, pepper, and ornamentals like chrysanthemum and alstroemeria (Parrella et al., 2003). As the name implies, common symptoms on tomato include leaf wilting, and in addition often leads to chlorotic or necrotic ringspots on leaves and fruits, as well as stunted growth. TSWV affects both the quality and the yield of the harvested products, attributing to costly crop losses (Goldbach and Peters, 1994; Srinivasan et al., 2017). This plant virus is transmitted by thrips in a propagative manner (Wijkamp et al., 1993), and is mainly spread by the western flower thrips (*Frankliniella occidentalis*) (Rotenberg et al., 2015). Since the 1980s, the worldwide spread of the western flower thrips has led to the worldwide re-emergence and distribution of TSWV (Gilbertson et al., 2015).

The genome of TSWV consists of three single stranded RNA molecules of negative/ambisense polarity (Figure 1.3). The large (L) segment codes for the viral RNA-dependent RNA polymerase (RdRp). The medium (M) segment encodes the precursor of the two glycoproteins (Gc, Gn) and the non-structural cell-to-cell movement protein (NSm),
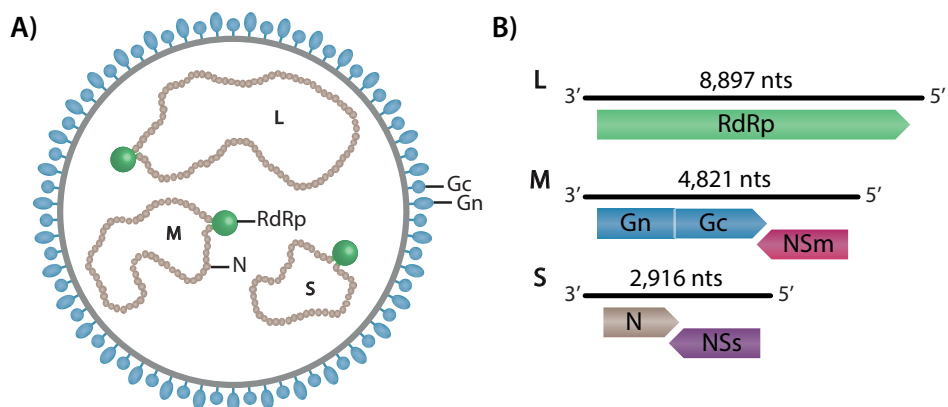
Figure 1.3 – Genome organization of Tomato spotted wilt virus (TSWV). A) Schematic of an enveloped virus particle TSWV. **B)** Organization of the tripartite RNA genome of TSWV. *Abbreviations: Gc, glycoprotein processed from the carboxy-terminus of the glycoprotein precursor; Gn, glycoprotein processed from the amino-terminus of the glycoprotein precursor; L, large RNA segment; M, middle RNA segment; N, nucleocapsid protein; NSm, nonstructural protein encoded by the M RNA segment, presents the movement protein; NSs, nonstructural protein encoded by the S RNA segment, presents the RNA silencing suppressor protein; nts, nucleotides; RdRp, RNA-dependent RNA polymerase; S, small RNA segment.*

while the small (S) codes for the nucleocapsid protein (N) and the non-structural RNA silencing suppressor protein (NSs) (de Haan et al., 1991; Kormelink et al., 1991, 1994, 2011; Adkins et al., 1995; van Knippenberg et al., 2002). The spherical TSWV particles are 80-120 nm in diameter and are enveloped by a phospholipid membrane in which the two glycoproteins are embedded. At the core, the particles contain the tripartite genome encapsidated by the N protein and a few copies of RdRp to form replication competent ribonucleoproteins (RNPs). RNPs can also associate with the movement protein NSm, which facilitates their spread towards neighboring cells through the plasmodesmata. The non-structural protein NSs binds long dsRNA as well as miRNA and siRNA, inhibits long dsRNA cleavage by DCLs and, as a consequence of all these actions, suppresses gene silencing (Schnettler et al., 2010; Hedil et al., 2017). NSs is an RSS and contains a typical GW/WG motif identified in various other RSS proteins, associated with the binding of Argonaut by these RSS proteins (Giner et al., 2010). Whether TSWV NSs genuinely interacts with AGO1 is still not known. However, mutation of the GW/WG motif within TSWV NSs does diminish its ability to locally silence RNA (de Ronde et al., 2014b; Hedil et al., 2015).

Resistance against tomato spotted wilt virus

To avoid introduction and spread of viruses through thrips, pesticides are often used against these insects. However, thrips have become resistant to many pesticides, urging the need for alternatives to combat the virus. One way is by deploying resistance genes. Two resistance genes have been found to confer resistance against TSWV and are commercially available: *Sw-5b* in *Solanum lycopersicum* and *Tsw* in *Capsicum chinense*. Both genes are NLRs (Brommonschenkel et al., 2000; Spassova et al., 2001; Kim et al., 2017a) and provide resistance by inducing an HR response, but they are activated by

different viral effectors. *Sw-5b* confers resistance against a few strongly related American-type tospoviruses, which include TSWV, Groundnut ringspot virus, and Tomato chlorotic spot virus (Folkertsma et al., 1999; Brommonschenkel et al., 2000; Leastro et al., 2017; de Oliveira et al., 2018). The Sw-5b protein is triggered by the presence of a 21 amino acid (aa) region that is highly conserved in American-type tospoviruses, but not in Euro-Asian tospoviruses (Zhu et al., 2017). Mutations within this 21-aa region affect the intracellular movement of TSWV NSm (Peiró et al., 2014; Feng et al., 2016), suggesting that this 21-aa region, in addition to other, essential, domains, is important for cell-to-cell movement of TSWV. While avirulence can be affected upon mutation of domains critical for movement, these regions are considered to be independent from the avirulence of NSm towards Sw-5b, and to only indirectly affect avirulence (Zhao et al., 2016; Leastro et al., 2017; Huang et al., 2018).

Tsw on the other hand only confers resistance against TSWV (Boiteux and de Ávila, 1994; Boiteux, 1995; Jahn et al., 2000). Tsw is activated by the presence of the RNA silencing suppressor NSs (Margaria et al., 2007; de Ronde et al., 2013). The *Tsw* and *Sw-5b* genes have both been introgressed into commercially available lines of pepper and tomato respectively to combat the spread of TSWV. Various resistance breaking (RB) TSWV strains have been reported for both *Tsw* (Margaria et al., 2007; de Ronde et al., 2013; Almási et al., 2015; Ferrand et al., 2015; Jiang et al., 2017a; Ronde et al., 2019) and *Sw-5b* (Thompson and van Zijl, 1996; Latham and Jones, 1998; Canady et al., 2001; Aramburu and Marti, 2003; Ciuffo et al., 2005; Gordillo et al., 2008; Zaccardelli et al., 2008; Batuman et al., 2017). Furthermore, Tsw is not functional at temperatures higher than 30-32 °C (Chung et al., 2012; Ronde et al., 2019). Increased global warming and the worldwide spread of TSWV thus emphasize the importance of gaining knowledge on TSWV resistance genes and, towards the future, finding/developing novel resistance methods.

Thesis outline

At the outset of the presented research, Kim et al. published their discovery of *Tsw* and the *R* gene cluster in *C. chinense* (Kim et al., 2017a). In that paper, the cloned gene was confirmed to provide resistance against TSWV, while the mechanism behind the recognition of NSs and the resulting downstream signaling remained unknown. The major objective of this thesis was to unravel Tsw-mediated resistance and its activation by NSs. This would create a further understanding of Tsw-mediated resistance and assist in finding new targets for antiviral strategies, necessitated by increased resistance breaking and rising day temperatures leading to a reduced effectiveness of Tsw.

Preceding the experimental analysis of these topics, an overview is presented on the molecular biology and tospoviral life cycle, and the distinct layers of defense against tospoviruses in **Chapter 2**. The mechanism behind the more extensively studied Sw-5b-conferred resistance against TSWV is examined in detail, which provides a theoretical starting point to facilitate the study of Tsw.

While the amino acid sequence of the functional resistance gene *Tsw* and the susceptible homolog *tsw* had been published by Kim et al., their respective nucleotide sequences – including non-coding, intron, sequences – had not. These non-coding sequences might provide insight into the potential transcriptional regulation by transcription factors in the 5' untranslated region (UTR), as well as possible translation inhibition through miRNA targets in the introns. The creation of a bacterial artificial chromosome (BAC) library of *C. chinense* PI152225 provided a targeted approach towards obtaining the full sequence of *Tsw* (**Chapter 3**).

As the aa sequences as published by Kim et al. indicated, Tsw belongs to the 10% of NLRs that contain an unusually large LRR domain. In addition to the sequence of Tsw obtained through our sequencing of the BAC library (Chapter 3), next 3D model predictions were made of Tsw and tsw. The folding structure of Tsw obtained was compared with the resistance gene ZAR1 and other NLRs (**Chapter 4**), and questions discussed regarding the interaction between Tsw, NSs and, potentially, host factors and RNA.

Earlier work already showed that the TSWV NSs protein binds si- and miRNAs that can target genes for inhibition of translation, thereby possibly influencing gene expression. Using the non-coding sequences of Tsw uncovered through the BAC sequencing (Chapter 3) *in silico* analysis of the transcriptional and translational regulation was performed, as well as a pilot study of the effect of miRNA silencing on TSWV infection (**Chapter 6**). Furthermore, to enable functional gene analyses a method was devised for future construction and expression of Tsw and variants with truncations or mutations based on its predicted 3D model.

The exact activation and downstream signaling pathways of Tsw currently remain unknown. As the search for proteins interacting with Tsw is still ongoing, the availability of NSs (wild type and variant) gene constructs provided a way to potentially close the knowledge gap from the other side. In **Chapter 6**, NSs-GFP constructs were used in a co-immunoprecipitation (co-IP) assay followed by mass spectrometry and in confocal microscopy to uncover host proteins potentially interacting with NSs. The interaction with detected host proteins was verified and the biological relevance of confirmed interactions further investigated.

In **Chapter 7**, the findings of this thesis are summarized and discussed in light of the current knowledge on *R* gene activation and downstream signaling, and future perspectives and questions remaining, or resulting from this thesis, are presented.

Chapter 2

Paving the way to tospovirus infection: multilined interplays with plant innate immunity

M. Zhu, I.L. van Grinsven, R. Kormelink, and X. Tao

Abstract

Tospoviruses are among the most important plant pathogens and cause serious crop losses worldwide. Tospoviruses have evolved to smartly utilize the host cellular machinery to accomplish their life cycle. Plants mount two layers of defense to combat their invasion. The first one involves the activation of an antiviral RNA interference (RNAi) defense response. However, tospoviruses encode an RNA silencing suppressor that enables them to counteract antiviral RNAi. To further combat viral invasion, plants also employ intracellular innate immune receptors (e.g., Sw-5b and *Tsw*) to recognize different viral effectors (e.g., NSm and NSs). This leads to the triggering of a much more robust defense against tospoviruses called effector-triggered immunity (ETI). Tospoviruses have further evolved their effectors and can break *Sw-5b-/Tsw*-mediated resistance. The arms race between tospoviruses and both layers of innate immunity drives the coevolution of host defense and viral genes involved in counter defense. In this review, a state-of-the-art overview is presented on the tospoviral life cycle and the multilined interplays between tospoviruses and the distinct layers of defense.

Introduction

Tospoviruses belong to the genus *Orthotospovirus*, the family *Tospoviridae*, and the order *Bunyavirales*, of which most members are restricted to animal and human hosts. Tomato spotted wilt virus (TSWV), the type species of the tospoviruses (siglum from **to**mato **spo**tted), ranks within the top 10 most important plant viruses (Scholthof et al., 2011), with estimated annual crop losses of more than one billion US dollars. Tospoviruses cause increasing problems in agricultural and ornamental crops worldwide (Oliver and Whitfield, 2016; Turina et al., 2016) and are transmitted by thrips in a persistent, circulative–propagative manner (Whitfield et al., 2005; Rotenberg et al., 2015). The western flower thrips (*Frankliniella occidentalis*) is known as one of the most important vectors of tospoviruses, and its worldwide spread starting in the 1980s has contributed to the reemergence and worldwide distribution of TSWV (Gilbertson et al., 2015).

In the past decade, excellent reviews have appeared on various aspects of the epidemiology, molecular biology, cytopathology, virus–vector interactions, and resistance strategies of tospoviruses (Whitfield et al., 2005; Kormelink, 2011; Gilbertson et al., 2015; Oliver and Whitfield, 2016; Turina et al., 2016). In the following sections, an overview of the molecular biology and life cycle of tospoviruses is presented, based on the most important and best-studied member of the tospoviruses, TSWV, with a focus on the achievements obtained in recent years and the multilined interplays with plant innate immunity. Most of this overview is representative for all members of the tospoviruses unless stated otherwise.

Molecular biology and life cycle of tospoviruses

The genome of TSWV consists of three negative/ambisense-strand genomic RNA segments (Figure 2.1) that are denoted large (L), medium (M), and small (S) and code for

A)
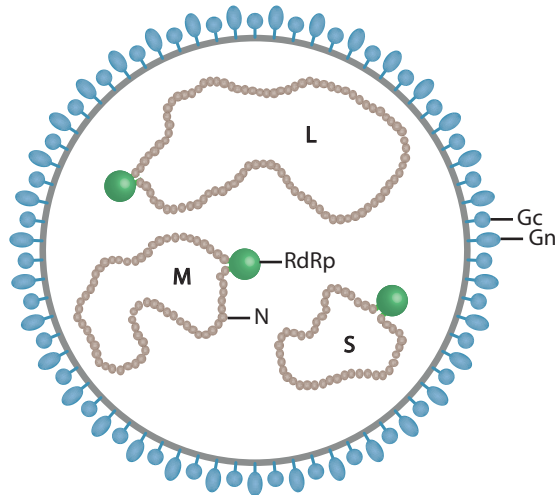
B)
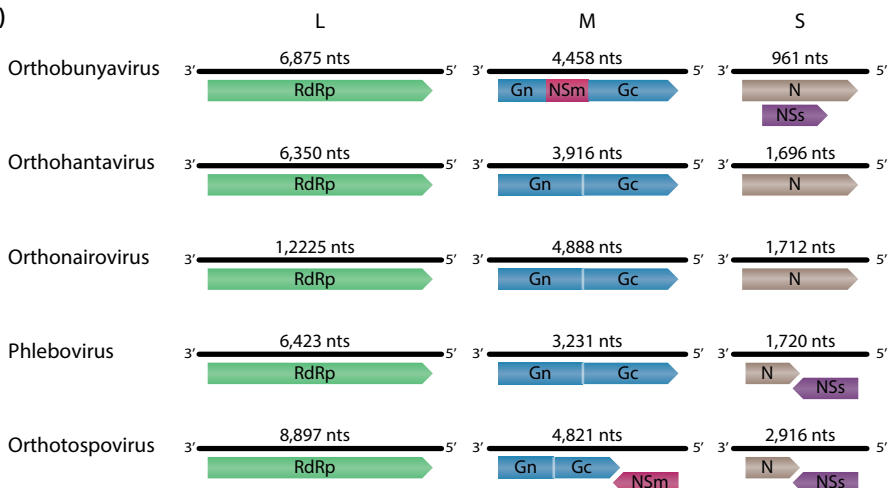
Figure 2.1 – Particle morphology and genome organization of tospoviruses. A) Schematic of an enveloped virus particle of a tospovirus. **B)** Organization of the tripartite RNA genome of representatives from the animal-infecting bunyaviruses and plant-infecting orthotospoviruses. *Abbreviations: Gc, glycoprotein processed from the carboxy-terminus of the glycoprotein precursor; Gn, glycoprotein processed from the amino-terminus of the glycoprotein precursor; L, large RNA segment; M, middle RNA segment; N, nucleocapsid protein; NSm, nonstructural protein encoded by the M RNA segment, presents the movement protein; NSs, nonstructural protein encoded by the S RNA segment, presents the RNA silencing suppressor protein; nts, nucleotides; RdRp, RNA-dependent RNA polymerase; S, small RNA segment.*

the viral RNA dependent RNA polymerase (RdRp; 330 kDa), the cell-to-cell movement protein (NSm; 33 kDa), the precursor to the glycoproteins (GP; 127 kDa), the nucleocapsid protein (N; ~28 kDa), and a nonstructural protein (NSs; 52 kDa) (de Haan et al., 1991; Kormelink et al., 1991, 1994, 2011; Adkins et al., 1995; van Knippenberg et al., 2002). The GP precursor is processed by host cellular proteases into two mature viral Gn (58 kDa) and Gc (78 kDa) glycoproteins (n and c referring to their amino- and carboxy-terminal positions within the precursor).
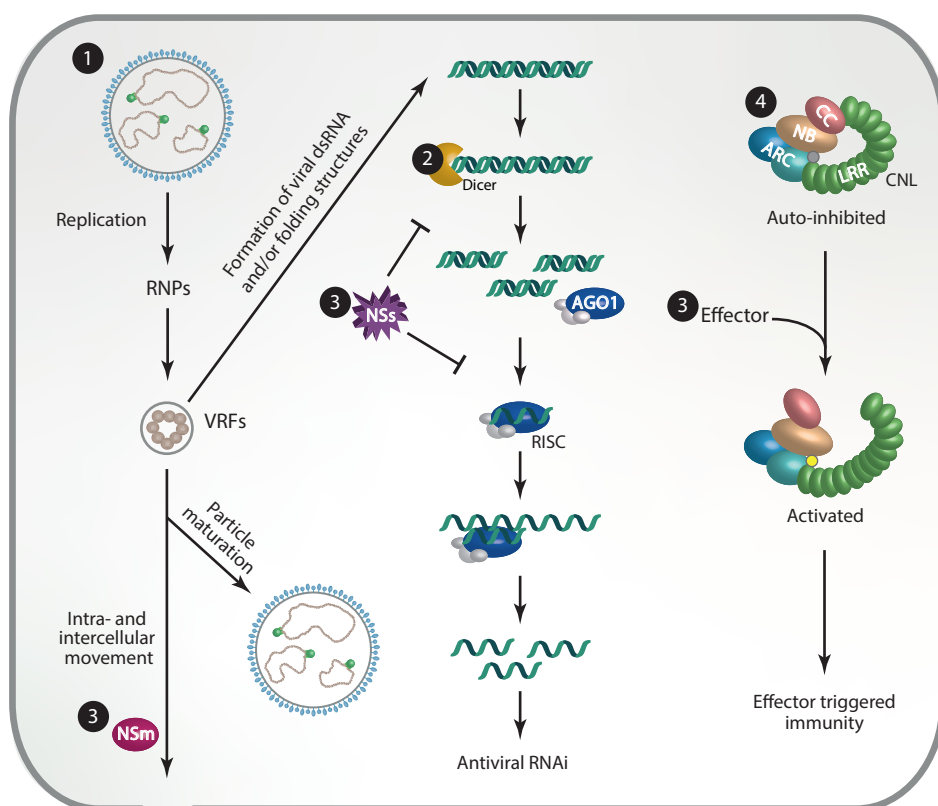
Figure 2.2 – Tospovirus life cycle and interplay with plant innate immunity. ❶ Tospoviruses depend on host cellular machineries to accomplish their life cycle. Ribonucleo(capsid)protein complexes (RNPs) recruit host cellular factors, including eEF1A, to build viral replication factories (VRFs) for genomic RNA replication and transcription. RNPs engaged in transcription snatch capped RNA leader sequences from cytoplasmic mRNAs to prime viral genome transcription. VRFs and/or N/RNPs move along the actin/ endoplasmic reticulum (ER) network to meet Gn and/or Gc at endoplasmic reticulum exit sites (ERES) and comigrate to the Golgi complex for assembly of enveloped particles. N/RNPs associating with NSm move intra- and intercellularly via a continuous ER network and through NSm-derived tubule structures to spread systemically. **❷** Viral dsRNA, produced from either viral mRNA by host RNA-dependent RNA polymerases (RdRps) or secondary folding structures of viral mRNA, triggers the first layer of plant innate immunity to defend against tospovirus invasion. dsRNA is cleaved into 21 24-nucleotide viral small interfering RNAs (vsiRNAs) by Dicer-like (DCL) endoribonucleases. Incorporation of siRNAs into Argonaute (AGO) activates the RNA-induced silencing complex (RISC) and enables the sensing and degradation of viral target RNAs. **❸** To counteract antiviral RNAi and successfully establish a viral infection, tospoviruses produce nonstructural proteins NSs and NSm. NSs proteins sequester long dsRNA and siRNAs to prevent the biogenesis of siRNAs and their uploading into RISC, respectively. NSm facilitates movement between plant cells, which allows the systemic spread of the virus. **❹** To further combat tospovirus invasion, tomato and pepper are able to sense the effectors NSm and NSs and mount a second, robust effector-triggered immunity (ETI) mediated by their nucleotide-binding leucine-rich repeat (NLR) receptors, Sw 5b and Tsw, respectively. *Abbreviations: CC, coiled-coil domain; CNL, coiled-coil nucleotide-binding leucine-rich repeat receptor; LRR, leucine-rich repeat domain; NB, nucleotide-binding domain; NB-ARC, Apaf-1, R-protein, and CED-4 domain; NSm, nonstructural protein encoded by the M RNA segment, presenting the movement protein; NSs, nonstructural protein encoded by the S RNA segment, presenting the RNA silencing suppressor protein.*

Like all animal-infecting members of the former *Bunyaviridae*, tospoviruses are now reclassified as a separate family within the order *Bunyavirales* (Maes et al., 2018). Tospovirus virions (~80-120 nm in diameter) are spherical and enveloped by a phospholipid membrane (Figure 2.1A). The two viral glycoproteins (Gn and Gc) are embedded in the lipid membrane and form spikes on the surface. The core of virus particles contains the three genomic RNA elements tightly packaged by the N protein and a few copies of the viral RdRp into ribonucleo(capsid)proteins (RNPs), the minimal infectious unit (Kormelink et al., 2011; Oliver and Whitfield, 2016; Turina et al., 2016). Crystal structures of TSWV N were recently obtained and point toward N trimerization prior to assembly into RNPs (Guo et al., 2017; Komoda et al., 2017), in which critical binding sites for RNA localize in a cleft of the N protein and suggest deep embedding of the genomic RNAs (Li et al., 2015).

2

During a successful infection and dissemination of progeny tospoviruses, three main stages can be recognized *in planta* (Figure 2.2): (a) replication and transcription of the genetic viral elements to produce massive amounts of infectious RNPs; (b) their subsequent intra/intercellular trafficking; and (c) envelopment by Golgi membranes to support virus acquisition by thrips and dissemination into neighboring healthy crop plants. To accomplish these stages, tospoviruses have evolved to smartly utilize the host cellular machinery, but it is primarily during the first stage that the inducers of antiviral RNAi and *R* gene–based host defense are produced, namely double-stranded RNA (dsRNA) and viral effector proteins, respectively (Figure 2.2).

Tospovirus Genome Replication and Transcription

For tospoviruses, which have a negative/ambisense RNA genome, RNPs present the minimal infectious unit to initiate a multiplication cycle. Replication and transcription of tospoviruses occur in the cytoplasm, but purified RNPs of TSWV, containing the viral RdRp, also support both replication and transcription *in vitro* (Adkins et al., 1995; van Knippenberg et al., 2002). *In vitro* transcription–replication requires the presence of translational machinery, not to support ongoing viral protein synthesis but for ribosome scanning of nascent viral transcripts to prevent premature transcription termination (van Knippenberg et al., 2002, 2004). Eukaryotic elongation factor 1A (eEF1A) has been identified to play an important role in the enhancement of TSWV RNA replication and transcription (Komoda et al., 2014). In *F. occidentalis* thrips, a transcription factor (FoTF) associates with TSWV RdRp, which stimulates replication but not transcription. This factor is also found to turn human cell lines permissive to TSWV replication (de Medeiros et al., 2005).

The 5′ ends of (subgenomic) TSWV mRNAs contain a non-viral, capped nucleotide (nt) sequence of ~12-20 nt (Kormelink et al., 1992; van Poelwijk et al., 1996). These sequences derive from host cellular mRNAs by a process called cap snatching, as first discovered, and described for the influenza virus (Bouloy et al., 1978). During this process, the viral transcriptase complex binds to the 5′ cap structure of a host mRNA and performs a

cleavage further downstream by an endonuclease activity encompassed in the viral RdRp. The resulting capped RNA leader aligns at the 3′ end of a viral genomic RNA template and is elongated by the viral RdRp to generate messenger transcripts.

During *in vitro* transcription of TSWV, using either purified virus particles or RNPs, reticulocyte lysate also provides a source of capped RNA (globin mRNAs) (van Knippenberg et al., 2002). During *in vivo* genome transcription, the virus is able to take any available source of capped RNA. This can also originate from a coinfecting, cytoplasmic-replicating RNA virus like Alfalfa mosaic virus (AMV) (Duijsings et al., 1999). Extensive *in vitro* and *in vivo* transcription studies have shown that TSWV prefers capped RNA leaders with the ability for multiple base pairing to the 3′ end three residues of the viral RNA template, and these outcompete similar leaders with only one base pairing residue (van Knippenberg et al., 2005). Transcription studies with influenza A (Geerts-Dimitriadou et al., 2011a, 2011b) and tenuiviruses (Yao et al., 2012), now classified as the plant-infecting members of the *Phenuiviridae* within the *Bunyavirales*, have revealed similar findings and provide support for a paradigm for cap snatching generic to all segmented, negative-strand RNA viruses that employ cap snatching, in which (multiple) base pairing between capped RNA leaders and the viral genomic RNA template promotes their usage during genome transcription initiation.

Although the endonuclease domain has not been functionally mapped within the TSWV RdRp, crystal structure analysis of the influenza virus polymerase PA subunit has pointed toward the identification of the endonuclease domain (Yuan et al., 2009). Similar analysis of the La Crosse orthobunyavirus RdRp, combined with biochemical assays using RdRp mutants and a structure-based sequence alignment, has revealed that the endonuclease domain of the RdRp is highly conserved among all bunyaviruses, including plant-infecting tospoviruses, emaraviruses, and tenuiviruses, and maps to the N-terminus of the RdRp protein (Reguera et al., 2010).

Whereas influenza virus replicates in the nucleus and snatches capped RNA leaders from nascent RNA chains produced by the nuclear RNA polymerase II system (Chan et al., 2006), TSWV replicates in the cytoplasm and snatches capped RNA leaders from the existing pool of cytoplasmic host cellular messengers. Where and how this takes place is not known, but studies with the related Sin nombre Hantavirus (SNV, a former joint member of the *Bunyaviridae*) point toward cytoplasmic RNA processing (P) bodies because the SNV N protein binds to 5′ caps with high affinity and colocalizes with P bodies (Mir et al., 2008).

Tospovirus Intracellular and Intercellular Movement

Once cytosolic TSWV RNPs are produced, cellular machineries for macromolecular transport are hijacked to take RNPs to the plasmodesmata for spreading into neighboring/ systemic plant tissue (via intercellular trafficking) or to the endoplasmic reticulum (ER) and

Golgi, where particle maturation occurs (via intracellular trafficking). During intracellular movement, TSWV N/RNPs form motile cytoplasmic protein bodies that depend on the actin cytoskeleton for its movement and are driven by the myosin XI-K (Feng et al., 2013; Ribeiro et al., 2013). Application of latrunculin B, an actin-depolymerizing agent, not only abolishes the intracellular movement of N inclusions but also strongly inhibits the local and systemic movement of TSWV in tobacco (Feng et al., 2013). In the presence of cytochalasin D, a functionally similar drug, but not colchicine (a microtubule-depolymerizing agent), TSWV N protein bodies are smaller but appear more abundantly, indicating the importance of the actin skeleton in the formation of N inclusions (Ribeiro et al., 2013). TSWV N/RNP bodies also move along the ER membrane network (Feng et al., 2013). The fact that TSWV RNPs play a critical role in viral replication leaves open the possibility for a model in which viral replication factories (VRFs) are not stationary but rather are constantly mobile within the cells. The VRFs moving along the actin/ER network can easily collect cellular host factors required for viral RNA replication–transcription. Trafficking of VRFs along the actin/ER network can also greatly facilitate the envelopment of TSWV RNPs by the ER or Golgi membrane. The observation that the TSWV N protein interacted with ER-resident Gc and concentrated at ER export sites before being rescued by Gn to the Golgi complex (an actin-dependent transport process) nicely fits into this model (Ribeiro et al., 2013).

During the early stages of infection, RNPs transiently associate with the movement protein NSm. Although these complexes can be widely distributed in the cytoplasm, they ultimately concentrate near and at plasmodesmata (Kormelink et al., 1994; Storms et al., 1995, 2002; Soellick et al., 2000) to assist in cell-to-cell and long-distance movement of infectious RNPs through plasmodesmata. Owing to the action of the NSm movement protein, NSm forms a tubule structure within the plasmodesmata to guide the transport of RNPs into neighboring healthy cells (Storms et al., 1995; Li et al., 2009). Plasmodesmata containing NSm also exhibit an enlarged size exclusion limit (Storms et al., 2002). NSm is also expressed in viruliferous thrips but is not considered to play a role in the infection cycle in thrips (Storms et al., 1995). Because animal-infecting bunyaviruses lack a homolog of NSm, the NSm protein reflects the adaptation of TSWV to infect plant hosts.

Cell-to-cell movement of plant viruses is considered a relatively conserved process. As observed with viral movement proteins from other viruses, the TSWV NSm protein is also able to transcomplement movement-deficient viruses that have a plus-strand RNA genome, e.g., Cucumber mosaic virus (CMV), Tobacco mosaic virus (TMV), and AMV, in their cell-to-cell and long-distance movement (Lewandowski and Adkins, 2005; Li et al., 2009; Peiró et al., 2014; Shen et al., 2014). The TSWV NSm protein tightly associates with the ER membrane and can move by itself from cell to cell along the ER network (Feng et al., 2016). In plants, the ER interconnects between neighboring cells via plasmodesmata and forms a continuous ER membrane network throughout the entire plant (Stefano et al., 2014). Mutations within NSm that impair its ER association or pharmacologically disrupt

2

the ER network severely inhibit the cell-to-cell movement of NSm. Likewise, *Arabidopsis thaliana* rhd3 mutants containing an impaired ER network show a severe reduction of the intercellular movement of TSWV NSm and, as a result, show reduced levels of systemic viral infection (Feng et al., 2016). Although NSm facilitates cell-to-cell movement of infectious RNPs in a tubule-guided manner via plasmodesmata, an interplay between NSm and ER clearly plays an important role in the intra- and intercellular movement of viral RNPs.

Tospovirus Particle Maturation

During later stages of infection, TSWV RNPs become surrounded by the lipid membrane of the Golgi apparatus. In contrast to the animal-infecting bunyaviruses, whose RNPs bud into lumenized Golgi stacks, TSWV RNPs are enwrapped by an entire Golgi stack, leading to the formation of double-enveloped virus particles (Kikkert et al., 1999). These double-enveloped particles subsequently fuse with each other and ER membranes and generate large vesicles containing accumulating amounts of mature singly enveloped virus particles (Kikkert et al., 1999). The viral glycoproteins Gn and Gc are processed from the precursor protein on ER membranes. From there, Gn continues trafficking toward the Golgi complex as a monomer or dimerized with Gc. Upon single transient expression, both glycoproteins are able to induce membrane curving of ER/Golgi membranes (Ribeiro et al., 2008), a feature typical of reticulons, i.e., proteins that predominantly reside in the ER and play an important role in promoting membrane curvature (Voeltz et al., 2006). Whether membrane curving reflects the formation of a spherical enveloped virus particle remains speculative.

Gc expressed on its own is retained in the ER and requires dimerization with Gn to escape and move toward the Golgi complex (Ribeiro et al., 2008). ER arrest of Gc is not a matter of improper folding but is due to its (relatively short) transmembrane domain (TMD). The exchange of this domain by the larger TMD of Gn allows Gc to escape from ER arrest and move toward the Golgi (Ribeiro et al., 2009b). This observation fits with the idea that the TMD length of membrane proteins has a major influence on their final destination within the endomembrane system, with shorter TMDs restricting the proteins to the ER (Brandizzi et al., 2002).

During the early stages of glycoprotein synthesis and processing, the glycoproteins localize at ER export sites (ERESs), distinct foci in the ER, from where concentrated glycoproteins continue in their anterograde trafficking (Stefano et al., 2006; Ribeiro et al., 2008). This observation suggests that Gn recruits and subsequently concentrates ER-resident Gc at ERESs, from which anterograde transport to the Golgi complex continues. Inhibition of COPII vesicle formation at ERESs aborts trafficking of Gn and Gn-Gc dimers toward the Golgi complex, which indicates that viral glycoprotein trafficking from ER to Golgi involves COPII-dependent vesicle transport (Ribeiro et al., 2009b).

TSWV N/RNPs rely on ER/Golgi membrane for their envelopment. Intracellular movement of TSWV N/RNPs along the ER membrane network is dependent on actin and myosin (Feng et al., 2013; Ribeiro et al., 2013). The Golgi complex is known to move as a mobile unit along the ER membrane network in an actomyosin-dependent manner as well, meanwhile picking up ER cargo from ERESs in a stop-and-go mode (Boevink et al., 1998; Nebenfu̇hr et al., 1999; daSilva et al., 2004). In plants, TSWV N protein colocalizes and interacts with both Gn and Gc at the interface of ER and Golgi (Ribeiro et al., 2009a, 2013). The cytosolic N protein of TSWV is able to recruit ER-resident (reticular) Gc proteins to ERESs through interaction with the cytoplasmic tail of Gc (Ribeiro et al., 2013). From there N/RNPs-Gc complexes are rescued by an interaction with Gn and exit the ER. Alternatively, N/RNPs interact with preformed Gn-Gc heterodimers either at ERES, and traffic to the Golgi, or at the Golgi stacks where particle assembly matures (Ribeiro et al., 2013).

2

The actomyosin-dependent movement of RNPs along the ER network ensures RNPs meet ER-associated NSm during the early stages of infection, leading to cell-to-cell movement, or encounter Gn-Gc complexes at either ERESs or Golgi complexes during late stages of infection, triggering membrane envelopment.

RNA interference: the first layer of plant defense against tospovirus

RNA silencing or RNA interference (RNAi) is known to play a key role in antiviral defense in plant cells. The mechanism is triggered by the formation of dsRNA, which, during a viral infection, arises from viral replication intermediates or secondary folding structures in viral (m)RNA. These are processed by Dicer-like (DCL) endoribonucleases into (primary) 21-24 nt small interfering RNAs (siRNAs). One strand of siRNA loads into an RNA-induced silencing complex (RISC) and guides the RISC to viral RNA target molecules, leading to degradation of the target (Guo et al., 2019). Aberrant RNAs, resulting from target RNA cleavage, are either further degraded by the cellular decay machinery localized in P bodies or used as a substrate by host-encoded RdRps for further amplification of the antiviral RNAi response (Hedil and Kormelink, 2016).

Small Interfering RNAs Derived from Tospovirus

Tospovirus-specific siRNAs (viral siRNAs (vsiRNAs)) from infected *Nicotiana benthamiana* and tomato show the production of 21 and 22 nt vsiRNAs from all three RNA segments (Mitter et al., 2013; Margaria et al., 2015). The vsiRNA profiles from both hosts are quite similar and show most reads from the M and S RNA and only low amounts from the L RNA. In tomato, vsiRNAs have been found that potentially target host genes involved in many pathways, including those related to plant-pathogen interactions (Ramesh et al., 2017). Only a small amount of vsiRNAs is derived from the intergenic region (IGR) sequence of the ambisense M and S RNA segments. Although synthetic IGR transcripts are recognized and processed into small RNAs by Dicer from *Drosophila melanogaster* embryo extracts,

during a natural infection the predicted folding structures within the IGRs are likely not accessible because of masking by other viral/host factors (Hedil et al., 2014).

RNAi-Mediated Resistance Against Tospovirus

In the past, pathogen-derived resistance has been engineered against TSWV. Although transgenic plants expressing the N protein protected against a tospovirus infection (Gielen et al., 1991; MacKenzie, 1992; Prins, 1995), resistance was also obtained after expression of untranslatable (partial) N, NSm, NSs, or RdRp genes, demonstrating an RNA-mediated defense mechanism (Pang et al., 1993; Prins et al., 1997; Jan et al., 2000; Sonoda and Tsumuki, 2004; Peng et al., 2014; Yang et al., 2015; Yazhisai et al., 2015). Transgenic resistance holds against the homologous virus only and not any other distinct or even relatively closely related tospoviruses (Haan et al., 1992; Hassani-Mehraban et al., 2009). Stable tobacco transformants expressing an NSm transgene are completely immune at the plant level, but protoplasts collected from these still support replication of TSWV, indicating that antiviral RNAi triggered by a tospovirus infection acts on viral messengers and not (anti)genomic RNA strands (Prins et al., 1997). Viral (anti)genomic RNAs are not being targeted, likely because these (anti)genomic RNAs are only present in RNPs and tightly encapsidated with N protein and are thereby protected from DCL cleavage.

Suppression of antiviral RNAi: viral counter-defense

As a viral counter-defensive strategy, almost all plant viruses encode RNA silencing suppressor (RSS) proteins to inhibit key steps of host antiviral RNAi. Viral RSSs suppress RNA silencing pathways through sequestering dsRNA to inhibit biogenesis of siRNA, sequestering siRNA duplexes to prevent the loading of small RNA into the Argonaut (AGO) proteins, or inhibiting antiviral activity of the host RISC machinery (Guo et al., 2019).

Tospovirus Encodes an RNA Silencing Suppressor to Counteract Host RNAi Antiviral Defense

Using a GFP-based transient silencing suppression assay, the NSs protein of TSWV and other tospoviruses such as Impatiens necrotic spot virus (INSV), Groundnut ringspot virus (GRSV), and Tomato yellow ring virus (TYRV) has been identified as the RSS (Takeda et al., 2002; Bucher et al., 2003; Hassani-Mehraban et al., 2009; Schnettler et al., 2010). Electrophoretic mobility shift assays with NSs from different tospoviruses, either using NSs-containing crude plant/insect cell extracts or *Escherichia coli*–expressed purified NSs, have shown that tospoviral NSs proteins are able to bind dsRNA independent of size; siRNAs and microRNA (miRNA)/miRNA* duplexes as well as long dsRNA (Schnettler et al., 2010; Hedil et al., 2017). In agreement with the binding of long dsRNA, TSWV NSs protein inhibits the cleavage of long dsRNA by Dicer *in vitro* and suppresses gene silencing induced by an inverted repeat GFP (Schnettler et al., 2010; Hedil et al., 2017). Together, these findings suggest that tospovirus NSs exerts RNAi suppressor activity by sequestering long and small dsRNA to prevent the biogenesis of small RNAs, their uploading, and the activation of antiviral RISCs. Using alanine substitution analysis of the NSs protein, both the N and

C-terminal domain of NSs have been shown to be required for RNA silencing suppression, but mutations introduced in the N-terminus almost always rendered NSs dysfunctional in the suppression of local GFP silencing assays (de Ronde et al., 2014b). TSWV NSs also contains a GW/WG motif, a dipeptide sequence that has been reported to be involved in the interaction of several RSS proteins with AGO1 (Giner et al., 2010). Mutation of this motif within NSs also impaired its (local) RNA silencing suppression activity (de Ronde et al., 2014b) and implies a possible interference at AGO1-RISC. Mutations introduced into NSs affect the ability to suppress local RNAi and do not always render this protein completely dysfunctional. Several dysfunctional NSs mutants are still able to suppress GFP silencing in systemic leaves, most likely because of their ability to sequester siRNAs from local leaves and to prevent their movement and subsequent activation of antiviral RISC in systemic tissue. This also applies to the NSs GW/WG mutant and supports the idea that this mutant is functionally hampered from suppressing local GFP silencing further downstream from the biogenesis (and binding) of siRNAs (Hedil et al., 2015).

The ability of NSs to sequester siRNAs independent of sequence also implies that it might interfere in the antiviral RNAi response toward a coinfecting virus. This is nicely exemplified with stable transformants of *N. benthamiana* containing a partial N-gene sequence of TYRV-t (tomato strain) that are resistant to this virus but not to TYRV-s (soybean strain). When these transformants are co-challenged with TYRV-t and TRYV-s, a systemic infection of both viruses is observed. Delivery of TSWV-s NSs protein from a Potato virus X (PVX) replicon prior to TYRV-t infection also allows the latter to break the resistance. This shows that even engineered resistance that provides total immunity can be directly broken when the virus target strain is trans-complemented by a viral RNAi suppressor from a coinfecting related strain, or possibly from an unrelated virus as well (Hassani-Mehraban et al., 2009). NSs has been observed to trans-complement (heterologous) RNAi-suppressor-deficient viruses as well. A nice example of this is observed in Turnip mosaic virus (TuMV). When the NSs gene was inserted into a HcPro-defective TuMV-GFP, NSs compensated for the functional loss of HC-Pro and supported a systemic TuMV infection of the plant (García-Cano et al., 2006; Garcia-Ruiz et al., 2018).

Tospovirus NSs Suppresses RNAi in insects

Because RNAi also acts as an antiviral defense mechanism in arthropods, NSs may also suppress RNAi in insects and even ticks. Evidence for this has come from studies in which TSWV NSs recombinantly expressed from a baculovirus enhanced the replication of this virus in Sf9 insect cells (Oliveira et al., 2011) and increased its virulence in caterpillars (de Oliveira et al., 2015). In tick and mosquito cells, expression of TSWV NSs from a Semliki forest virus replicon impaired RNAi induced by a Semliki forest virus replicon expressing luciferase (Blakqori et al., 2007; Hedil and Kormelink, 2016).

During infection of arthropods with arboviruses, the activity of RNAi in the midgut epithelium has been postulated to play an important role in the midgut barrier and vector

competence (Lan et al., 2016). In thrips, antiviral RNAi has not yet been investigated, but it is generally assumed that the virus encounters an antiviral RNAi response and requires NSs to counter-defend during propagative transmission. Recent studies by Margaria et al. (2014) pointed toward an important role for TSWV NSs during acquisition/transmission by *F. occidentalis*. A TSWV isolate containing a truncated NSs protein and compromised in its ability to suppress RNAi could still be acquired by *F. occidentalis*. However, this virus was not able to establish a persistent infection and was not transmitted by *F. occidentalis* (Margaria et al., 2014), suggesting that the observed reductions in vector competence and virus titers are likely due to the lack of RNAi suppression by these NSs-defective isolates.

Effector-triggered immunity: a robust second layer of plant defense against tospovirus invasion

Plants employ diverse intracellular innate immune receptors to recognize invading pathogens and to induce host defense (Jones et al., 2016). The largest class of intracellular receptors is presented by the nucleotide-binding leucine-rich repeat receptors (NLRs). During pathogen invasion, plant NLRs recognize specific pathogen effectors to trigger so-called effector-triggered immunity (ETI) (Caplan et al., 2008b; Jones et al., 2016), which often comes along with a programmed cell death (PCD)–based hypersensitive response (HR) at the initial infection site. Effectors recognized by NLRs are referred to as elicitors or avirulence (Avr) determinants. Based on the difference at the N-terminus, plant NLRs are further classified into Toll/Interleukin-1 (TIR)-NLRs (TNLs) and coiled-coil (CC)-NLRs (CNLs).

Tomato Immune Receptor Sw-5b-Mediated Resistance Against Tospoviruses

The tomato *Sw-5* gene is the most commonly used resistance (*R*) gene in tomato resistance breeding to control tospoviruses. Two groups have independently cloned this single dominant *R* gene (Brommonschenkel et al., 2000; Spassova et al., 2001). Sw-5b belongs to the CNL-type immune receptors, which contain a CC domain, a central nucleotide-binding adaptor shared by **A**paF-1, **r**esistance proteins, a **C**ED-4 (NB-ARC) domain, and a C-terminal leucine-rich repeat (LRR) domain (Brommonschenkel et al., 2000; Spassova et al., 2001; Chen et al., 2016; De Oliveira et al., 2016).

Sw-5b confers broad-spectrum resistance to American-type tospoviruses through recognition of a conserved 21-amino acid (aa) epitope in NSm

Currently, almost thirty recognized and tentative tospovirus species are known (Turina et al., 2016), and these classify into American and Euro-Asian clades based on their geographic origin and the amino acid sequence of the N protein. Analyses of *Sw-5b* stable transgenic *N. benthamiana* plants and (near-isogenic) tomato lines carrying *Sw-5b* showed that *Sw-5b* confers broad-spectrum resistance to various American-type tospoviruses, but not to Euro-Asian-type tospoviruses (Figure 2.3) (Leastro et al., 2017; Zhu et al., 2017).
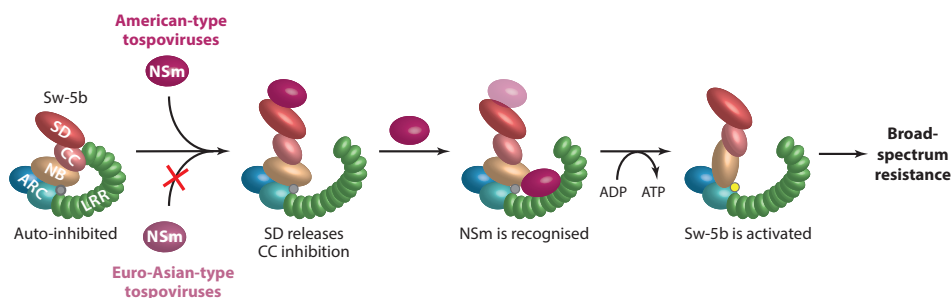
Figure 2.3 – Model for Sw-5b-mediated resistance against tospoviruses. Sw-5b confers broad-spectrum resistance to various American-type tospoviruses, but not to Euro-Asian-type tospoviruses, through the recognition of a conserved 21-aa (amino acid) peptide region of the nonstructural protein NSm. NSm binds to the extended SD of Sw-5b, which releases the inhibition of the CC domain on the central nucleotide-binding leucine-rich repeat domain (NB-ARC-LRR). The NB-ARC-LRR further recognizes NSm, exchanges ADP (gray dot) for ATP (yellow dot), and switch activates the receptor, leading to a robust defense response against tospoviruses. *Abbreviations: CC, coiled-coil; CNL, coiled-coil nucleotide-binding leucine-rich repeat receptor; LRR, leucine-rich repeat domain; NB, nucleotide binding domain; NB-ARC, Apaf-1, R-protein, and CED-4 domain; NSm, nonstructural protein encoded by the M RNA segment, presenting the movement protein; NSs, nonstructural protein encoded by the S RNA segment, presenting the RNA silencing suppressor protein; SD, Solanaceae domain.*

Sw-5b resistance is triggered by the cell-to-cell movement protein NSm (Lopez et al., 2011; Hallwass et al., 2014; Peiró et al., 2014; Leastro et al., 2017; Zhu et al., 2017). The activation does not rely on the presence of motifs required for plasmodesmata targeting, tubule formation, or cell-to-cell movement (Zhao et al., 2016; Leastro et al., 2017). Instead, the central domain, specifically a 21-aa region positioned at aa 115-135 in TSWV NSm (NSm21), is sufficient to trigger Sw-5b-mediated HR. This 21-aa region is highly conserved among the American-type tospoviruses only. Importantly, the Euro-Asian-type tospovirus NSm protein can be converted into an inducer of Sw-5b-mediated HR by introducing the NSm21 peptide to their sequence (Zhu et al., 2017). Mutations C118Y and T120N, which have been reported in two natural TSWV NSm resistance-breaking (RB) isolates, map within this NSm21 region. Introduction of these mutations into either the full-length NSm of a resistance-inducing (RI) strain or the NSm21 peptide only impairs its ability to elicit Sw-5b-mediated HR (Leastro et al., 2017; Zhu et al., 2017). The mutations in this conserved peptide region have been shown to strongly affect intercellular movement of TSWV NSm (Feng et al., 2016). This suggests that the conserved NSm21 region plays a critical role in viral cell-to-cell movement. Introduction of the T120N mutation into NSm compromises the avirulence and leads to a fitness cost during viral infection (Peiró et al., 2014).

Plant pattern-recognition receptors (PRRs) typically recognize conserved pathogen-associated molecular patterns (PAMPs) to provide broad-spectrum pathogen resistance, whereas plant NLRs generally recognize strain-specific pathogen effectors and confer race-specific resistance. The findings showing that Sw-5b NLR recognizes a conserved NSm21 peptide region to confer broad-spectrum resistance to American-type tospoviruses strongly support the notion that this conserved NSm21 peptide acts as a

PAMP (Zhu et al., 2017), similar to the plant PRRs FLS2 and EFR, which (extracellularly) recognize a conserved 22-aa peptide of flagellin (flg22) (Chinchilla et al., 2006) and an 18-aa peptide (elf18) of elongation factor Tu (EF-Tu) (Zipfel, 2014), respectively, and provide broad-spectrum resistance to bacterial invasions. The intracellular triggering of Sw-5b NLR by NSm21 provides the first case of a plant NLR conferring broad-spectrum resistance by means of a PAMP-like structure–triggered immune response. The Mi-1.2 NLR, a close Sw-5b homolog from tomato, also provides broad-spectrum resistance to four completely different pests: nematodes, aphids, whiteflies, and psyllids (Milligan et al., 1998; Vos et al., 1998; Nombela et al., 2003; Casteel et al., 2006). Although speculative, it is not unlikely that triggering the broad-spectrum Mi1.2-mediated resistance, as with Sw-5b, might also involve a conserved PAMP-like structure in the cognate pest effectors.

Multilayered auto-inhibition and activation of Sw-5b

Plant NLR immune receptors typically function as molecular switches to initiate robust defense responses against pathogen invasion. The central NB-ARC domain of NLRs binds to ADP or ATP to switch from an auto-inhibited "off" state to an activated "on" state (van Ooijen et al., 2008; Takken and Tameling, 2009; Maekawa et al., 2011; Williams et al., 2011) The CC domain and the LRR domain suppress NB-ARC nucleotide exchanging through intracellular domain interaction and maintain the NLR protein in an auto-inhibited state. Only once the presence of the pathogen is perceived does the status change to an activated one. Although Sw-5b belongs to the CNLs, like other *R* genes from *Solanaceae*, it contains an extended N-terminal Solanaceae domain (SD) (Mucyn et al., 2006). With four domains, Sw-5b has evolved a more complex multilayered mechanism to regulate and switch between its auto-inhibited, or "off," and activated, or "on," state (Chen et al., 2016). In the absence of NSm, Sw-5b LRR suppresses the NB-ARC domain and keeps the NB-ARC-LRR in an auto-inhibited state. The N-terminal SD and the CC domain actively suppress the NB-ARC-LRR to prevent activation of the receptor and concomitant onset of PCD (Chen et al., 2016; De Oliveira et al., 2016). In the presence of NSm, the N-terminal SD, CC, and LRR-mediated inhibition of NB-ARC are sequentially relieved and switch Sw-5b to an activated form, leading to a (visual HR) defense response (Figure 2.3) (Chen et al., 2016; De Oliveira et al., 2016).

Domain swaps between Sw-5b and Sw-5 paralogs from *Solanum peruvianum* or orthologs from susceptible *Solanum lycopersicum* (domesticated Heinz Tomato) have shown that LRR domains can be exchanged and revert susceptible/nonfunctional homologs into functional resistance-conferring copies (De Oliveira et al., 2016; Zhu et al., 2017). Identification and characterization of the NSm binding site in the LRR domain by means of site-directed mutagenesis and domain swaps discerned four polymorphic sites in the Sw-5b LRR domain that are critical for NSm or NSm21 recognition. Structure modeling of the Sw-5b NB-ARC-LRR revealed that these four polymorphic sites are all clustered and surface exposed (Zhu et al., 2017). Furthermore, the R927 contact site between the NB-ARC and LRR domain was also identified; it is required to maintain Sw-5b NB-ARC-

LRR in an autoinhibited state and is located adjacent to the four polymorphic sites in the LRR domain of the homology model. Considering that the R927 contact site and the NSm21 binding sites are located next to each other on the LRR domain surface, the recognition of NSm21 by Sw-5b LRR likely disrupts the R927 contact site and weakens the intramolecular interaction between the NB-ARC and LRR domain, thus translating NSm21 ligand recognition into activation of the receptor (Zhu et al., 2017).

Although many plant NLRs recognize pathogen effectors via their N-terminal domain (Dodds et al., 2001; Collier and Moffett, 2009b; Jones et al., 2016) and switch to an activated state in the presence of the effector, the situation with Sw-5b seems more complex. Recently, Li et al. (2019) demonstrated that, in addition to the LRR domain, the extended N-terminal SD of Sw-5b also plays a critical role in sensing NSm and enhances the ability of NB-ARC-LRR to detect a low amount of NSm. These findings suggest that Sw-5b NLR has evolved an extra pathogen sensor (SD) next to the main switch activator (NB-ARC-LRR) in one single immune receptor and uses two distinct domains to detect the viral movement protein NSm. In this way, Sw-5b has adopted a two-step recognition mechanism to significantly enhance its sensitivity to tospovirus infections.

Evolutionary selection of Sw-5b homologs from wild tomato species in South America

The resistance allele Sw-5b originates from a wild species of *S. peruvianum* from South America. American-type tospoviruses (de Avila et al., 1990; Bezerra et al., 1999; Pozzer et al., 1999; Williams et al., 2001) and tomato species (Sato et al., 2012; Lin et al., 2014) both originate from South America. The observation that Sw-5b confers resistance to the American-type tospoviruses only supports the idea that Sw-5b and these tospoviruses coevolved. The importance of the four polymorphic sites (M3–6) and R927 within Sw-5b LRR for NSm recognition and activation of NB-ARC-LRR is emphasized by the observed natural variation among Sw-5b homologs analyzed from ~90% of all wild tomato species from South America (Zhu et al., 2017). The polymorphic site 8 (M8) in Sw-5 SD (Li et al., 2019a) has recently been identified and has also shown to be critical for NSm recognition (Li et al., 2019a). Natural covariation analysis showed that only those Sw-5 homologs containing conserved M8 sites in the SD and conserved M3–6 sites in the LRR domain were able to induce HR in the presence of NSm or NSm21 and confer resistance to TSWV infection (Zhu et al., 2017; Li et al., 2019a). These findings support the idea that co-selection of the critical NSm recognition sites in both the SD and LRR domain is necessary for conferring the resistance against tospoviruses.

Pepper Immune Receptor Tsw-Mediated Resistance Against Tospovirus

In contrast to the broad-spectrum resistance mediated by *Sw-5b* to various American-type tospoviruses, *Tsw*, originating from *Capsicum chinense*, only confers resistance to TSWV isolates (Boiteux and de Ávila, 1994; Boiteux, 1995). Similar to *Sw-5b*, HR is triggered

in *Tsw* cultivars upon TSWV infection to prevent local and systemic viral spread (Boiteux and de Ávila, 1994).

Tsw recognizes tospovirus RNA silencing suppressor to elicit a hypersensitive response

In the past, the identity of the effector of Tsw has been debated. Margaria et al. (2007) obtained indirect evidence indicating that NSs triggered Tsw-mediated resistance. Lovato et al. (Lovato et al., 2008), however, found that the N protein triggered PCD in *C. chinense* plants carrying the *Tsw* gene after transient expression from a PVX vector. In 2013, de Ronde et al. indisputably identified NSs as the effector of Tsw-mediated resistance. Using *Agrobacterium tumefaciens* transient expression assays (ATTAs), de Ronde and colleagues found that NSs, but not N, triggers HR in *Tsw* pepper plants

Considering that NSs from TSWV acts as an Avr factor as well as an RSS suggests an arms race during coevolution between the virus and plant innate immunity. Analysis of a series of NSs mutants indicated the importance of the N-terminus of the protein for both functions, although some mutations do uncouple the RSS activity and HR triggering of NSs (de Ronde et al., 2014b). Chimera of TSWV NSs and GRSV NSs, two closely related tospoviruses, do not trigger Tsw-mediated resistance, indicating that Avr functionality cannot simply be transferred. Furthermore, the NSs GW/WG motif mutant that was previously shown to be impaired in its local but not systemic RNA silencing suppression activity also lost the ability to trigger Tsw-mediated resistance (de Ronde et al., 2014b; Hedil et al., 2015).

Tsw is a CC-NLR immune receptor

Screening of *Capsicum spp*. Lines identified three *C. chinense* accessions (PI152225, PI159236, and 'Panca') resistant to TSWV (Black, 1991; Boiteux, 1995). Allelic relationship analysis of these lines showed that resistance was provided by the same gene (Boiteux, 1995). Subsequent genetic mapping of the *Tsw* locus linked the gene to the marker SCAC568 c. on chromosome 10 (Jahn et al., 2000). Meanwhile, the *Tsw* gene has been cloned using a comparative genome-guided fine-mapping approach. Within a 295-kb region of chromosome 10, eight intact CNLs are found, of which only one candidate induced HR upon co-expression with NSs in *N. benthamiana*. Transient and stable expression of this candidate in *N. benthamiana* plants conferred resistance to TSWV and was therefore confirmed as *Tsw* (Kim et al., 2017a).

The 6,351-bp-long *Tsw* gene has nine exons and encodes an NLR protein of 2,116 amino acids with a CC domain, a nucleotide-binding domain, and eight LRRs (Kim et al., 2017a). Of these eight LRRs, seven are highly similar and differ by only a few amino acids (Chapter 3 and 4). The susceptible allele *tsw* lacks four of these similar LRRs and does not induce an HR response. Tsw also encodes a different C-terminal end than *Tsw*. As the LRR domain of NLRs is generally associated with the recognition of effectors (Dodds et al., 2001; Collier

and Moffett, 2009b; Jones et al., 2016), the absent LRRs, as well as the different C-terminal domain in tsw, are likely involved in the recognition of NSs. The role of the LRR as effector recognizer is further illustrated upon comparison of *Tsw* and Pvr4, an NLR receptor from *Capsicum annuum* with high structural homology to *Tsw* that confers resistance against potyviruses upon recognition of the potyviral RdRp (Kim et al., 2017a).

Strains of Tomato spotted wilt virus that break Tsw-mediated resistance

Various *Tsw* RB strains of TSWV have been reported (Margaria et al., 2007; de Ronde et al., 2013, 2019; Almási et al., 2015; Ferrand et al., 2015; Jiang et al., 2017a). Almási et al. (2017) discovered that a single-point mutation at amino acid position 104 of wild-type NSs was sufficient to break *Tsw* resistance. When position 104 of wild-type and RB NSs was changed, their phenotypes reversed, which suggests that this point mutation likely affects the (in)direct interaction with Tsw. In addition to this single point mutation, other mutations associated with RB strains were found that localized further toward the C-terminus of NSs.

At temperatures above 30-32 °C, the Tsw NLR is no longer functional (Chung et al., 2012; de Ronde et al., 2019) and any TSWV isolate, including a wild-type (RI) TSWV, is able to establish a systemic infection. Besides the common RB strains, De Ronde et al. (2019) discovered another class of strains that break Tsw mediated resistance at 28 °C, a temperature at which Tsw is still functional. At standard greenhouse conditions (21-22 °C), these temperature-dependent RB strains induce Tsw-mediated resistance. Temperature-shift assays indicated that *de novo* synthesis of this RB NSs protein at lower temperatures was required to induce resistance.

Perspectives

Employment of Recessive or Dominant Resistance Genes to Prevent Tospovirus Infection

Plant viruses are obligate pathogens whose life cycles are dependent on host-plant machineries. Disruption or mutation of cellular host factors essential for virus replication, transcription, or movement can result in a recessive resistance against virus infection in plants. Many recessive resistance genes against virus infection have been identified in diverse plant species, including Arabidopsis, pea, pepper, lettuce, barley, melon, white lupin, and wild tomato (Hashimoto et al., 2016). One of the best examples is eIF4E: Its isoform, eIF(iso)4E, carries a mutation in the recessive allele and fails to interact with viral VPg (viral protein genome-linked), thereby conferring resistance to potyviruses (Hashimoto et al., 2016). Later, this gene was found to confer resistance to a broad spectrum of other plant viruses as well (Takken and Tameling, 2009; Sanfaçon, 2015). Several plant proteins have now been characterized as important factors needed for plant virus life cycles (Hashimoto et al., 2016) and are potential resources of recessive resistance. For TSWV, RHD3 has been shown to play a critical role in movement (Feng et al., 2016),

whereas TSWV N/RNPs move intracellularly in an actomyosin-dependent manner. Viral replication and transcription are enhanced by eEF1A (Komoda et al., 2014). Whether these host factors present interesting targets for the development of recessive resistance against TSWV/tospoviruses remains to be further investigated.

Although several new sources of tospovirus resistance are being described (Turina et al., 2016), e.g., Sw-2, Sw-3, Sw-4, and Sw-7, a continuing search for new sources of resistance is needed. Several wild species of tomato do not possess the Sw-5b resistance allele but could confer resistance to TSWV infection and present interesting targets for further analysis (Zhu et al., 2017). Further resistance screening can be done to combat not only TSWV in tomato but also Euro-Asian-type tospoviruses in a wide range of other host plants/crops.

Breeding for Resistant Cultivars Against Tospoviruses Through Targeted Genome Editing

Clustered regularly interspaced short palindromic repeat (CRISPR)-Cas9 technology has emerged as a powerful tool for precise targeted genome editing, including gene knock-out and knock-in (Yin et al., 2017; Zhang et al., 2017). For tomato breeding, TSWV-susceptible cultivars are known to contain a Sw-5b homolog with only one or a few SNPs relative to the functional Sw-5b. Because tomato is suitable for genome editing by CRISPR-Cas9 technology (Shimatani et al., 2017; Dahan-Meir et al., 2018), susceptible/dysfunctional Sw-5 alleles can now be (ex)changed to reintroduce Sw-5b-mediated TSWV resistance. The same applies to the susceptible *tsw* allele from pepper, which lacks four LRR domains compared with the functional *Tsw*. Their reintroduction could rescue Sw-5b or Tsw-mediated resistance in susceptible tomato or pepper and simultaneously allow these cultivars to enter introgression breeding programs.

Molecular Engineering of NLR Receptors to Expand the Spectrum of Pathogen Recognition

Artificial evolution has been successfully used to expand the spectrum of pathogen recognition by plant NLRs, which is nicely exemplified with Rx1. This TIR-NLR receptor confers resistance to several strains of PVX. A natural mutation at amino acid residue 121 in the PVX coat protein (CP) sequence allows the virus to overcome Rx-mediated resistance in potato (Kavanagh et al., 1992; Goulden et al., 1993; Kohm et al., 1993). Using an in vitro artificial evolution approach, a newly generated (mutant) Rx not only recognized the non-eliciting PVX strain but also, along with other strains of PVX, the distantly related Poplar mosaic virus (Farnham and Baulcombe, 2006; Harris et al., 2013). Applying this approach to Sw-5b and Tsw is tempting, considering that RB isolates have been documented for both genes (Latham and Jones, 1998; Aramburu and Marti, 2003; Ciuffo et al., 2005; Lopez et al., 2011; Turina et al., 2016; Jiang et al., 2017a) and the narrow resistance spectrum has been observed for Tsw.

Although many plant NLRs share high amino acid sequence identity, they mostly recognize very different pathogens. Several good examples exist: Rpi-blb2 NLR shares 82% amino acid sequence identity with Mi-1.2 but confers resistance to *Phytophthora*, whereas Mi-1.2 confers resistance to a nematode and several insects (Milligan et al., 1998; Vos et al., 1998; Nombela et al., 2003; Vossen et al., 2005; Casteel et al., 2006). Rpi-vnt1.1 and Tm-2² share 72% identity but confer resistance to potato late blight and Tomato mosaic virus, respectively (Foster et al., 2009). Rx1 and Gpa (88% amino acid sequence identity) respectively confer resistance to PVX and the nematode *Globodera pallida* (van der Vossen et al., 2000). R8 shares 89% amino acid sequence identity with Sw-5b but confers resistance to potato late blight disease (Vossen et al., 2016). Pvr4 shares 66% amino acid sequence identity with Tsw, but both genes recognize different viruses (Kim et al., 2017a). It is likely that these homologs share a common ancestor but have evolved to recognize different pathogens/pests through natural selection. Recently, the Arabidopsis immune receptor RPS5 has been engineered to switch from recognizing bacteria to viruses (Kim et al., 2016). In light of this achievement, and with the availability of new and fast gene editing tools, there is now the possibility to molecularly engineer plant NLRs with a changed or expanded pathogen recognition spectrum. Additional NLRs also increase the number of genetic resources that can be used for crop resistance (introgression) breeding programs to defend against different tospoviruses or plant pathogens/insect pests.

2

Summary points

o   Tospoviruses are important plant pathogens and pose a serious threat to agricultural and ornamental industries worldwide.

o   To accomplish their propagation cycles, tospoviruses depend on host cellular machineries for (almost) every step, including replication, transcription, movement, and particle maturation.

o   As the first layer of defense, the plant immune system utilizes an RNA silencing mechanism to prevent tospovirus invasion.

o   Tospoviruses are known to encode RSSs to counteract the host's first defense layer and facilitate virus infection in plants.

o   To further combat tospoviruses, the plant immune system has evolved NLR immune receptors (e.g., Sw-5b and Tsw) to recognize different viral effectors (e.g., NSm and NSs) and induce robust ETI as the second layer of defense against tospovirus infection.

o   Sw-5b and Tsw-mediated resistance has been broken by naturally evolved new TSWV strains. The arms race between tospoviruses and the plant innate immune system is endless and drives the coevolution of host defense mechanisms and virally encoded proteins. New strategies, including targeted gene editing and artificial evolution of NLRs, can be adopted to reprogram plant innate immunity to control tospovirus diseases.

Chapter 3

Localization of the resistance gene *Tsw* on *Capsicum chinense* BAC library contigs

I.L. van Grinsven, P. Butterbach, E. Schijlen, R. Kormelink

Abstract

The coding sequence of the pepper resistance gene *Tsw*, conferring resistance to tomato spotted wilt virus (TSWV), has been previously recovered from *Capsicum chinense* PI152225. However, the non-coding sequences of this gene had not been published initially, while these potentially provide valuable information regarding the transcriptional regulation of *Tsw* via non-coding promoter and/or intron sequences. In this study, a bacterial artificial chromosome (BAC) library with 11x coverage was constructed of the *C. chinense* PI152225 genome. Starting from markers P, Q and L, known to be in close proximity of *Tsw*, genome walking was undertaken in search of contigs containing the full-length sequence of *Tsw*. Using this strategy, several BACs were collected spanning this genomic region, of which clone Cc100 contained the entire resistance gene. Sequence analysis revealed an exceptionally long LRR domain, and an overall similarity of the coding sequence to the published coding sequence of *Tsw*.

Introduction

Plants have acquired various strategies to combat the posing dangers of pests and diseases. Their multilayered defense mechanism includes everything from generic non-host resistance, involving for instance cell wall thickening, to (extracellular) recognition of conserved molecules of intruders, and the highly specific (intracellular) recognition of pathogen effectors. Effector triggered immunity (ETI) often relies on resistance (*R*) genes that are responsible for the identification of pathogens as well as the subsequent triggering of defense pathways (Chapter 1 and 2).

Most *R* genes encode for proteins containing a nucleotide binding (NB) and a leucine-rich repeat (LRR) domain, also known as NLR proteins. Plants generally contain a range of 600-1,000 resistance genes, most of which reside in gene clusters with homologous (pseudo) genes (Meyers et al., 2003; Seo et al., 2018; van Wersch and Li, 2019). The exact purpose of *R* gene clusters has not been fully elucidated. One hypothesis is that co-expression of certain *R* genes has provided benefits through for instance guard-decoy NLR pairs, coupling these genes, often in head-to-head orientation, during evolution, even when these co-expression benefits might not be present any longer (Adachi et al., 2019). While this explains the occurrence of paired genes, for instance for sensor and helper NLRs, this does not explain large gene clusters. Recently, it has been proven that ZAR1, a coiled coil (CC) NLR (CNL) of *Arabidopsis thaliana*, is able of homo-pentamerization and forms a resistosome that is similar in structure to the inflammasome in humans (Wang et al., 2019a). It is hypothesized that the close proximity of multiple *R* genes in gene clusters could have created more options for hetero-(penta)oligomerization of their gene products.

Clusters containing highly similar homologs have most likely arisen from gene duplication events and unequal recombinations during meiosis (van Wersch and Li, 2019). Pathogen stress can trigger quick expansion or reduction of *R* gene clusters, potentially resulting

in great differences between gene clusters in different ecotypes and cultivars. The result of such adaptations is for instance the divergent evolution of the potato *R* genes *Gpa2* and *Rx1* (Bakker et al., 2004), or the tomato NLR genes *Sw-5a* and *-b* (De Oliveira et al., 2016; Sharma et al., 2021b). The divergent evolution is further illustrated by the recently found *R* gene clusters in *Capsicum annuum* and *C. chinense,* respectively harboring the NLR gene *Pvr4* and *Tsw* (Kim et al., 2017a). The homologous genes *Prv4* and *Tsw* share a common ancestor and are located in the same locus on chromosome 10 of two different cultivars of pepper, but their gene products recognize structurally different viral effectors. Furthermore, activation of Pvr4 results in a durable extreme resistance (ER) against a range of potyviruses, while recognition by Tsw of the RNA silencing suppressor of tomato spotted wilt virus leads to a hypersensitive resistance (HR) and resistance is easily broken. The gene cluster of *Tsw* in *C. chinense* PI159236 contains 14 NLR genes, 10 of which, including *Tsw*, result from duplication events of a single ancestral gene, and are thus highly similar.

Here we report the successful isolation of the full *Tsw* gene sequence from a constructed bacterial artificial chromosome (BAC) library with 11x coverage of the genome of *Capsicum chinense* cultivar PI152225. Through genome walking from markers known to reside relatively close to the *Tsw* gene (Jahn et al., 2000), as well as primers based on the meanwhile published amino acid sequence of *Tsw* (Kim et al., 2016), a BAC clone has been obtained that contains the full-length gene of *Tsw*. In addition, BAC clones have been isolated that close the gap between the Q marker and *Tsw*.

Materials and Methods

Plant material

For the construction of the BAC library, *Capsicum chinense* cultivar PI152225 plants were grown for 6-8 weeks under greenhouse conditions (24 °C, 16 h light/8 h dark regime). Young shoots and leaves were snap-frozen in liquid nitrogen and stored at -80 °C until use.

BAC library construction

C. chinense nuclei were isolated and purified using a modified protocol of Liu et al, (1994). Purified nuclei were imbedded in 1.4% inCert Agarose (Biozym) to create agarose plugs. After lysis treatment, the plugs were partially digested with 10 U of *Hind*III (NEB) for 1 hour. The plugs were run on a Clamped Homogeneous Electric Field (CHEF) electrophoresis agarose gel at 4 °C for 18 hours, at 6 V/cm and a linear increasing pulse time of 60 seconds. The region in the gel containing fragments of 100-150 kbp was excised and used for a second size selection in a 1% SeaPlaque Agarose CHEF gel run at 4 °C for 10 hours, at 4 V/cm and a constant pulse time of 5 seconds. Fragments of 100 kbp were excised, dialyzed in TE buffer, and melted at 70 °C. Melted bands were treated with 1 U GELASE/100 mg agarose gel. The ~100 kbp fragments were cloned in pCC1BAC Copy Control Vector

and transformed into *Escherichia coli* DH10β via electroporation. Approximately 400,000 colonies were selected and divided over 800 pools of 500 clones stored in LB freezing buffer, constructing a BAC library of the *C. chinense* genome at 11x coverage. Each pool was grown overnight in 5 ml LB with appropriate antibiotics at 37 ˚C. To facilitate BAC library walking, miniprepped DNA (ThermoFisher) was pooled further and stored at -20 ˚C.

BAC library screening

Primers used for library screening were designed to generate products of 100-500 bp in a PCR with cycling conditions of 95 ˚C for 2 min, followed by 35 cycles of 95 ˚C for 15 s, 50-65 ˚C for 30 s, and 72 ˚C for 30 s, followed by 10 minutes of 72 ˚C. PCR products were subsequently run on a 1% agarose electrophoresis gel to determine presence or absence of fragments. Positive pools were grown in 3-5 ml cultures, diluted to contain 25-100 colonies/100 µl, plated and screened for single positive colonies.

Starting primers were designed for the Q marker and used to screen the BAC library pools. New primers were designed on BAC end sequences of positive identified clones to allow for BAC walking, as well as to determine the overlap and orientation of the new clones compared to known clones. Primers based on the published Tsw aa sequence (Kim et al., 2017a) were also created and used for BAC pool screening. All primers used and their respective annealing temperatures can be found in Table S3.1.

BAC clone sequencing and assembly

The selected BAC clones were grown overnight in 5 ml LB with appropriate antibiotics at 37 ˚C. The pellet was resuspended (Resuspension Buffer, ThermoScientific) and equal parts of 0.4 M NaOH and 2% SDS were added to the resuspended cell pellet. After addition of 1.5 parts 3 M potassium acetate, samples were frozen at -80 ˚C for 15 minutes. Samples were thawed and centrifuged at 12,000 x g for 15 minutes. Supernatant was transferred and centrifuged for 10 minutes. Ice cold isopropanol was gently mixed and frozen at -80 ˚C for 15 minutes. Samples were once again thawed at room temperature and centrifuged at 12,000 x g for 15 minutes. Pellet was washed with ice cold 70% ethanol, centrifuged, and allowed to air dry after removal of the supernatant. The pellet was resuspended in TE buffer and stored at -20 ˚C until sequencing with PacBio systems. Sequences were analyzed with Geneious Prime version 2017 and up. LRR motifs were predicted using LRRpredictor (Martin et al., 2020a).

Results

BAC library construction

To facilitate the search for the resistance gene *Tsw*, a bacterial artificial chromosome (BAC) library of the genomic DNA *of C. chinense* PI152225 was constructed in 2010. The BAC library, composed of approximately 400,000 clones with an average size of 100 kbp, was
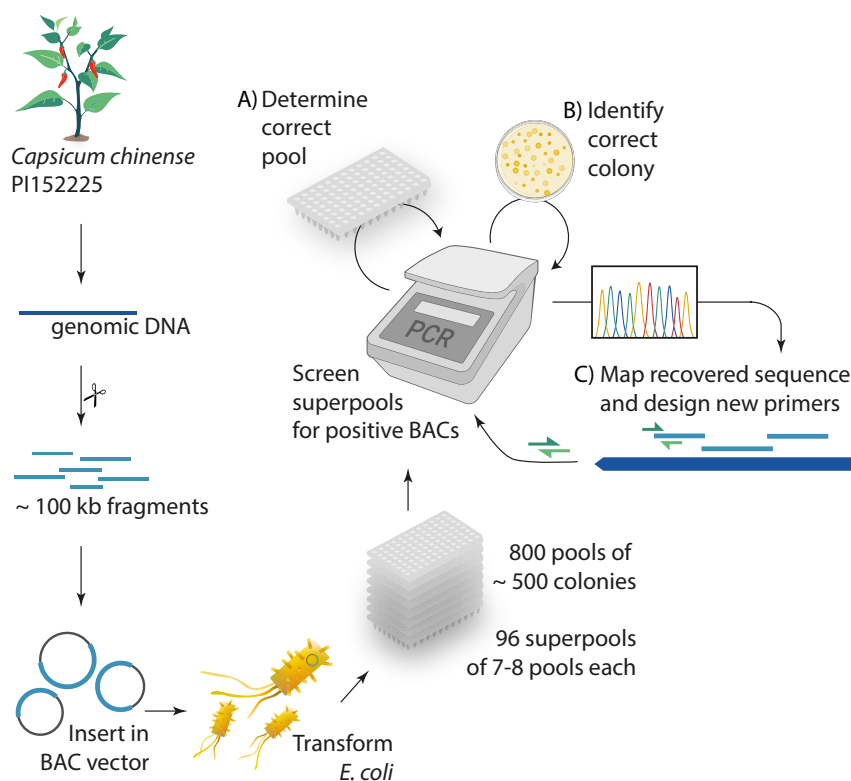
Figure 3.1 – Workflow of the creation and screening of a bacterial artificial chromosome (BAC) library containing the genomic sequence of *C. chinense* PI152225. Genomic DNA fragmented into 100 kbp pieces was inserted into BAC vector pCC1BAC Vector and transformed to *E. coli*. Colonies of *E. coli* were pooled per ~500, resulting in 800 pools. Pools were further pooled into 96 superpools to facilitate speedy screening of the BAC library. By screening superpools, pools, and colonies with PCR for the target sequence a single colony was finally found harboring the correct contig. The BAC ends of the contig were subsequently sequenced, enabling construction of a general map of the recovered contigs in the region, as well as the design of new primers to use for the further search of (i.e. walking towards) the resistance gene *Tsw*.

estimated to cover the *C. chinense* genome roughly 11 times. The 400,000 clones after transformation of *E. coli* were pooled per 500 colonies, resulting in 800 pools. To accelerate the speed of screening, samples of isolated DNA of all pools were further pooled into so-called superpools on a 96-wells plate (Figure 3.1).

BAC library walking and sequencing

Primers made to anneal to the Q and L marker, both indicated to be in relative close proximity to *Tsw* (Jahn et al., 2000), were used to screen the BAC library superpools. Only the 7/8 pools used compose the superpools that were indicated by PCR to contain a sequence of interest were subsequently screened. Colonies of positive pools were screened to uncover the single positive BAC clone. End sequences of positively identified BACs were sequenced and used to create new primer pairs to screen the BAC library (Figure 3.1). Primer pairs were designed to amplify fragments of 200-400 bp.
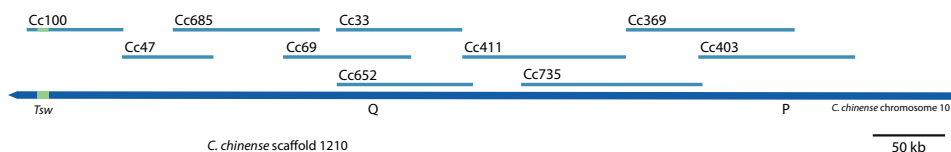
Figure 3.2 – A physical map of contigs recovered from the BAC library of *C. chinense*. Recovered contigs (light blue) spanning chromosome 10 of *C. chinense* PI152225 (dark blue) between the P and Q marker, and the resistance gene *Tsw* (green). Scaffold1210 (grey) of the published contig sequences of *C. chinense* v1.2 partly matches.

A general map was created to visualize the respective orientations and overlaps of the recovered BAC contigs. Upon the publication of the Tsw amino acid sequence by Kim *et al.*, additional *Tsw* specific primers were created to speed up uncovering the BAC contig containing the *Tsw* sequence. Contigs presumed to link the Q marker with *Tsw*, namely Cc100, Cc47, Cc685, Cc69, Cc652, and Cc411, were sequenced using PacBio sequencing. Sequenced contigs were *de novo* assembled and the location and orientation of these and previously recovered BAC contigs was confirmed (Figure 3.2). As expected, Cc100 contained the full sequence of *Tsw*. Assembly showed that the overlap between Cc100 and Cc47 is 277 bp. When delving further into the fully sequenced region, the 550 kbp was shown to contain 870 regions of at least 50 bp that were repeated at least once, of which 214 were longer than 100 bp.

Comparison to the published Tsw coding sequence from Kim et al. (2016)

The complete sequence of Tsw and its flanking untranslated regions (UTRs) were found on Cc100 (Figure 3.3). The uncovering of the full sequence, rather than just the coding sequence, provides opportunity to use the non-coding sequences for transcription regulation analyses (see Chapter 5). The full gene sequence, spanning 18,651 bp from start to stop codon, codes for an NLR gene. The coding sequence of 6.351 bp assembled from nine exons code for the coiled coil (CC) domain, nucleotide binding **A**paf-1, **R**-protein, and **C**ED-4 domain (NB-ARC) domain, and the leucine-rich repeat domain (LRR). The NB-ARC domain contains all motifs typically found in these domains (P-loop/Kinase-1, RNBS-A, Kinase-2, RNBS-B/Kinase-3, GLPLA and RNBS-D (Meyers et al., 1999)). At amino acid (aa) level, the conserved CC and NB-ARC domains of the Tsw protein are very alike that of other NLR genes. Using the LRRpredictor (Martin et al., 2020a), it could be determined that the LRR domain, starting from aa 511, contains 57 LRR motifs. These motifs make up one core LRR segment, followed by seven almost identical blocks of LRR repeats. In total, the LRR domain is 1,632 aa long, longer than the LRR of most known NLRs (see Chapter 4).

The published coding sequence (CDS) of *Tsw* and the CDS of the *Tsw* gene found on BAC contig Cc100 only differed by two non-synonymous point mutations in exon 5 and 6 (Figure S3.1). The intron sequences of the previously published *Tsw* variant (Kim et al.,

2017a) were recovered by alignment of this sequence to the *C. chinense* genome v1.2 scaffold sequences, as well as the later published *C. chinense* genome (MCIT00000000, (Kim et al., 2017b)). The comparison with the full-length sequences obtained in the current study revealed a few 1-6 nt insertions or deletions in intron regions, in addition to a few single nucleotide mutations. A manual search for potential splice sites in the *Tsw* sequence recovered via our BAC library also did not indicate a different intron/exon structure.

Following the recovery of *Tsw* from the BAC library, the assembled sequences of the BAC contigs Cc100, Cc47, CC685, Cc69, and Cc652 were mapped against the published Scaffold1210 of *C. chinense* PI152225. The sequences corresponded very well, despite multiple 1 bp differences. Differences caused by large gaps were mainly the effect of gaps in the published sequence, the only exception being a large insertion in Scaffold1210 at the left end of Cc652, disrupting the rest of the sequence.
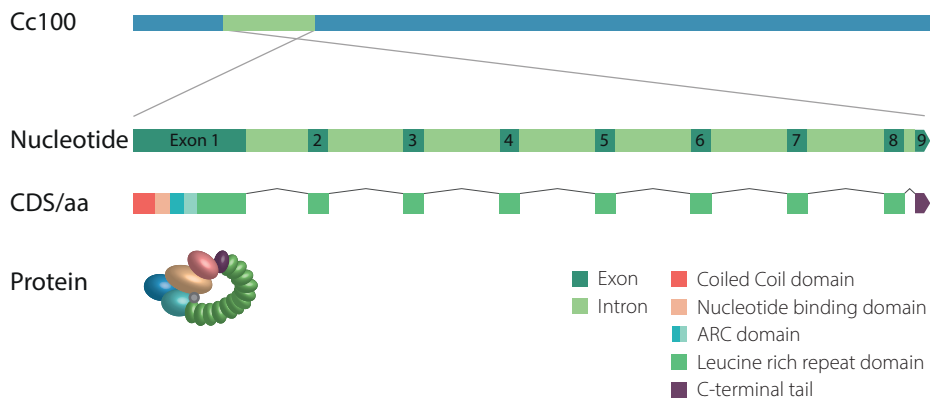
Figure 3.3 – An overview of the *Tsw* gene and the resulting protein. A physical map of the Tsw gene on the *C. chinense* BAC library contig Cc100 (blue), the gene organization of *Tsw* comprised of 9 exons (dark green) and 8 introns (light green), the domains and their location on the coding sequence (CDS) of *Tsw*, and *Tsw* stylized as a 3D protein.

Discussion

BAC libraries have been constructed for genomic studies of countless different organisms. While this technique is relatively old, it is still a relevant tool in the sequencing of genomes, investigation of gene functions, and positional cloning. We have constructed a BAC library of the genome of *C. chinense* and successfully screened it for the complete sequence of the resistance gene *Tsw*. With a genome size of approximately 3.2 Gb (Kim et al., 2017b), and 400,000 contigs of ~ 100 kbp, this BAC library was estimated to have a 11x genome coverage. When screening all 96 superpools for a unique sequence, around 10-11 positive superpools were expected to be found. However, quite often the number of positive superpools returned was higher, which possibly resulted from repeat sequences associated with *R* gene clusters, as well as the LRR repeats found in *R* genes. Sequencing of the contigs spanning the region between the P and Q marker,

and the *Tsw* gene verified the high similarity and considerable number of large repeats found in the region. The similarity of sequences in this region has likely been the cause of a longtime struggle to move away from the previously uncovered contigs, resulting in many overlapping BAC contigs of the same region.

Chances of finding the coveted resistance gene were increased by also searching for *Tsw* from the other side of the genome using primers based on the L marker. However, genome walking from this location did not result in the discovery of *Tsw* or other genes residing in the *R* gene cluster. Mapping of BAC end sequences on the published genome of *C. chinense* revealed that these did not map to either scaffold on which the *Tsw R* gene cluster was indicated to preside, but did align with sequences of Scaffold71. However, due to the incomplete assembly of the publicly available *C. chinense* genome v1.2, it is unknown how far removed this scaffold is from the respective *R* gene cluster.

Upon publication of the coding sequence of *Tsw* by Kim *et al.*, the discovery of a contig harboring *Tsw* was sped up by using gene-specific primers, resulting in the finding of Cc100. The subsequent sequencing of uncovered BACs showed that the overlap between the contig and the published CDS of *Tsw*, Cc100, and its neighbor Cc47, was a mere 276 bp. Primer pairs used for screening were designed to result in 200-400 bp fragments. For all primer pairs designed based on the BAC end of Cc47, one of the primers from a pair targeted the sequence outside of contig Cc100 and therefore would never amplify any fragment to indicate its vicinity to Cc47.

The exact activation mechanism of *Tsw* and subsequent triggering of downstream immune pathways remain to be determined. It is known though, that *R* genes, *Tsw* included, are often less effective at higher temperatures. Increased temperatures as a result of climate change, as well as increasingly prevalent *Tsw*-resistance breaking strains of TSWV emphasize the importance of unravelling the details of *Tsw*-mediated resistance. Furthermore, a search and functionality assessment of allelic variants of *Tsw* could possibly result in the finding of an added measure against TSWV infection, for instance in the discovery of variants whose resistance is less easily broken, or by finding crucial domains or host factors that could be employed in creating more resistant, or less susceptible pepper plants. Sequence comparison of the published *Tsw* and our version revealed their almost identical similarity. Comparison of our BAC contigs with the published Scaffold1210 of *C. chinense* also revealed highly identical sequences, despite large gaps found in Scaffold1210.

The complete sequence of *Tsw* including non-coding sequences makes it possible to search for regulatory elements affecting its transcription to increase insight into its regulation (Chapter 5) and to facilitate cloning of the *Tsw* gene into plant expression vectors. In light of its unusually large LRR domain, analysis of 3D folding predictions of Tsw will support the construction Tsw variants for functional gene studies (Chapter 4).

Supplemental data

Figure S3.1 – Comparison of the two nucleotide sequences of *Tsw*. Assembly of the full sequence of Tsw with 2000 nt upstream of the start site from the public *C. chinense* genome database v1.2 mapped against the recovered full-length sequence of *Tsw* from the *C. chinense* BAC library contig Cc100 described in this chapter. " . . . " represents a match in the sequence. Green sequences represent exons, black represents non-coding sequences.

```
Tsw Cc100   GTAAATCTAATCGAAAACATCAAAAAAGAAAAAAGAAACACTGAAAACCAGAACGTAAGAACCCATAGGAAAGAGGTTGATCTCCAATGATACAACCACAAGTAAGAATTACT  -1921
Tsw Kim '17 ...................................................................................................................  -1920

Tsw Cc100   TCATGAATTGTGCAGATGAATAAACATGCATAACGGTAGAAGAAGGGTATGCTACACTAAGTAAATTAAATTGTATAATAATCACATCAAAGTTTCAGAAAGCATGTGACTC  -1801
Tsw Kim '17 ...................................................................................................................  -1800

Tsw Cc100   CGCATCTAAGCTGAGAAGGTGTAGTCTTAAACTAATTGTTGTTTTCAATTGATTGTTTCATAATTTCACGTTGAATTTTGTTCTTTATGTATAAAAGAATCCTAAACCTAAA  -1681
Tsw Kim '17 ...................................................................................................................  -1680

Tsw Cc100   GAAACCTAATTCTTAACCTAAACAAAGAAAAGATATATATCTATAATATATTAAAAGTGTGTAGGATATTAGAAATGTTGTTTGAACTTTTCGAAAAGACATATATCTATAA  -1561
Tsw Kim '17 ...................................................................................................................  -1560

Tsw Cc100   AAGTGTGAAGGATGTTAGAAATGTTATTTGAACTTTTTGCCTTTCATTAAAAGTGTTTGTTTAGATAAAATCGGCTTTTCACTATTTTTTTTCTAATATTTAAGAGCTAAAA  -1441
Tsw Kim '17 ...................................................................................................................  -1440

Tsw Cc100   AATATATTTATGGTAAATCTTTTCTTTATTAGAAGTTATCAAAATTAAAAACGGGGTGGTTCAACACGTCACCCCTTAAGCAAGAGGTCGGGGGTTCGATCCCCACCTCTGG  -1321
Tsw Kim '17 ...................................................................................................................  -1320

Tsw Cc100   AAAAACTCTGTGGTCAATTCCCTACCGCTTAATGCGCTGACCGTAGGAATGAAGGATTAGTCTCACGACTTCCGGATGTGGGGATACCTTGAAAAACAAAAATAAATAAATA  -1201
Tsw Kim '17 ...................................................................................................................  -1200

Tsw Cc100   AATAAATTTAAGAGGTGTGTCAAGTCAATAATAATATAGTAGTAGTACCAATTTAATTGGTGACAGTTCATGTTGTGGGGGTAGATGAGAGTAATAATAGAGTAGTAACTCA  -1081
Tsw Kim '17 ...................................................................................................................  -1080

Tsw Cc100   TAACGCTGAAAATGGTAGGTGGAAGATGAAATTGTGGTCGAGACCAACATCTCAATTGAAAAAAAATTGAAGGACAATTGTTTTTGAGATATTACAATTATATAATTATCAA  -961
Tsw Kim '17 ...................................................................................................................  -960

Tsw Cc100   CAAATGTTACAAATAATTCAAAATGCAAGGAAGTTAAGCTTAATATTTATGAGAGACCAGTAATAAGATCAAAACTGGTATTTGTACATGTATGAGAGTCATTTTCAATTTA  -841
Tsw Kim '17 ...................................................................................................................  -840

Tsw Cc100   ATTTTGATTTGATAAGTTTTAAATTAAAGATATGTATAATGTATCATAGTATCTTTTAAATTTGTGGTCCTTAACATGTTCTGCGAAAAGTTAGAATAATGAATTGCCAGAA  -721
Tsw Kim '17 ...................................................................................................................  -720

Tsw Cc100   GAAATATTCTTTTCTAAACGGACTAAAAAAATAGGATAAATAAAAATTGAAACGAATGGAGTTTACCAACTGCAAAGTGAGAGGAAGGAAAGCAAAGAAAAGTCATTACTTC  -601
Tsw Kim '17 ...................................................................................................................  -600

Tsw Cc100   CATTTTATGTGTGACATATTTTGATCGGATACGGAGTTTAAGAAATAAGAAAAAACTTGATGATATTTACTAAATTATCTTTTATTAAAAAAATGTAATACATTAACTTTAA  -481
Tsw Kim '17 ...................................................................................................................  -480

Tsw Cc100   GACCTAAAATTGAAATTGTGTCTTTAAATAATTACCTGAAATTGTGTCTTTAAATAGTTACTAAATAGGAAAAAGGTTACATTCTTTTTAGGACAAATCAAAAAGAAAATAG  -361
Tsw Kim '17 ...................................................................................................................  -360

Tsw Cc100   TTATATTTTTTAATAAAGGATAATTTAATAAAACGCCAACAAATCAATAATTTTTTTTTTTTTTTTAAATTTCGTGCTCAATCAAAATATAGCAAATAAAATTGAACGAGGGAT  -241
Tsw Kim '17 ..........-........................................................................................................  -241

Tsw Cc100   ATGGCATTATAATACTATTTCCCTCATAGCCCTTAAAGTCCAAGTGCGAGTTTTCAATCACTCTCAAGCCGTTTTCATCTAATTCCTACTCCAAAAAAAAATTTGATCTCCT  -121
Tsw Kim '17 ...................................................................................................................  -121

Tsw Cc100   GTAAATCTAATCGAAAACATCAAAAAAGAAAAAAGAAACAACTTTGGCAGTTCAGACTTCAGAGACTGTTTTTGTTTATTGGACTAGCAAGTGAAGAATTAATTAAGGAGAT  -1
Tsw Kim '17 ...................................................................................................................  -1

Tsw Cc100   AAATGGTAACTAATATTTTGAGCCCAGTCGCAGAACACTTGATCATACCGCCAGTTGTGCGACAGATTGGGTATTTGTTTTACTATAGGCGCAACATCAGGTCTCTGGATGA  120
Tsw Kim '17 ...................................................................................................................  120

Tsw Cc100   ATAAGCTCGAGAATATCAGAAGTGGGGTGAATCAAAGAGTGGAGGCTGCCCGGAGAAACTTACAAGTCATTTCACCCAATGTTGAGGCTTGGTTAACTAGTGTTGATACTAC  240
Tsw Kim '17 ...................................................................................................................  240

Tsw Cc100   ATGTGGCAGTTGTAATGGGAGGTCGAACTGAGGTTGAAAGGGGTTGCATCTACGGTTGGTGTCCAACATTGAAGTCACGTTACTTGCTGAGCAGGAGAGCCAAGAAAATTAC  360
Tsw Kim '17 ...................................................................................................................  360

Tsw Cc100   TTACTAAACTTCAAAAGGAAGGTAACGAGTATGCTGTTTTCTGCTATCCGGTTCCAGCTGATGAAATTGAAGCTATACCTAGTAACACTAATGAGGAGTTTGTCTCTAGAAA  480
Tsw Kim '17 ...................................................................................................................  480

Tsw Cc100   AGGAAGAGGTCATGACAGCTTTGAAAGATGAGGAGATCACTATTGTTGGGATATGTGGTATGGGTGGTGTTGGTAAAACATCACTGGCTGAGAAAGTAAGAGCAAGGGCAAA  600
Tsw Kim '17 ...................................................................................................................  600

Tsw Cc100   GGTTATTTAATGATGTTGTCATGGTAACTGTCAGTCAACAACAAGACCTGAAAAGAATTCAGGGCGAGATCGCCGGAGAGGTCGGGCTGACGTCATTGCAAGGAGACAATTT  720
Tsw Kim '17 ...................................................................................................................  720

Tsw Cc100   GTGGAGATCAGCTGCGTGCCAAGGTTAATGCAAGAAGGAAGTCGCGCTCCTTGTAATTTTGGATGATGTTTGGGAGGCTCTTGATGATCTAGAGAAACTTGGAATTCCCACTGG  840
Tsw Kim '17 ...................................................................................................................  840

Tsw Cc100   ACAACTATCAGTGCAAAGTGGCATTGACGACGCGCCGCCAAGATGTTTGTGACGCTATGAAGGCTCAAAAGACTGTAGATGTTGTAATCTTATCTGATGAAGAAGCATGGGT  960
Tsw Kim '17 ...................................................................................................................  960

Tsw Cc100   GGCAGAAAGCCGGTTATTCAGCTGATGATCCCTCTCTTTCTGAAGTAGCAAAAGAAGTTGCCAAAGAATGCAAGGGGCTGCCACTTGCAATTGTTATTGTTGCAGGAGCACT  1080
Tsw Kim '17 ...................................................................................................................  1080

Tsw Cc100   AAACCAAGCCTTCATGGGATGATGCCCTTGTAGAATTAAGAAAAGCATCACCAAAAAATATCCCAAGAGTGCTTACAAAGGTGTATCAACCTCTGAAGACAAGCTATGATCA  1200
Tsw Kim '17 ...................................................................................................................  1200

Tsw Cc100   GTGATGAAGCCAGGTATGTCTTTTTGCTTTGTTCCTTGTTTGAGGAAGATAGTGATATTTGGACCGAAGAATTACTTAAATATGGAGTGGGACTTGGCATCTTTTCGGAACT  1320
Tsw Kim '17 ...................................................................................................................  1320

Tsw Cc100   TAGAATGTGCAAGAAATAGGGTGTGTAATCTGTTAGAAACATTGAAAAATTGTTTCTTGTTATCCCAAGGTAAAGACAAAAATTATGTCAAAATGCATGATGTGGTCCGTGA  1440
Tsw Kim '17 ...................................................................................................................  1440

Tsw Cc100   TATATATTGCGTCTGAGGGTGAGCATAAATTTCTGGTAAATCACAATGTGAACTCAAACGTGTTCTTGAGAAAAGTCTCTTATGAGCAATACAGTCACATGTCAATTGTTGC  1560
Tsw Kim '17 ...................................................................................................................  1560

Tsw Cc100   TTGATGAGCGTCCCACACCAATATTTTGCCCACGATTGAAGCTTCTGATGTTAAAACTCCGTTTCGAAGAGGGTTTCAAATTACAGGATGATTTTTTTGATGGAATGAGTGA  1680
Tsw Kim '17 ...................................................................................................................  1680

Tsw Cc100   TCATAAAACTGAGTGGATATGATCGAAACTCCATTCTGCCCTTTCCATCATCCATTCAGAGGTTGTCAAATCTGAGCACGCTTTGGCTGAGTAATCTAAGGTTGGATGACGT  1800
Tsw Kim '17 ...................................................................................................................  1800
```

```
Tsw Cc100   TTGGGAAACTTGTCACTTTAGAAATTCTCAGCATCAGAGGTTCTGACTTACAGGAGCTTCCAGTGGAGATCGGAAACTTGGCCAATCTAACTATGTTAGAGTATTGGAATAC   1920
Tsw Kim '17 ............................................................................................................   1920

Tsw Cc100   GAAAACGTATGAGGATTTCACCAGGGGTCTTATCAAGACTAGTTCGATTGGAGGAACTACATATGGTGGGAGTAGAAGATTGTAGTTACTCCACCTTGAGGGAGCTGGAATC   2040
Tsw Kim '17 ............................................................................................................   2040

Tsw Cc100   GATTGACTGCACTGGCATTTGATGAATGTTCCGTAGATGTGATTTACAGTAACTTGGGCCTTTCCTCCAAGTTGACACGGTACGCTCTTAAAATGGGTAGACACTACACGTT   2160
Tsw Kim '17 ............................................................................................................   2160

Tsw Cc100   TCATGGAAACTTACAACAAGGCTATCGATCTAGACGTCACCAAGGGCACCCCATTGGGTGATTGGATCTGCCTCCTGTTGAGGAATAGCGAAGTTGTACATTCAAGAGGAAA   2280
Tsw Kim '17 ............................................................................................................   2280

Tsw Cc100   AGAATGTAATGGTCGAGTTGCAGAATGTGAAAGATCTCATGCTGTCTGATTGTGATTCATTGAATATCCACTATCAGAATAATATTTCATTCCCCGAACTCGAAAGGCTGGA   2400
Tsw Kim '17 ............................................................................................................   2400

Tsw Cc100   ATTGTGATTATCTACGGCATCTTTTTTGTGTGTCCTTGGCTTGCCCTGATGAGGGGACCTCTCGAAGGACACACATCAGACCTGACGTAATCAAGTTCCCCAATTTACATAG   2520
Tsw Kim '17 ............................................................................................................   2520

Tsw Cc100   TTCGTAATCTGGAATTCTTCACACACTTTTACAGTGACACTGTTGAGGGCATTGAGTTCCCTCTGTTACGGGTAATATATCTCAGGGGATTACCAGAGTTCCAAAATTTCTG   2640
Tsw Kim '17 ............................................................................................................   2640

Tsw Cc100   CCAACAATGCTATCACCGACTCAAATCCTCTTTTTAATGAAAAGGTTTGCTTCTACACGAATAAGTTTTTTTTTTGAAGGGGTTTTATGAATTAATATTTTAGCTTTAACTT   2760
Tsw Kim '17 ............................................................................................................   2760

Tsw Cc100   ATATTTATAAGATGTTACTCTCTTTGTTCCAACTTTTATAATATTGATAATTAATTGGGGGTGACATGATAAAATTACGATTGAAGCAGCAAATGATACATGTCTTAAATTA   2880
Tsw Kim '17 ............................................................................................................   2880

Tsw Cc100   AACTAATTAGAAGTGATGTAGCAAAATTACTTTAAAAAATACTTGTGGAAAAATAATTAAGTGGGCGCCATATAAGACAGGTTTGACATTAAATTTTATTAATTTTTTAATA   3000
Tsw Kim '17 ............................................................................................................   3000

Tsw Cc100   ACTTTATAATAAGTTTTTAATATTAAATTTTTTAATACTTAAATGAATTGTAATAATTAATTAGGGGTGATAGTAAAATTATGGTTGAAGCAGATATGTCAAAAATGTTTAT   3120
Tsw Kim '17 ............................................................................................................   3120

Tsw Cc100   TTCAATTAAATATTAATAATTTTCTAATTCGTTAATGCTATGTAATAATTAATTAGGATTGATATAGAAAAATCATGGAGCAAGCAATTGAATGGTCGACAAGCAGAACGAC   3240
Tsw Kim '17 ............................................................................................................   3240

Tsw Cc100   GTTCACTCCCTTTTATAATATAAGAGAATGCCTATTTTTTAAATTATATACAACAACATACCTAGTTCTTTATCGTATTTTTTTTAATTGTAATTAATAAAGCATTTATCAA   3360
Tsw Kim '17 ............................................................................................................   3360

Tsw Cc100   TTGAAATTCTATTGAAGTACAATAGAGTAGCCGTTGTCAATATTTTTAATGCGTGATTAGAGAGACTGTGTGTACATGGGACTTTTTTGATCTTATGTCACATGAACATTAT   3480
Tsw Kim '17 ............................................................................................................   3480

Tsw Cc100   TAATTTTATGAAGTTACAAATAAAGTTTTAATAAATCTATAAAAGTTTAGTGGTAACTGGCAAATTGACAAATACATATATACTGTTAAACACATTATAGTTATTATTTTAA   3600
Tsw Kim '17 ............................................................................................................   3600

Tsw Cc100   ATAACTTTGCAATGATGTACGTATTCAAACTTGACCAAATTTTAAAATAATATATAAAAGTCATTAAAAACCCAAAAACTAATAAACTTTGTTCAATTTTAAAAATTACCGA   3720
Tsw Kim '17 ............................................................................................................   3720

Tsw Cc100   CATCTCAAATTCGAAGAAAATATCTTGAAATATGAATATGAGTTTGAAAGAACAAATGAATTGACACTTTTCATCCTCCTTTATGATGAAAAGATGTATTTCTACTGTTCTAA   3840
Tsw Kim '17 ............................................................................................................   3840

Tsw Cc100   TAAGAAAAAATAATTTCTTCGAATGTTTCTCAAGTAGAATATTTATTAACTTCTACCATTTTATTCGAAATAGAGCGGTCAAATTCATAATTAAATGGATAGTACAATTAATT   3960
Tsw Kim '17 ............................................................................................................   3960

Tsw Cc100   GCATATTCATATATTTACATTTTTATTGATATCTAATCGTTTAATCTAAAAAAAGATTTAACTTACTGTAATAAAATAACATTAAAAGACACACTAGAAATAAGAATTTTACTA   4080
Tsw Kim '17 ............................................................................................................   4080

Tsw Cc100   TTGCAAGACATGAGAAATGGAGTTGATACCTTGTCAATGTTTTGCAGGTTTCTTGTCCCAACCTGAAAGTGCTAAAACTCCATGAGGCTAACAGCATAACTGCTCTGTGCTC   4200
Tsw Kim '17 ............................................................................................................   4200

Tsw Cc100   TTCCAACTACCTACTTCAGCAAACTTGAGACATTGAAAGTAGAAAATTGTGGAAAATTGAGACACTTGATGTCTCCATCAGTGGCCAGAGGTCTTCTGAATCTCCGAATATT   4320
Tsw Kim '17 ............................................................................................................   4320

Tsw Cc100   AATACTGCGAATCAATGGAAGAAGTGATCATAGAAGAAGAACAAGAAGGAGATGAAATCATGTGTAACGAACCCTTATTTCCCCAATTGGAAGAGCTGGAACTTCAAAATCT   4440
Tsw Kim '17 ............................................................................................................   4440

Tsw Cc100   TGAGGCATTTCATTTTGACGAAGCAAGCTCTTGAATTTCCATTCCTCATAGAAGTGAAGATTCGTAACTGCCCTGAAATGAAGATGCTTGTCCAACGGAGATTTGTCAGTAC   4560
Tsw Kim '17 ............................................................................................................   4560

Tsw Cc100   TCAAAAGTGTGAACAATGATGATGAGTTGAAAGTTGTTGATCTCAACAAAGCGATGTTTAATTCTAAGGTTTGTCTTGTGCCGCTTTATAACTAAATGATATACATATTTGA   4680
Tsw Kim '17 ............................................................................................................   4680

Tsw Cc100   CTAACAGATCCCTGGTCTATTTCTCTCTTTCTGGTTTAAGAAATGGTGCTTATGAAAAATGCCACATAAGCAGACCATCTTGTTTGATTGCTTGTTCAATTGTCTTGATGAC   4800
Tsw Kim '17 ............................................................................................................   4800

Tsw Cc100   GAGAAAACGCTGTTGACTGGAACTGAAAACCCAATGAATTTCCTAACCCAAATTCAAGAATTAAACACACCCACACTAAATAAACCACAGACCAATACGAAGCAAGCCAAAG   4920
Tsw Kim '17 ............................................................................................................   4920

Tsw Cc100   ACTGAAACCATGCTATTCAGCCATTTTTATGTTCAATTAGTCAAACATAGGAATGCTTTAAGTATTAGACATAGATCGGGGATGAAGGTAATATTACGCCCCAAGTTTTGGA   5040
Tsw Kim '17 ............................................................................................................   5040

Tsw Cc100   ACCAGTTCAATCTAATTGTAATAGTAATCTTCATGTAGTTATTTTTTGTTTGTCACCTGATACCTATGTTGATGGAGGTAACAAACATTGGTGAATCTACAATATTATGATC   5160
Tsw Kim '17 ............................................................................................................   5160

Tsw Cc100   CATGTGAACCCAATAACTTTTTTCATAGACTTTTTATCTGTATTAATAATATCTAGAGATTTCTTGAGGTCATTGAAGATTTCGATAAATTTCAGATTTTAAATCTACCTCT   5280
Tsw Kim '17 ............................................................................................................   5280

Tsw Cc100   ACTCTTCTTTCATTATATGTAACAACTAATATATGGATTTGTTTTTCTCATTTCATCTTGCATTCTTTGTCATTTTAAAAGATCTTAAAGAATGCTATTTTAAAGATGCTTT   5400
Tsw Kim '17 ............................................................................................................   5399

Tsw Cc100   CTTCTTCTATTTGAGTAAACACCTCAACTAAAATAAGCACTTTTGGCTTCTCAACTTATCACTTAGGGCGTGTTTGGTATGAAGGAAAATGTTTTCATGGAAAGCAAGTTGAT   5520
Tsw Kim '17 ............................................................................................................   5519

Tsw Cc100   ATTTTCTTATGTTTGGTTATTGAGTGGAAAATATTTTTCGAAAAATATTTTCTAGTGTTTAGTTAGTGAATGAAAATATATTTTTCAGAAAATACATTTTAGTATTTAGCTA   5640
Tsw Kim '17 ............................................................................................................   5639

Tsw Cc100   ACTTTTGTAGAAAATAAGTTTGACGCTGAAAAATATTTTTTAGGTTGTGTTAGGTGGAAAACATTTTTTTTTATAAAAATAAGTTGTTTCTTGTTTATATTCTTGTGTTCGTT   5760
Tsw Kim '17 ............................................................................................................   5759

Tsw Cc100   ATTAAAATATTTTTTTAAAGTATTTGTATATGTTTAACATAAAACAATGGGAATGAATAGGGTAAGAAAAAAGTTTAAAATATTTATTAAATATTTTTTTTGAGGAGAGGGT   5880
Tsw Kim '17 ............................................................................................................   5879

Tsw Cc100   GGAGGGATGTATCAGAATATGTGGTAGGGTTAGGTGTGGGGTAAGAAAAAAAGTTTTGAATTTTTTTAAAGAAATTATTCTGTTTTTTTTTGGGGTGGGGGGGGGGGGGGTCA   6000
Tsw Kim '17 .............................................................................-..............................   5998

Tsw Cc100   TGGGGCGTAAAAAATTGTTTTTTTGGTTAACAAATCGTGTGTGTTTTATTGTAATTAATAAAATATTATCTAACTTGTAATTTGAAATTCTTTTGAAGTATAACAGAGCAGC   6120
Tsw Kim '17 ............................................................................................................   6118
```

```
Tsw Cc100   ATTTCTTACAAAATATGCCTCTATTTTAAATGCACAGTTAGAGAGACTGTGGACATAAGAAGTTTTGGTTCACTTTTCATCAAAATCAAACCAAATTAACTATATCGGTTTT   6240
Tsw Kim '17 ..........................................................................................................   6238

Tsw Cc100   ATTTATATATTATAAATTCATATATTTACATTTTATTGATATCTAATGATTTAACTTACTGTGATAAAATAACATTAAAAGACACACCTGGAATAAGAATTTTATTAGTAAT   6360
Tsw Kim '17 ..........................................................................................................   6358

Tsw Cc100   AGACATGAGAAATGGAGTTGATATTTTGTCAATGTTTTGCAGGTTTCTTGTCCCAACCTGAAAGTGCTAAAACTCCATGAGGCTAACAACATAACTGCTCTGTGCTCTCACC   6480
Tsw Kim '17 ..........................................................................................................   6478

Tsw Cc100   ACTACCTACTTCAGCAAACTTGAGACATTGGAAGTAGAAAATTGTGGAAAATTGAGACACTTGATGTCTCCATCAGTGGCCAGAGGTCTTCTGAATCTCCGAATATTACTAT   6600
Tsw Kim '17 ..........................................................................................................   6598

Tsw Cc100   TGTGAATCAATGGAAGAAGTGATCATAGAAGAAGAACAAGAAGGAGAAGAAAACATGACTAATGAGCCCTTATTTCCCCTGTTGGAAGAGCTGATACTCGATAAGCTACCAA   6720
Tsw Kim '17 ..........................................................................................................   6718

Tsw Cc100   CATTTCTTTCTGACGAAGCGTGCTCTTGAATTTCCATTCCTCAGAGAAGTGAAGATTCGTAAATGCCCTGAAATGAAGATGCTTGTCCAACAGAGATCTGTCAGTACATCAA   6840
Tsw Kim '17 ..........................................................................................................   6838

Tsw Cc100   AGTGTGAACAATGATGATGAGTTGAAAGTTGTTGATCTCAACAAAGCGATGTTTAATTCTAAGGTTTGTCTTGTGCCGCTTTATAACTAAATGATATACATATTTGAGTAAT   6960
Tsw Kim '17 ..........................................................................................................   6958

Tsw Cc100   AGATCCCTGGTCTATTTCTCTCTTTCTGGTTTAAGAAATGGTGCTTATGAAAAATGCCACATAAGCAGACCATCTTGTTTGATTGCTTGTTCAATTGTCTTGATGACAAAAT   7080
Tsw Kim '17 ..........................................................................................................   7078

Tsw Cc100   AACGCTGTTGACTGGAACTGAAAACCCAATGAATTTCCTAACCCAAATTCAAGAATTAAACACACCCACACTAAATAAACCACAGACCAATACGAAGCAAGCCAAAGATTTC   7200
Tsw Kim '17 ..........................................................................................................   7198

Tsw Cc100   AACCATGCTATTCAGCCATTTTTATGTTCAATTAGTCAAACATAGGAATGCTTTAAGTATTAGACATAGATCGGGGATGAAGGTAATATTACGCCCCAAGTTTTGGATTGTT   7320
Tsw Kim '17 ..........................................................................................................   7318

Tsw Cc100   TTCAATCTAATTGTAATAGTAATCTTCATGTAGTTATTTTTTGTTTGTCACCTGATACCTATGTTGATGGAGGTAACAAACATTGGTGAATCTACAATATTATGATCCAAGG   7440
Tsw Kim '17 ..........................................................................................................   7438

Tsw Cc100   GAACCCAATAACTTTTTTCATAGACTTTTTATCTGTATTAATAATATCTAGAGATTTCTTGAGGTCATTGAAGATTTCGATAAATTTCAGATTTTAAATCTACCTCTCGTTC   7560
Tsw Kim '17 ..........................................................................................................   7558

Tsw Cc100   TCTTTCATTATATGTAACAACTAATATATGGATTTGTTTTTCTCATTTCATCTTGCATTCTTTGTCATTTTAAAAGATCTTAAAGAATGCTATTTTAAAGATGCTTTTTAAA   7680
Tsw Kim '17 ..........................................................................................................   7678

Tsw Cc100   TCTATTTGAGTAAACACCTCAACTAAATAAGCACTTTTGGCTTCTCAACTTATCACTTAGGGCGTGTTTGGTATGAAGGAAAATGTTTTCATGGAAAGCAAGTTGATTTTTT   7800
Tsw Kim '17 ..........................................................................................................   7798

Tsw Cc100   CTTATGTTTGGGTTATTGAGTGGAAAATATTTTTCGAAAAATATTTTCTAGTGTTTAGTTAGTGAATGAAAATATATTTTTCAGAAAATACATTTTAGTATTTAGCTAAAGAA   7920
Tsw Kim '17 ..........................................................................................................   7918

Tsw Cc100   TGTAGAAAATAAGTTTGACGCTGAAAAATATTTTTTAGGTTGTGTTAGGTGGAAAACATTTTTTTTATAAAAATAAGTTGTTTCTTGTTTATATTCTTGTGTTCGGTTATGCA   8040
Tsw Kim '17 ..........................................................................................................   8038

Tsw Cc100   AATATTTTTTTAAAGTATTTGTATATGTTTAACATAAAACAATGGGAATGAATAGGGTAAGAAAAAAGTTTAAAATATTTATTAAATATTTTTTTGAGGAGAGGGTTGGTAG   8160
Tsw Kim '17 ..........................................................................................................   8158

Tsw Cc100   ATGTATCAGAATATGTGGTAGGGTTAGGTGTGGGGTAAGAAAAAAGTTTGAATTTTTTTAAAGAAATTATTCTGTTTTTTTTGGGGTGGGGGGGGGGGGGGGGGTCAGTAC   8280
Tsw Kim '17 ..........................................................................................-----...........   8275

Tsw Cc100   GCGTAAAAAATTGTTTTTTTGGTTAACAAATCGTGTGTGTTTTATTGTAATTAATAAAATATTATCTAACTTGTAATTTGAAATTCTTTTGAAGTATAACAGAGCAGCCGTT   8400
Tsw Kim '17 ..........................................................................................................   8395

Tsw Cc100   CTTACAAAATATGCCTCTATTTTAAATGCACAGTTAGAGAGACTGTGGACATAAGAAGTTTTGGTTCACTTTTCATCAAAATCAAACCAAATTAACTATATCGGTTTTTTCA   8520
Tsw Kim '17 ..........................................................................................................   8515

Tsw Cc100   ATATATTATAAATTCATATATTTACATTTTATTGATATCTAATGATTTAACTTACTGTGATAAAATAACATTAAAAGACACACCTGGAATAAGAATTTTATTAGTAATTATT   8640
Tsw Kim '17 ..........................................................................................................   8635

Tsw Cc100   ATGAGAAATGGAGTTGATATTTTGTCAATGTTTTGCAGGTTTCTTGTCCCAACCTGAAAGTGCTAAAACTCCATGAGGCTAACAGCATAACTGCTCTGTGCTCTCACCAACT   8760
Tsw Kim '17 ..........................................................................................................   8755

Tsw Cc100   CCTACTTCAGCAAACTTGAGACATTGGAAGTAGAAAATTGTGGAAAATTGAGACACTTGATGTCTCCATCAGTGGCCAGAGGTCTTCTGAATCTCCGAATATTACTATTAGG   8880
Tsw Kim '17 ..........................................................................................................   8875

Tsw Cc100   AATCAATGGAAGAAGTGATCATAGAAGAAGAACAAGAAGGAGAAGAAAACATGACTAATGAGCCCTTATTTCCCCTGTTGGAAGAGCTGATACTCGATAAGCTACCAAAGCT   9000
Tsw Kim '17 ..........................................................................................................   8995

Tsw Cc100   TCTTTCTGACGAAGCGTGCTCTTGAATTTCCATTCCTCAGAGAAGTGAAGATTCGTAAATGCCCTGAAATGAAGATGCTTGTCCAACAGAGATCTGTCAGTACATCAAGTCT   9120
Tsw Kim '17 ..........................................................................................................   9115

Tsw Cc100   TGAACAATGATGATGAGTTGAAAGTTGTTGATCTCAACAAAGCGATGTTTAATTCTAAGGTTTGTCTTGTGCCGCTTTATAACTAAATGATATACATATTTGAGTAATGCAC   9240
Tsw Kim '17 ..........................................................................................................   9235

Tsw Cc100   CCCTGGTCTATTTCTCTCTTTCTGGTTTAAGAAATGGTGCTTATGAAAAATGCCACATAAGCAGACCATCTTGTTTGATTGCTTGTTCAATTGTCTTGATGACAAAATTTTG   9360
Tsw Kim '17 ..........................................................................................................   9355

Tsw Cc100   CTGTTGACTGGAACTGAAAACCCAATGAATTTCCTAACCCAAATTCAAGAATTAAACACACCCACACTAAATAAACCACAGACCAATACGAAGCAAGCCAAAGATTTCTAGA   9480
Tsw Kim '17 ..........................................................................................................   9475

Tsw Cc100   ATGCTATTCAGCCATTTTTATGTTCAATTAGTCAAACATAGGAATGCTTTAAGTATTAGACATAGATCGGGGATGAAGGTAATATTACGCCCCAAGTTTTGGATTGTTTACA   9600
Tsw Kim '17 ..........................................................................................................   9595

Tsw Cc100   ATCTAATTGTAATAGTAATCTTCATGTAGTTATTTTTTGTTTGTCACCTGATACCTATGTTGATGGAGGTAACAAACATTGGTGAATCTACAATATTATGATCCAAGGGTTC   9720
Tsw Kim '17 ..........................................................................................................   9715

Tsw Cc100   CCAATAACTTTTTTCATAGACTTTTTATCTGTATTAATAATATCTAGAGATTTCTTGAGGTCATTGAAGATTTCGATAAATTTCAGATTTTAAATCTACCTCTCGTTCTAAA   9840
Tsw Kim '17 ..........................................................................................................   9835

Tsw Cc100   TCATTATATGTAACAACTAATATATGGATTTGTTTTTCTCATTTCATCTTGCATTCTTTGTCATTTTAAAAGATCTTAAAGAATGCTATTTTAAAGATGCTTTTTAAAATAC   9960
Tsw Kim '17 .........................................................................................-...............   9954

Tsw Cc100   TTTGAGTAAACACCTCAACTAAATAAGCACTTTTGGCTTCTCAACTTATCACTTAGGGCGTGTTTGGTATGAAGGAAAATGTTTTCATGGAAAGCAAGTTGATTTTTTCTTA   10080
Tsw Kim '17 ..........................................................................................................   10074

Tsw Cc100   TGTTTGGTTATTGAGTGGAAAATATTTTTCGAAAAATATTTTCTAGTGTTTAGTTAGTGAATGAAAATATATTTTTCAGAAAATACATTTTAGTATTTAGCTAAAGAAAATA   10200
Tsw Kim '17 ..........................................................................................................   10194

Tsw Cc100   GAAAATAAGTTTGACGCTGAAAAATATTTTTTAGGTTGTGTTAGGTGGAAAACATTTTTTTTATAAAAATAAGTTGTTTCTTGTTTATATTCTTGTGTTCGGTTATGCAAGTA   10320
Tsw Kim '17 ..........................................................................................................   10314

Tsw Cc100   TTTTTTTAAAGTATTTGTATATGTTTAACATAAAACAATGGGAATGAATAGGGTAAGAAAAAAGTTTAAAATATTTATTAAATATTTTTTT-GAGGAGAGGGTTGGTAGTTG   10439
Tsw Kim '17 ........................................................................................T.................   10434
```

51

```
Tsw Cc100  TATCAGAATATGTGGTAGGGTTAGGTGTGGGGTAAGAAAAAAGTTTTGAATTTTTTTAAAGAAATTATTCTGTTTTTTTTGGGGTGGGGGGGGGGGGGGGGGGGTCAGTACAG 10559
Tsw Kim '17 .......................................................................................---........ 10551

Tsw Cc100  GTAAAAAATTGTTTTTTTGGTTAACAAATCGTGTGTGTTTTATTGTAATTAATAAAATATTATCTAACTTGTAATTTGAAATTCTTTTGAAGTATAACAGAGCAGCCGTTGA 10679
Tsw Kim '17 .............................................................................................. 10671

Tsw Cc100  TACAAAATATGCCTCTATTTTAAATGCACAGTTAGAGAGACTGTGGACATAAGAAGTTTTGGTTCACTTTTCATCAAAATCAAACCAAATTAACTATATCGGTTTTTCAAC 10799
Tsw Kim '17 .............................................................................................. 10791

Tsw Cc100  ATATTATAAATTCATATATTTACATTTTATTGATATCTAATGATTTAACTTACTGTGATAAAATAACATTAAAAGACACACCTGGAATAAGAATTTTATTAGTAATTATTTG 10919
Tsw Kim '17 .............................................................................................. 10911

Tsw Cc100  GAGAAATGGAGTTGATATTTTGTCAATGTTTTGCAGGTTTCTTGTCCCAACCTGAAAGTGCTAAAACTCCATGAGGCTAACAGCATAACTGCTCTGTGCTCTCACCAACTTC 11039
Tsw Kim '17 ....................................................................A.................... 11031

Tsw Cc100  TACTTCAGCAAACTTGAGACATTGGAAGTAGAAAATTGTGGAAAATTGAGACACTTGATGTCTCCATCAGTGGCCAGAGGTCTTCTGAATCTCCGAATATTACTATTAGGAT 11159
Tsw Kim '17 .............................................................................................. 11151

Tsw Cc100  TCAATGGAAGAAGTGATCATAGAAGAAGAACAAGAAGGAGAAGAAAACATGACTAATGAGCCCTTATTTCCCCTGTTGGAAGAGCTGATACTCGATAAGCTACCAAAGCTGG 11279
Tsw Kim '17 .............................................................................................. 11271

Tsw Cc100  TTTCTGACGAAGCGTGCTCTTGAATTTCCATTCCTCAGAGAAGTGAAGATTCGTAAATGCCCTGAAATGAAGATGCTTGTCCAACAGAGATCTGTCAGTACATCAAGTCTCA 11399
Tsw Kim '17 .............................................................................................. 11391

Tsw Cc100  AACAATGATGATGAGTTGAAAGTTGTTGATCTCAACAAAGCGATGTTTAATTCTAAGGTTTGTCTTGTGCCGCTTTATAACTAAATGATATACATATTTGAGTAATGCACTA 11519
Tsw Kim '17 .............................................................................................. 11511

Tsw Cc100  CTGGTCTATTTCTCTCTTTCTGGTTTAAGAAATGGTGCTTATGAAAAATGCCACATAAGCAGACCATCTTGTTTGATTGCTTGTTCAATTGTCTTGATGACAAAATTTTGAG 11639
Tsw Kim '17 .............................................................................................. 11631

Tsw Cc100  GTTGACTGGAACTGAAAACCCAATGAATTTCCTAACCCAAATTCAAGAATTAAACACACCCACACTAAATAAACCACAGACCAATACGAAGCAAGCCAAAGATTTCTAGACT 11759
Tsw Kim '17 .............................................................................................. 11751

Tsw Cc100  GCTATTCAGCCATTTTTATGTTCAATTAGTCAAACATAGGAATGCTTTAAGTATTAGACATAGATCGGGGATGAAGGTAATATTACGCCCCAAGTTTTGGATTGTTTACACC 11879
Tsw Kim '17 .............................................................................................. 11871

Tsw Cc100  CTAATTGTAATAGTAATCTTCATGTAGTTATTTTTTGTTTGTCACCTGATACCTATGTTGATGGAGGTAACAAACATTGGTGAATCTACAATATTATGATCCAAGGGTTCAT 11999
Tsw Kim '17 .............................................................................................. 11991

Tsw Cc100  AATAACTTTTTTCATAGACTTTTTATCTGTATTAATAATATCTAGAGATTTCTTGAGGTCATTGAAGATTTCGATAAATTTCAGATTTTAAATCTACCTCTCGGTTCTAAACT 12119
Tsw Kim '17 .............................................................................................. 12111

Tsw Cc100  ATTATATGTAACAACTAATATATGGATTTGTTTTTCTCATTTCATCTTGCATTCTTTGTCATTTTAAAAGATCTTAAAGAATGCTATTTTAAAGATGCTTTTTAAAATACTT 12239
Tsw Kim '17 .............................................................................................. 12231

Tsw Cc100  TGAGTAAACACCTCAACTAAATAAGCACTTTTGGCTTCTCAACTTATCACTTAGGGCGTGTTTGGTATGAAGGAAAATGTTTTCATGGAAAGCAAGTTGATTTTTTCTTATT 12359
Tsw Kim '17 .............................................................................................. 12351

Tsw Cc100  TTTGGTTATTGAGTGGAAAATATTTTTCGAAAAATATTTTCTAGTGTTTAGTTAGTGAATGAAAATATATTTTTCAGAAAATACATTTTAGTATTTAGCTAAAGAAAATACT 12479
Tsw Kim '17 .............................................................................................. 12471

Tsw Cc100  AAATAAGTTTGACGCTGAAAAATATTTTTTAGGTTGTGTGTTAGGTGGAAAACATTTTTTTTATAAAAATAAGTTGTTTCTTGTTTATATTCTTGTGTTCGTTATGCAAGTATT 12599
Tsw Kim '17 .............................................................................................. 12591

Tsw Cc100  TTTTTAAAGTATTTGTATATGTTTAACATAAAACAATGGGAATGAATAGGGTAAGAAAAAAGTTTAAAATATTTATTAAATATTTTTTTTGAGGAGAGGGTTGGTAGTTGGAG 12719
Tsw Kim '17 .............................................................................................. 12711

Tsw Cc100  CAGAATATGTGGTAGGGTTAGGTGTGGGGTAAGAAAAAAGTTTTGAATTTTTTTAAAGAAATTATTCTGTTTTTTTTGGGGTGGGGGGGGGGGGGGGGGGGGTCAGTACAGGG 12839
Tsw Kim '17 .......................................................................................-------- 12823

Tsw Cc100  AAAAAATTGTTTTTTTGGTTAACAAATCGTGTGTGTGTTTTATTGTAATTAATAAAATATTATCTAACTTGTAATTTGAAATTCTTTTGAAGTATAACAGAGCAGCCGTTGAGA 12959
Tsw Kim '17 .............................................................................................. 12943

Tsw Cc100  CAAAATATGCCTCTATTTTAAATGCACAGTTAGAGAGACTGTGGACATAAGAAGTTTTGGTTCACTTTTCATCAAAATCAAACCAAATTAACTATATCGGTTTTTTCAACCA 13079
Tsw Kim '17 .............................................................................................. 13063

Tsw Cc100  ATTATAAATTCATATATTTACATTTTATTGATATCTAATGATTTAACTTACTGTGATAAAATAACATTAAAAGACACACCTGGAATAAGAATTTTATTAGTAATTATTTGCA 13199
Tsw Kim '17 .............................................................................................. 13183

Tsw Cc100  GAAATGGAGTTGATATTTTGTCAATGTTTTGCAGGTTTCTTGTCCCAACCTGAAAGTGCTAAAACTCCATGAGGCTAACAGCATAACTGCTCTGTGCTCTCACCAACTTCCA 13319
Tsw Kim '17 ....................................................................A................. 13303

Tsw Cc100  CTTCAGCAAACTTGAGACATTGGAAGTAGAAAATTGTGGAAAATTGAGACACTTGATGTCTCCATCAGTGGCCAGAGGTCTTCTGAATCTCCGAATATTACTATTAGGATAC 13439
Tsw Kim '17 .............................................................................................. 13423

Tsw Cc100  AATGGAAGAAGTGATCATAGAAGAAGAACAAGAAGGAGAAGAAAACATGACTAATGAGCCCTTATTTCCCCTGTTGGAAGAGCTGATACTCGATAAGCTACCAAAGCTGGGG 13559
Tsw Kim '17 .............................................................................................. 13543

Tsw Cc100  TCTGACGAAGCGTGCTCTTGAATTTCCATTCCTCAGAGAAGTGAAGATTCGTAAATGCCCTGAAATGAAGATGCTTGTCCAACAGAGATCTGTCAGTACATCAAGTCTCAAA 13679
Tsw Kim '17 .............................................................................................. 13663

Tsw Cc100  CAATGATGATGAGTTGAAAGTTGTTGATCTCAACAAAGCGATGTTTAATTCTAAGGTTTGTCTTGTGCCGCTTTATAACTAAATGATATACATATTTGAGTAATGCACTAAC 13799
Tsw Kim '17 .............................................................................................. 13783

Tsw Cc100  GGTCTATTTCTCTCTTTCTGGTTTAAGAAATGGTGCTTATGAAAAATGCCACATAAGCAGACCATCTTGTTTGATTGCTTGTTCAATTGTCTTGATGACAAAATTTTGAGAA 13919
Tsw Kim '17 .............................................................................................. 13903

Tsw Cc100  TGACTGGAACTGAAAACCCAATGAATTTCCTAACCCAAATTCAAGAATTAAACACACCCACACTAAATAAACCACAGACCAATACGAAGCAAGCCAAAGATTTCTAGACTGA 14039
Tsw Kim '17 .............................................................................................. 14023

Tsw Cc100  TATTCAGCCATTTTTATGTTCAATTAGTCAAACATAGGAATGCTTTAAGTATTAGACATAGATCGGGGATGAAGGTAATATTACGCCCCAAGTTTTGGATTGTTTACACCAG 14159
Tsw Kim '17 .............................................................................................. 14143

Tsw Cc100  AATTGTAATAGTAATCTTCATGTAGTTATTTTTTGTTTGTCACCTGATACCTATGTTGATGGAGGTAACAAACATTGGTGAATCTACAATATTATGATCCAAGGGTTCATGT 14279
Tsw Kim '17 .............................................................................................. 14263

Tsw Cc100  TAACTTTTTTCATAGACTTTTTATCTGTATTAATAATATCTAGAGATTTCTTGAGGTCATTGAAGATTTCGATAAATTTCAGATTTTAAATCTACCTCTCGGTTCTAAACTCT 14399
Tsw Kim '17 .............................................................................................. 14383

Tsw Cc100  TATATGTAACAACTAATATATGGATTTGTTTTTCTCATTTCATCTTGCATTCTTTGTCATTTTAAAAGATCTTAAAGAATGCTATTTTAAAGATGCTTTTTAAAATACTTCT 14519
Tsw Kim '17 .............................................................................................. 14503

Tsw Cc100  AGTAAACACCTCAACTAAATAAGCACTTTTGGCTTCTCAACTTATCACTTAGGGCGTGTTTGGTATGAAGGAAAATGTTTTCATGGAAAGCAAGTTGATTTTTTCTTATTTT 14639
Tsw Kim '17 .............................................................................................. 14623

Tsw Cc100  TGGTTATTGAGTGGAAAATATTTTTCGACAAATATTTTCTAGTGTTTAGTTAGTGAATGAAAATATATTTTTCAGAAAATACATTTTAGTATTTAGCTAAAGAAAATACTTT 14759
Tsw Kim '17 .............................................................................................. 14743
```

```
Tsw Cc100    ATAAGTTTGACGCTGAAAAATATTTTTTAGGTTGTGTTAGGTGGAAAACATTTTTTTTATAAAAATAAGTTGTTTCTTGTTTATATTCTTGTGTTCGTTTTGCAAGTATTAA 14879
Tsw Kim '17  ..............................................................................................................  14863

Tsw Cc100    ATTAAAGTAATTGTATATGTTTAACATAAAACAATGGGAATGAATAGGGTAAGAAAAAAGTTTAAAATATTTATTAAATATTTTTTTGAGGAGAGGGTTGGTAGTTGGAGGG 14999
Tsw Kim '17  ..............................................................................................................  14983

Tsw Cc100    GAATATGTGGTAGGGTTAGGTGTGGGGTAAGAAAAAAGTTTTGAATTTTTTTAAAAAAATTATTCTGTTTTTTTTTTGGGGGGGGGGGGGG—TCAGTACAGGGTGGGGCGTA 15117
Tsw Kim '17  ..........................................................................................GG................  15103

Tsw Cc100    GTTTTTTTGGTTAACAAATCGTGTGTGTTTTATTGTAATTAATAAAATATTATCTAACTTGTAATTTGAAATTCTTTTGAAGTATAACAGAGCAACCGTTGAGAATTTCTTA 15237
Tsw Kim '17  ..............................................................................................................  15223

Tsw Cc100    GCCTCTATTTTAAATGCACAGTTAGAGAGACTGTGGACATAAGAAGTTTTGGTTCACTTTTCATCAAAATCAAACCAAATTAACTATATCGGTTTTTTCAACCAATTTATAT 15357
Tsw Kim '17  ..............................................................................................................  15343

Tsw Cc100    TTCATATATTTACATTTTATTGATATCTAATGATCTAACTTACTGTGATAAAATAACATTAAAAGACACACCTGGAATAAGAATTTTATTAGTAATTATTTGCAAGACATGA 15477
Tsw Kim '17  ..............................................................................................................  15463

Tsw Cc100    GTTGATACCTTGTCAATGTTTTGCAGGTTTCTTGTCCCAACCTGGAATTGCTACGACTCTATAAGGCTAACAGCGTAGGTTCTCTATTCTCTCACCAACTTCCAACTACCTA 15597
Tsw Kim '17  ..............................................................................................................  15583

Tsw Cc100    AACTTGAGACATTGGAAGTAGAAAATTGTGGAAAATTGAGACACTTGATGTCTCCATCAGTGGCCAGAGGTCTTCTGAATCTCCGAATATTAATATTAAGAGACTGTGAATC 15717
Tsw Kim '17  ..............................................................................................................  15703

Tsw Cc100    AAGTGATCACAGAAGAGGAACAAGAAGGAGATGAAATCATGTGTAATGAACCATTATTTCCCCAATTGGAAGAGCTGAAACTTGAAAATCTGCCAAAGCTGAGGCATTTCAT 15837
Tsw Kim '17  ..............................................................................................................  15823

Tsw Cc100    AGCAAGCTCTTGAAATTCCCTTCCTCATAGAAGTGCAGATTCGTAACTGCCCTGAAATGAAGATGCTTGTCCAACAGAGATCTGTCAGTACATCAAGTCTCAAAAGTGTGAA 15957
Tsw Kim '17  ..............................................................................................................  15943

Tsw Cc100    ATGAGTTGAAAGTTGTTGATCTCAACAAAGCGATGTTTAATTCTAAGGTTTGTCTTGTGCCGCTTTATAACTAAATGATATCATATTTGAGTAATGCACTAACAGATCCCTG 16077
Tsw Kim '17  ..............................................................................................................  16063

Tsw Cc100    CTCTCTTTCTCGGTTTAAGAAATGGTGCTTATGAAAAATGCCACATAAGCAGACCATCTTGTTTGATTGCTTGTTCAATTGTCTTGATGACAAAATTTTGAGAAAACGCTGTT 16197
Tsw Kim '17  ..............................................................................................................  16183

Tsw Cc100    CTGAAAACCCAATGAATTTCCTAACCCAAATTCAAGAATTAAACACACCCACACTAAATAAACCACAGACCAATACGAAGCAAGCCAAAGATTTCTAGACTGAAACCATGCT 16317
Tsw Kim '17  ..............................................................................................................  16303

Tsw Cc100    ATTTTTATGTTCAATTAGTCAAACATAGGAATGCTTTAAGTATTAGACATAGATCGGGGATGAAGGTAATATTACGCCCCAAGTTTTGGATTGTTTACACCAGTTCAATCTA 16437
Tsw Kim '17  ..............................................................................................................  16423

Tsw Cc100    AGTAATCTTCATGTAGTTATTTTTTGTTTGTCACCTGATACCTATGTTGATGGAGGTAACAAACATTGGTGAATCTACAATATTATAATCCAAGGGTTCATGTGAACCCAAT 16557
Tsw Kim '17  ..............................................................................................................  16543

Tsw Cc100    TTCATTGACTTTTTATCTGTATTAATAATATCTAGAGATTTCTTGAGGTCATTGAAGATTTCGATAAATTTCAGATTTTAAATCTACCTCTCGTTCTAAACTCTTCTTTCAT 16677
Tsw Kim '17  .-............................................................................................................  16662

Tsw Cc100    ACAACTAATATATGGATTTGTTTTTCTCATTTCATCTTGCATTCTTTGTCATTTTAAAAGATCTTAAAGAATGCTATTTTAAAGATGCTTTTTAAAATACTTCTTCTATTTG 16797
Tsw Kim '17  ..............................................................................................................  16782

Tsw Cc100    CCTCAACTAAATAAGCACTTTTGGCTTCTCAACTTATCACTTAGGGCGTGTTTGGTATGAAGGAAAATGTTTTCATGGAAAGCAAGTTGATTTTTTCTTATTTTCTTATGTT 16917
Tsw Kim '17  ..............................................................................................................  16902

Tsw Cc100    GAGTGGAAAATATTTTTCGAAAAATATTTTCTAGTGTTTATTAGTGAATGAAAATATATTTTTCAGAAAATACATTTTAGTATTTAGCTAAAGAAAATACTTTTGTAGAAAA 17037
Tsw Kim '17  ..............................................................................................................  17022

Tsw Cc100    ACGCTGAAAAATATTTTTTAGGTTGTGTTAGGTGGAAAACATTTTTTTTTATAAAAATAAGTTGTTTCTTGTTTATATTCTTGTGTTCGTTATGCAAGTATTAAAATATTTT 17157
Tsw Kim '17  ..............................................................................................................  17142

Tsw Cc100    ATTTGTATATGTTTAACATAAAACAATGGGAATGAATAGGGTAAGAAAAAAGTTTAAAATATTTATTAAATATTTTTTTGAGGAGAGGGTTGGTAGTTGGAGGGATGTGTCA 17277
Tsw Kim '17  ..............................................................................................................  17262

Tsw Cc100    GGTAGGGTTAGGTGTGGGGTAAGAAAAAAGTTTTGAATTTTTTTAAAAAAATAATTCTTTTTTTTTTGGGTGGGGGGGGGGGGGGGGG—TCAGTACAGGGTGGGGCGTAAAA 17396
Tsw Kim '17  .................................................................................G........................  17382

Tsw Cc100    TTTTGGTTAATAAATCGTGTGTGTTTTATTGTAATTAATAAAATATTATCTAACTTGTAATTTGAAATTCTTATGAAGTATAACAGACCAGTCGTTGAGAATTTCTTACAAA 17516
Tsw Kim '17  ..............................................................................................................  17502

Tsw Cc100    CTATTTTAAATGCACATTTAGAGAGACTGTGGACATAAGAAGTTTTGGTTCACTTTTTATCAAAATCAAACCAAATTAACTATATCGGTTTTTTCAACCAATTTATATATTA 17636
Tsw Kim '17  ..............................................................................................................  17622

Tsw Cc100    TATATTTACATTTTATTGATATCTAATGATTTAGTCTAAAAATGATTTAACTTACTCTGATAAAATAACATTAAAAGACACACCTGAAATAAGAATTTTATTAGTAATTCTT 17756
Tsw Kim '17  ..............................................................................................................  17742

Tsw Cc100    ATGAGAAATGGAGTTGATACCTTGTCAATGTTTTGCAGGTTTCTTGTCCCAACCTGGAATTGCTACGACTCTATAAGGCTAACAGCGTAGGTTCTCTATTCTCTCACCAACT 17876
Tsw Kim '17  ..............................................................................................................  17862

Tsw Cc100    CCTACTTCAGCAAACTTATGCAACTAGGAATACGGAATTGTGGAAAATTGAGACACTTGATGTCTCCATCAGTGGCCAGAGGTCTTCTGAATCTCCGAATATTAATATTAGG 17996
Tsw Kim '17  ..............................................................................................................  17982

Tsw Cc100    AATCAATGGAAGAAGTGATCACAGAAGAGGAACAAGAAGGAGAAGAAAACATGACTAATGAGTCCTTATTTCCCTGCTTGGAAATGCTGAAACTTGAAAATCTGCCAAAGCT 18116
Tsw Kim '17  ..............................................................................................................  18102

Tsw Cc100    TCTTTCTGACGAAGCGTGCTCTTGAATTTCCATTCCTCAGAGAAGTGAAGATTCGTACATGCCCTGAAATGAAGATATTTGTCCAACATGGATCTGTGAGTACACCAAGTCT 18236
Tsw Kim '17  ..............................................................................................................  18222

Tsw Cc100    TGAACAATGATGATGAGGTAAAAGTAGATGATCTGAACGAATGGATACATCAGAGGTTCAATTCTAAGGTTTGTCTTGTGGTGCTAAATGACTGTATAACTAATTAACTCTA 18356
Tsw Kim '17  ..............................................................................................................  18342

Tsw Cc100    TTTCCATCTTAACTTGATAGTAATAGTAATATTTACATACTCGTTTATGTTTGTCACTTGAATTAGCTTCCTTTTACCTGCCCCAAATGTTTTATCTTTCAACGAAATATTG 18476
Tsw Kim '17  ..............................................................................................................  18462

Tsw Cc100    GTTTTTTTGTAATCTCTACTCTCAGGAAGAAGATGGAAGTGAATCTGAATCTTCTCAGGAAGAAGATGGAAGCGAATCTGAAGCTTCTCATGAAGATGGAAGCAAATGTGAAG 18596
Tsw Kim '17  ..............................................................................................................  18582

Tsw Cc100    GAAGAAGATAGAAGCGAATCTGAAGATTCAATGAGTGACGGCAGAAGCTAA                                                          18651
Tsw Kim '17  .................................................                                                          18637
```

Table S3.1 – Primers used in this research

ID	Sequence	Purpose
CC20 T7 F	GGCATGACTCAGGTTGTTGC	BAC library screening
CC20 T7 R	ACACTTGAATACCAGGCGGG	BAC library screening
CC33 PCCR F	AGCGTGTGTGATCCGCACCA	BAC library screening
CC33 PCCR F	CCGCCTTTCTCCTCTAATCCCGC	BAC library screening
CC33 PCCR F	CTCCATGGAAAGGATTGAAAGT	BAC library screening
CC33 PCCR F	GCAGGCAGCCCACCTCCTTG	BAC library screening
CC33 PCCR R	ACCCCAACCCGTTAAAGAAAAGCCCA	BAC library screening
CC33 PCCR R	CAAGGAAGGTGGGGCTGCCTGC	BAC library screening
CC33 PCCR R	TGGGACTCTTGGCATGCGGC	BAC library screening
CC33 PCCR R	TGGTGCGGATCACACACGCT	BAC library screening
CC33 T7 F	GCCGCCTATCCACTGCACCA	BAC library screening
CC33 T7 F	GTTGGGAAGCACCTATAACAG	BAC library screening
CC33 T7 F	GTTGGTGGTGGTCACAGTGTTCGT	BAC library screening
CC33 T7 R	CCCCACCCCTCCAACCAACC	BAC library screening
CC33 T7 R	CCGAGCTGGAGACTCCTCCCC	BAC library screening
CC33 T7 R	GCAGTGGATAGGCGGGCTCCC	BAC library screening
CC33 T7 R	TACCTCTCATCACTTCCTAGTG	BAC library screening
CC33 T7 R	TGGTTTCAGACATGTTGGTGGTGGT	BAC library screening
CC369 T7 F	AAGATCCTCACGCCTCT	BAC library screening
CC369 T7 F	CCCACGTGTAATGGGAGCTT	BAC library screening
CC369 T7 F	GCTCGTGAGGATCTTGTTGCC	BAC library screening
CC369 T7 F	GCTTCCGCTGCGCCACTCTG	BAC library screening
CC369 T7 F	GGGAAGAAGCTGATGGCTGT	BAC library screening
CC369 T7 R	ACTCTCACGAGCACCAGTTG	BAC library screening
CC369 T7 R	CTATGCGGCGGCATCAGAGCAGA	BAC library screening
CC369 T7 R	GCAGCTTGATCGTAAGCACG	BAC library screening
CC369 T7 R	GCTTCCGCTGCGCCACTCT	BAC library screening
CC37 PCCR F	CGGGACTGTAGGTGTTTGGGT	BAC library screening
CC37 PCCR R	GGGATGGCAAAACTGATCAAC	BAC library screening
CC386 PCCR F	ACACCCTGCATTTCGGGCTAG	BAC library screening
CC386 PCCR F	CCATACGTTCATACGCCA	BAC library screening
CC386 PCCR F	GGGACGAATGATCCCAAGGGGG	BAC library screening
CC386 PCCR F	TCGGTCACAGATCCATGC	BAC library screening
CC386 PCCR F	TGCGTTGGTTATCCCAGAGCC	BAC library screening
CC386 PCCR R	ACGCGCGATTGGTTAGGCGA	BAC library screening
CC386 PCCR R	CGCCGCACTTTGCATCGCAT	BAC library screening
CC386 PCCR R	CTCTGGGATAACCAACGC	BAC library screening
CC386 PCCR R	CTTGTCTTGAGTTGCAGG	BAC library screening
CC386 PCCR R	GCGCCGCATTGTGTATCGCA	BAC library screening

ID	Sequence	Purpose
CC386 T7 F	AATGCAGGTGCTCACCGCCC	BAC library screening
CC386 T7 R	CCCCTAGCACCCCCACCACA	BAC library screening
CC403 PCCR F	GCCAGGCCAAGTCGCTCCAG	BAC library screening
CC403 PCCR R	GGCAAAGCATGCCGAACTGGC	BAC library screening
CC403 T7 F	GGCCACATCGTCACACTGTTGGT	BAC library screening
CC403 T7 R	AGGAGGCAGAGGTGGCGTCC	BAC library screening
CC411 PCCR F	CCTGTCCCGCACCACTTGCA	BAC library screening
CC411 PCCR F	CGGTGGACGGGGTGGAGGAA	BAC library screening
CC411 PCCR F	GCTGTCATTTCGACCCCACTT	BAC library screening
CC411 PCCR R	AGCGCGTCGTCAATGCCCTT	BAC library screening
CC411 PCCR R	CCCGGGCGTGTTTTGGGACT	BAC library screening
CC411 PCCR R	TGTGGCCAGAGAGCAAGGCTG	BAC library screening
CC411 T7 F	CGAGGCTCAGCCCATCCAACA	BAC library screening
CC411 T7 R	TGTGCTGACTTAGGCCGCCA	BAC library screening
CC424 PCCR F	ACGAGTCCAAARGCTGCTTCA	BAC library screening
CC424 PCCR F	ACGAGTCCAAATGCTGCTTCA	BAC library screening
CC424 PCCR F	GGTTGCAGTGCTCTAGGAC	BAC library screening
CC424 PCCR F	TCCTCCGGGATCAAGTGGACA	BAC library screening
CC424 PCCR F	TGCCCGTAGACAGGGCTCTCC	BAC library screening
CC424 PCCR F	TGCTGCTTCACCAATCCTCTC	BAC library screening
CC424 PCCR R	ACTAGACCGTAGCGCCGGGG	BAC library screening
CC424 PCCR R	AGCCTTACGTAGGTGGCTCCA	BAC library screening
CC424 PCCR R	ATTTCTTGCAAATGTTCCTG	BAC library screening
CC424 PCCR R	GCGCCGGGGCAAAGAGAAGA	BAC library screening
CC424 PCCR R	GTAGCGCCGGGGCAAAGAGA	BAC library screening
CC424 PCCR R	TGCCTCGTGCGGGTTCACAG	BAC library screening
CC441 PCCR F	CGGTGGACGGGGTCGAGAGGAA	BAC library screening
CC441 PCCR R	TGTTTTGCTTTCCCCCAACTTCCCG	BAC library screening
CC441 T7 F	TGTTTCGCCTAGACCCCTTGGCA	BAC library screening
CC45 PCCR R	TGTGCAACGCGTGGACATCACT	BAC library screening
CC45 T7 F	CCACAAGCTTCGCCTTCTGGGG	BAC library screening
CC45 T7 F	TGCACAGTGAGCCTCATCTCCAGT	BAC library screening
CC45 T7 R	GGCCGGTATGACACCTCATT	BAC library screening
CC45 T7 R	TTGCCAGGCCAAGTCGCTCC	BAC library screening
CC47 T7 F	GTTGCAGTAATGATGAGGCGC	BAC library screening
CC47 T7 F	TCTCCAGCAGCAACAGTAGC	BAC library screening
CC47 T7 R	ACGGGACTGTAGGTGTTTGG	BAC library screening
CC47 T7 R	GATAGGGTTCTCTCCAGGCC	BAC library screening
CC525 PCCR F	ACGGGTTCATCTCATGACGG	BAC library screening
CC525 PCCR R	GGCATCGGCAAAACAACACT	BAC library screening

ID	Sequence	Purpose
CC525 T7 F	AGCTGCATGTTGTGCTTTGG	BAC library screening
CC525 T7 F	CACAACCCTCCCCCTACAAC	BAC library screening
CC525 T7 R	ACCGACATGAACAGGCAGAG	BAC library screening
CC525 T7 R	GGGTGGGAGTTCTTTCCCAG	BAC library screening
CC535 PCCR F	AGCCTTGCTGGACATCACAA	BAC library screening
CC535 PCCR F	AGGTTTACAACCCTGTTGATTCAT	BAC library screening
CC535 PCCR F	AGGTTTACAACCCTGTTGATTCATT	BAC library screening
CC535 PCCR R	GAGAGACTGCACGACCAACA	BAC library screening
CC535 PCCR R	TCCTCTGGGCGTTCACTTTC	BAC library screening
CC535 PCCR R	TGAGAGACTGCACGACCAAC	BAC library screening
CC632 PCCR F	TGCTAGACCACCTAGGCCAT	BAC library screening
CC632 PCCR R	ACCCCGTAGGTCGTGTCATA	BAC library screening
CC632 T7 F	ACCATGCAGAAGCACACAGA	BAC library screening
CC632 T7 R	ACACCCTACGGCTTTTTCGT	BAC library screening
CC652 F	AGGGGCTGCCGCTTGCAATT	BAC library screening
CC652 F	TGCAGAAGGACAGCCGCGTC	BAC library screening
CC652 R	AGCCACGTCACGGACCACAT	BAC library screening
CC652 R	GCCTCCCACGAAGGTGTGGC	BAC library screening
CC677 PCCR F	AGCTCGTAAACACCACGGCCT	BAC library screening
CC677 PCCR R	CGAGTCTCAAGTTCGACGCGC	BAC library screening
CC677 T7 F	CTCCCTCTTCCCGCACGGGT	BAC library screening
CC677 T7 R	AGGCTGAGGGACCAAGACGGT	BAC library screening
CC685 T7 F	TGCCTATTCAGATGAAGTACCCAC	BAC library screening
CC69 PCCR F	GGGGTGGGGAAGAGGGGTGA	BAC library screening
CC69 PCCR R	TGCATGTCATAGTCCACACAACCACC	BAC library screening
CC69 T7 F	TGCACGGGTCCAGGGAAGGG	BAC library screening
CC69 T7 R	AGGGGTCATGGGTTCAAGCCGA	BAC library screening
CC9 T7 F	ACAGAGTTGCAATCGCC	BAC library screening
CC9 T7 F	CAATCGCCCCGTTGACAAAA	BAC library screening
CC9 T7 R	AATCCTCAGCAGAAGGGCTG	BAC library screening
CC9 T7 R	GGATCCTCAGCAGAAGGGCT	BAC library screening
T7 seq	TAATACGACTCACTATAGGG	BAC contig sequencing
PCCR seq	TACGCCAAGCTATTTAGGTGAGA	BAC contig sequencing

Chapter 4

Tsw – a case study on structure-function puzzles in plant NLRs with unusually large LRR domains

I.L. van Grinsven*, E.C. Martin*, A.J. Petrescu, and R. Kormelink
* These authors have contributed equally to this work and share first authorship.

Abstract

Plant disease immunity heavily depends on the recognition of plant pathogens and the subsequent activation of downstream immune pathways. Nod-like receptors are often crucial in this process. *Tsw*, a Nod-like resistance gene from *Capsicum chinense* conferring resistance against Tomato spotted wilt virus (TSWV), belongs to the small group of Nod-like receptors with unusually large LRR domains. While typical protein domain dimensions rarely exceed 500 amino acids due to stability constraints, the LRR of these unusual NLRs range from 1,000 to 3,400 amino acids and contain over 30 LRR repeats. The presence of such a multitude of repeats in one protein is also difficult to explain considering protein functionality. Interactions between the LRR and the other NLR domains (CC, TIR, NBS) take place within the first 10 LRR repeats, leaving the function of largest part of the LRR structure unexplained. Herein we discuss the structural modeling limits and various aspects of the structure-function relation conundrums of large LRRs focusing on *Tsw*, and raise questions regarding its recognition of its effector NSs and the possible inhibition on other domains as seen in other NLRs.

Introduction

The basics of a properly working immune system in all organisms is the ability to distinguish self and non-self molecules, after which harmful intruders are forcefully removed while the own cellular integrity remains intact. Conserved microbial patterns, so-called microbe/pathogen associated molecular patterns (MAMPs, PAMPs), are recognized by pattern recognition receptors (PRRs) on the cell surface. However, effectors secreted by pathogens can suppress this pathway and lead to an infection. Plants in turn have acquired cytoplasmic resistance (*R*) genes, or nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), that can detect effectors and trigger an immune response (Jones and Dangl, 2006). NLRs are a diverse group of proteins. Most NLRs contain a variable N-terminal domain, as well as a nucleotide-binding domain (NBD), and a C-terminal leucine rich repeat (LRR) domain. The different N-terminal domains are used to classify NLRs into three different groups: the coiled-coil (CC) NLRs (CNLs), the Toll-interleukin receptor (TIR) NLRs (TNLs), and the Resistance to Powdery Mildew 8-like domain (RPW8)) NLRs (RNLs). The remainder, comprised of NLRs without an N-terminal domain, are classified as NBS-LRR, or NL, receptors (Duxbury et al., 2021).

The central NBD of NLRs contains a NB-ARC, a nucleotide binding (NB) adaptor shared by human **A**paf1, plant **r**esistance proteins, and nematode **C**ED-4 (ARC) (van der Biezen and Jones, 1998). The NB-ARC domain is a molecular activation switch that can bind adenosine 5′-diphosphate (ADP) and adenosine 5′-triphosphate (ATP) (van Ooijen et al., 2008; Takken and Tameling, 2009; Maekawa et al., 2011; Williams et al., 2011). In the absence of the pathogen effector, the LRR of an archetype NLR forms a tight horseshoe shape in which the N-terminal domain and the NB-ARC domain are nestled (Kobe and Kajava, 2001). The first few repeats of the LRR are in close interaction with conserved motifs found in other domains, which keeps the protein in its inactive ADP state and prevents auto-activity

(Slootweg et al., 2013; Wang et al., 2019a, 2019b). It is oftentimes the LRR that interacts (in) directly with effectors, which lifts the auto-inhibition of the NB-ARC by the LRR (Qi et al., 2012; Slootweg et al., 2013; De Oliveira et al., 2016). This recognition allows for a relaxation of the protein conformation, enabling ADP-ATP exchange, promoting oligomerization with (other) NLRs and triggering downstream immune pathways to limit pathogen spread.

Recently, cryogenic electronic microscopy (cryo-EM) studies showed that the Arabidopsis CNL HOPZ-ACTIVATED RESISTANCE 1 (ZAR1, confers resistance against *Pseudomonas syringae* (Laflamme et al., 2020), *Xanthomonas campestris* and *X. perforans* (Wang et al., 2015a; Schultink et al., 2019)) transforms upon activation from its inactive form, through an intermediate state, into a pentamer, with its N-terminal CC domain at the heart, while the C-terminal LRR domain extends outward (Wang et al., 2019a). This resistosome is structurally quite similar to the mammalian apoptosomes and inflammasomes that are formed in response to stress and trigger cell death upon activation. It is currently not known what exact downstream pathways are triggered by the formation of the ZAR1 pentamer. However, similarities of the resistosome with pore-forming proteins, as well as recent *in vivo* studies, suggests it triggers changes in Ca^+ ion flux, eliciting other signaling responses (Bi et al., 2021). Following the discovery of the ZAR1 resistosome, two other cyro-EM studies revealed resistosome formation by the Arabidopsis TNL *RPP1* (resistance against *Peronospora parasitica*) and the *Nicotiana benthamiana* TNL *Roq1* (resistance against *Xanthomonas* sp. and *Pseudomonas* sp.) (Ma et al., 2020; Martin et al., 2020b).

The LRR domain of all three aforementioned NLRs have a typical length (ZAR1: 313 amino acids (aa), RPP1: 462, *Roq1*: 516) and number of LRR repeats (ZAR1: 13, RPP1: 21, ROQ1: 21) that form the archetype horseshoe shape (Wang et al., 2019b, 2019a; Ma et al., 2020; Martin et al., 2020b). The CNL *Tsw* from *Capsicum chinense* (resistance against Tomato spotted wilt orthotospovirus; TSWV), has an LRR domain of an estimated 1,652 aa (Kim et al., 2017a), and thus belongs to the small group of NLRs with unusually large LRR domains. Typical protein domain dimensions rarely exceed 500 aa due to stability constraints (Xu and Nussinov, 1998), while these unusual LRRs range from 1,000 to 3,400 aa and contain over 30 LRR repeats. The added value of an extreme large LRR is unknown. It has been hypothesized that the LRR size (1,951 aa) of *Rps11* from soybean is linked to its ability to resistance against a broad range of *Phytophthora* sp. (Wang et al., 2021). However, (so far) *Tsw* only recognizes the effector NSs from TSWV, and its resistance can be broken with the change of 1 aa (Almási et al., 2017). The protein domains between which the (in)direct interaction of *Tsw* and NSs takes place are unknown, which leaves the function of largest part of the LRR structure unexplained. This paper discusses the structural modeling limits and various aspects of the structure-function relation conundrums of large LRRs while focusing on *Tsw*, and poses questions regarding its recognition of its effector NSs and the possible inhibition on other domains as seen in other NLRs.

Materials and Methods

Sequence analysis of Tsw and Tsw homologs

The amino acid and coding sequences of *Tsw* (A0A1C9TCM9_CAPCH) published by Kim et al. were used to search the *C. chinense* 'PI159236' scaffolds and assembled genome (Chinense.v.1.2 from the Pepper Genome Platform) for the full-length sequence of *Tsw*. The amino acid sequence of Tsw was also used to search the at the time most recent version of the genome of the susceptible *C. annuum* cv. CM334 version 1.55 for the full-length non-functional variant of *tsw* (Kim et al., 2017a).

Tsw homologs sharing more than 30% identity and 50% coverage were retrieved from the NLRscape database (NLRscape, 2022) and analyzed with NLRscape webserver tools. These were further curated by eliminating CC/NBS incomplete and motif missing sequences, and by retrieving their secondary structure propensities and LRR repeat profiles precomputed with LRRpredictor (Martin et al., 2020a) from NLRscape DB. Variability profiles were raised over this cleaned Tsw family set by computing Kullback–Leibler divergence of the alignment with respect to Blosum62 background distribution using Logomaker API (Tareen and Kinney, 2020) and secondary structure consensuses were generated for graphics display.

Large LRRs identification

A set of ~39,469 nonredundant LRR domains (90% identity cutoff) originating from plant NLR sequences were retrieved from NLRscape (NLRscape, 2022) and further processed for analysis. All sequences containing annotated LRR domains at 90% redunacy were collected using the online selection tools of NLRscape. Sequence data, domain annotations, and individual LRR repeat predictions computed with LRRpredictor (Martin et al., 2020a) were retrieved from NLRscape public data sets (NLRscape, 2022). Statistical analyses of the data were performed using in- house scripts. Figures were generated using Plotly (Plotly Technologies Inc., 2015), Matplotlib (Hunter, 2007) and ETE3 (Huerta-Cepas et al., 2016) libraries.

Homology modelling

Due to some problems that will be clarified below, homology models were built herein by two distinct methods: (1) by using the recently developed AI-based AlphaFold (Jumper et al., 2021) and (2) by using a system adapted modeling workflow that has previously been used successfully (Slootweg et al., 2013; Kozuki et al., 2017; Zhang et al., 2019). As RAM requirements grow exponentially with the sequence length, AlphaFold failed to generate a model for the long functional *C. chinense Tsw* form comprising of 2,116 aa. In addition, all models of the shorter susceptible (tsw) variant of Tsw in *C. annuum* (1,437 aa) proposed by AlphaFold were in the open ATP active state and none in the inactive ADP resting state, which seems unrealistic. Furthermore, we found that some of the delineation of LRR repeats as well as the local 3D structure in the LRR region proposed by AlphaFold is

problematic. This prompted us to develop alternate models allowing to explore a larger configurational space consistent with the local sequence propensity as well as the overall architecture of the LRR domain as seen in Figure 4.3.

The alternate models of the *Tsw*-CC, -NBS and -LRR domains were built using Modeller v10.1 (Webb and Sali, 2016) primarily based on ZAR1 ADP-bound cryo-EM conformation — PDB: 6j5w (Wang et al., 2019b). The CC and NBS domains were modelled based exclusively on the ZAR1 template as it shows the highest similarity to the *Tsw* sequence with: CC ~15% and NBS ~30% sequence identity. By contrast the LRR domain was modelled based on the repeat delineation provided by LRRpredictor (Martin et al., 2020a) using the Joint Fragment Remote Homology Modeling (JFRHM) approach (Slootweg et al., 2013) starting from: ZAR1-LRR, FLS2-LRR (PDB: 4mn8) (Sun et al., 2013) and SCHENGEN3-LRR (PDB: 6s6q)(Okuda et al., 2020). LRRpredictor indicates that the long functional form of TSW-LRR domain has 57 LRR repeats organized as follows: a starting core region of 15 repeats followed by 7 consecutive quasi-identical blocks consisting each of 6 LRR repeats (Figure 4.6). The first 10 repeats of the core region were modelled using ZAR1 as template by preserving its curvature and all its contacts with CC-NBS. The rest of the LRR domain was modeled using the other two templates depending on the local best match. To smooth out the transition between these two regions with distinct curvatures, a 'buffer' region consisting of repeats 10-15 was relaxed by gradually modifying the local radius using a template mixture protocol allowing the passage from the tighter ZAR1 region, r1-r10, to the broader r15-r57 rest. The C-terminal acidic tail of LRR was not covered in the model due to the lack of 3D templates and the low confidence of the secondary structure consensus within this area. The model of the shorter susceptible variant tsw was obtained using the optimized long Tsw model, as the two sequences are highly homologous (>83.6% identity). These raw models were subjected to an optimization workflow consisting of interactive rounds of energy minimization, simulating annealing and short-term Molecular Dynamics trials followed by modelling refinements. The compliance of geometric features to known 3D structures was probed using Molprobity (Williams et al., 2018). To further optimize its stability, the model was further subjected to an equilibration stage performed in explicit TIP3P water and a 0.5 nM NaCl neutral charged environment. The system was initially minimized and gradually heated from 0 to 300K and further subjected to 5 equilibration stages of gradually relaxing restraint profiles: from stronger backbone constraints in the first stage, to no constraints in the last stage. After the equilibration, the system was subjected to a long-run Molecular Dynamics stability trial. Simulations were performed using NAMD 3.0a8 (Phillips et al., 2005, 2020) and CHARMM 16m (Huang et al., 2017) forcefield using periodic boundary conditions and Particle Mesh Edwards electrostatics with a 12Å cutoff, 10Å switch and 14Å pairlist distances for nonbonded interactions. Constant temperature (300K) and pressure (1 atm) control was attained by using the Langevin integrator and Langevin piston. Trajectory analyses were performed using VMD (Humphrey et al., 1996) and the 3D molecular graphics displayed in

Chapter 4

the figures was generated using PyMOL (The PyMOL Molecular Graphics System, Version 2.2.3 Schrödinger, LLC).

Results

Large LRR domains in plant NLRs

To allow for the comparison of LRR domain lengths in plants, a set of ~39,469 LRRs from plant NLR sequences edged at 90% identity redundancy (NLRscape, 2022) was analyzed. Over three quarters of the retrieved LRRs span between 200-600 aa and are mainly comprised of less than 20 LRR repeats. The LRR length distribution peak is found at around 400 aa, corresponding to ~14-16 LRR repeats on average (Figure 4.1A). Both domain length and repeat number follow a unimodal distribution centered at 15 repeats and ~380 aa, with solid tails on both sides. The distribution is strongly right-sided skewed, with an elongated tail corresponding to large LRR domains. This roughly adds up to around 10% of LRR domains that are unusually long, with lengths of 1,000 to 3,400 aa that are organized in 30 or more repeats (Figure 4.1A, B). Large LRR domains are found mainly in CNL and NL receptors which display multiple examples of extremely large LRRs, of over 2,000 aa. By contrast, the RNL and TNL groups have far more regular LRRs (< 500 aa), with only few examples of more than 1,000 aa – all found in the TNL class. RNLs are quite rare, and hence far less represented in plants.

From a taxonomic point of view the vast majority of large LRRs originates from *Pentapetalae*, in particular from within the rosid clade: the fabids, with the bulk stemming from the *Fabales* order, and the malvids; *Sapindales, Malvales* and *Myrtales* orders. Interestingly, the *Brassicales* order in the malvids clade does not share this trend. Care should be taken in the interpretation of these results on the unequal statistical support given the uneven taxonomic distribution of NLRs in the database (shown in Figure 4.1C, last column).

Such large dimensions of LRR domains are puzzling as they are far larger than typical protein domains, which rarely exceed 500 aa due to stability problems (Xu and Nussinov, 1998). For instance, at 40% identity redundancy, 99.8% of the domains in the CATH database (Sillitoe et al., 2021) are shorter than 600 aa, with only 5 examples larger than 1,000 aa out of a total of 31,878, which is less than 0.02%. (Figure 4.1B, CATH column). These large dimensions are favored by the loosely and modular solenoidal LRR architecture, which, in contrast to the more globular architectures, offer flexibility and elasticity to the overall fold (Kobe and Kajava, 2001). In addition, large continuous LRR horseshoes are in some cases weakened or disrupted by relatively long insertions incompatible with the LRR fold, which potentially break the LRR domain into two or more sub-regions.

The manner in which such LRR subregions would fold with respect to each other is another still open-ended question. Shorter linkers between LRR fragments generally

64

tend to maintain the overall horseshoe shape of the domain with minor disruptions. By contrast, longer linkers that allow for rotations and translation of the downstream LRR elements might generate a wide range of overall potential conformations – from a perfect fusion of consecutive LRR fragments, with large insertion 'isles' as seen in some X-ray crystal structures such as the extracellular BRI1 (Hohmann et al., 2018) or PSKR1 (Wang et al., 2015b), to even a potential side-to-side LRR-LRR fragment clustering – similar to those seen in some protein-protein interactions. In assessing the LRR domain continuity, a special indicator of the local LRR stability is the embedment of the **L-x-x-L-x-L** pattern into a larger pattern L-x-x-(**L-x-x-L-x-L**)-x-x-N/C (Kajava, 1998) that provides further stability and structural reinforcement.

Large LRR domains are also difficult to explain from a functional point of view. Experimental data, from both solved 3D structures and *in vivo* experiments, indicate that in a large and diverse set of NLR examples interactions of the LRR with the rest of NLR domains (CC, TIR, NBS) primarily take place within the first four to eight LRR repeats (Qi et al., 2012; Slootweg et al., 2013). In ZAR1 cryo-EM structures, the LRR-CC interaction occurs
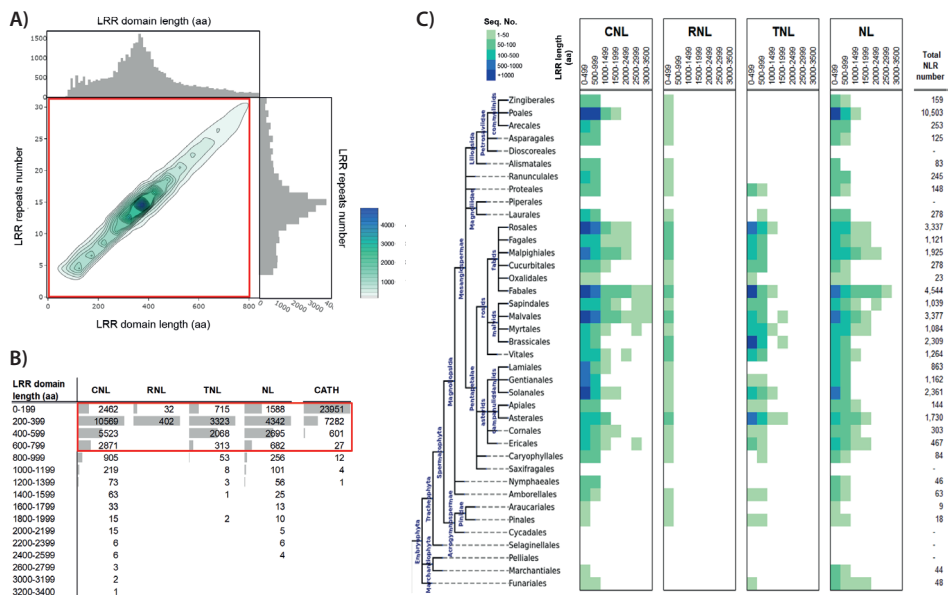
Figure 4.1 – Leucine rich repeat (LRR) domain length statistics over plant Nod-like resistance genes (NLRs). A) 2D histogram of the LRR domain length (aa) versus the corresponding number of LRR repeats. Individual 1D histograms of the LRR domain length and LRR repeats number are shown alongside the x and y axis of the 2D histogram plot. Heatmap color scale and contours legend is shown on the right side of the panel. **B)** LRR domain length (aa) distribution in the main NLR classes: the coiled-coil (CC) (CNLs), the Resistance to Powdery Mildew 8-like domain (RPW8)) NLRs (RNLs) the Toll-interleukin receptor (TIR) (TNLs), and NLRs without an N-terminal domain (NLs). The last column (CATH) shows the domain lengths computed on experimental structural data from PDB using the annotated domains from the CATH database at 40% identity redundancy. **C)** LRR domain length in the four main NLR classes (CNL, RNL, TNL and NL) spread across taxonomic orders. The species tree on the left shows the evolutionary relationships between plant orders and their parent clades.

within the first four LRR repeats, while the LRR-NB-ARC interaction is consolidated mainly within the first eight LRR repeats – in tight contact with the ARC2 subdomain of NB-ARC (Wang et al., 2019b, 2019a). Sparse contacts were also shown to shape up between the further 5 LRR repeats and the beta-turn-beta loop structure of the ARC2 domain. All these interdomain contacts are critical for a proper functioning of these receptors, as shown by swap experiments between sections of the LRR domain of different *R* genes. Compelling experimental data demonstrate the importance of the compatibility between the first LRR repeats and the rest of the NLR (Qi et al., 2012; Slootweg et al., 2013, 2017, 2018). For instance, LRR swaps between potato homologs *Rx1* and *Gpa2,* which despite their 88% sequence identity recognize unrelated pathogens, showed that successful phenotypic LRR exchanges were only obtained when preserving the original CC-NBS-first 3 LRR repeats, while the fusion of the foreign rest of the LRR domain converts the pathogen recognition (Slootweg et al., 2013, 2017, 2018).

The strong preference for 10-16 LRR repeats observed within the dataset of LRR domains (Figure 4.1A) is consistent with the existing experimental and structural data, yet it leaves a very large portion of the LRR structure unexplained in cases of extremely large LRR domains.

Tsw homologs group

Starting from the Tsw sequence from *C. chinense,* a set of distant homologs at above 30% identity with Tsw were retrieved using the NLRscape tools (NLRscape, 2022). The set was trimmed by eliminating the highly redundant sequences by imposing a threshold of 90% identity between the identified homologs obtaining a set of 200 sequences, which we further refer to as the extended homologs group. Based on the identity matrix of the alignment, a dense group emerges, corresponding to a set of close homologs of the *Capsicum* Tsw which consists of ~32 nonredundant sequences, which share between 55-90% identity with Tsw on the full-length protein and above 70% on the NBS domain span. Outside this group, the remaining homologs in the extended set share less than 40% identity with Tsw on the overall sequence and below 50% on the NBS domain span, generating a neat separation.

The close homologs are all part of the Solanaceae family, specifically of the *Capsicum*, *Solanum* and *Nicotiana* genera, while the more distant homologs show a broader taxonomic spread primarily across various taxa orders of the malvids and rosids clades (Supplementary File 1A). As expected, the close homologs of Tsw displays multiple highly conserved regions, some of which are also highly conserved in the extended homologs set (Supplementary File 1B, C).

Tsw 3D models

As specified under methods, AlphaFold automatic modelling fails for the long 2,116 aa functional version of Tsw. Additionally, the AlphaFold model does not account for
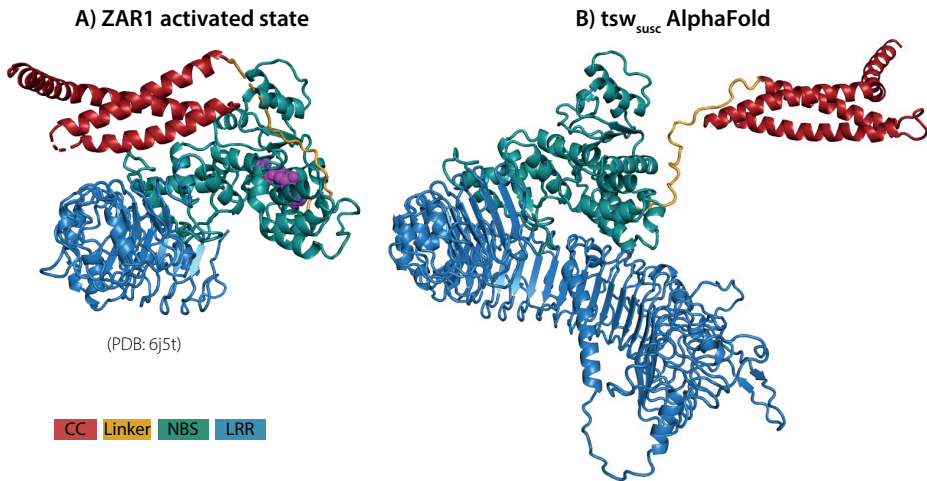
Figure 4.2 – Comparison of the overall model of the susceptible tsw from *C. annuum*. A) As proposed by AlphaFold in most likely its activated ATP state, with the ZAR1 cryo-EM structure **B)** In the same state (PDB: 6j5t). Configurations correspond to the superposition of NBS and the first 8 LRR repeats in both structures. As seen, the solution proposed by AF for the CC domain orientation is strikingly divergent to that experimentally determined by Cryo-EM in ZAR1. Domain color code is indicated in the figure legend.

the ADP resting state of the shorter susceptible 1,438 aa variant (Figure 4.2), nor does it unambiguously match with the motif delineation proposed by LRRpredictor. This prompted us to build additional system adapted models to overcome these limitations. Such adapted models furthermore allowed us to explore an opposed limiting scenario, which in conjunction with the AlphaFold model, defines the range in which we could expect such extensive LRR domain structures to stand. While the LRR curvature proposed by AlphaFold is wide – ~60 Å in diameter with a ~25° pitch of the supercoiled solenoid – the adapted models explore a wider range of diameter (40-70 Å), pitch angles (35-95°) of the supercoil bending and optimized repeat delineation. In a conformation with a wide radius, tight pitch and only 51 repeats, such as the one proposed by the AlphaFold model, the curvature translates into ~1.5 supercoiled turns. This is compared in Figure 4.3 with alternate Tsw-LRR models of 57 repeats in two extreme scenarios of relaxed vs. tight radius and pitch. These suggest that pitches narrower than ~25-35° potentially imply intra-domain interactions between consecutive turns of the LRR domain. It is therefore more probable to find the real structure standing and flexing in between these two extreme limits.

The Tsw-CC domain, aa 1-129, only displays remote homology to solved CC structures. Of these, ZAR1-CC is the closest with 13-17% identity depending on the alignment method. Secondary structure predictions indicate the presence of four helical segments in +/-1 turn agreement to ZAR1. Moreover, all four helices have an amphipathic pattern consistent with a 4-helical bundle fold with the hydrophobic regions hidden inside the bundle. As in ZAR1, the Tsw-CC N-terminal region of the first helical stretch, α1, is overall

hydrophobic, which might protrude the membrane in the active state similarly to that in ZAR1 (Figure 4.4). The next two helical segments α2 and α3 have an even higher similarity to ZAR1. Tsw also contains a partial EDVID motif (aa 81-85: "ADVAV") known to be involved in CC-LRR interaction. However, this is not a shared feature, as the third helix (α3) shows a significant variability even in the 32 close homologs of 50-90% identity (Figure 4.4). The absence of a conservation pressure within this region might indicate that the CC-LRR interaction pattern of Tsw family might be distinct from ZAR1.

Interestingly, the Tsw family displays highly conserved aromatic residues in positions equivalent to both the α1-α2 and the α3-α4 interconnecting loops in ZAR1. These two loops are close in distance in the 4-helical bundle structure, forming a contiguous surface exposed to the solvent. The location and degree of conservation of this patch might indicate a potential protein-protein interaction hotspot.
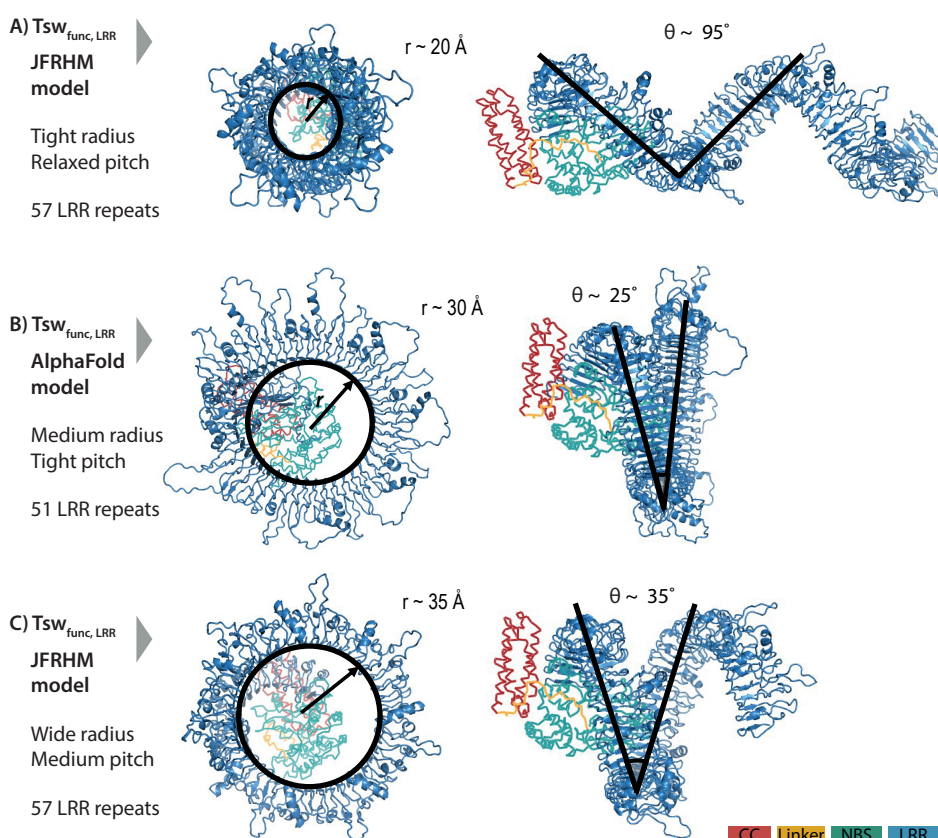
Figure 4.3 – Limiting scenarios of the LRR domain structure of the functional Tsw from *C. chinense* (1580 aa) with respect to the CC/NBS configuration in resting ADP state modeled starting from ZAR1. LRR model as **A)** proposed by AlphaFold automation; **B)** proposed by the adapted JFRHM LRR model in two limiting scenarios of tight/relaxed radius and pitch. Domain color code is indicated in the figure legend.

Lastly, α4 has the lowest homology with known templates, including ZAR1. In ZAR1 the predicted α4 is very long and was shown by cryo-EM to be unstable, as it folded differently with respect to α1, α2, and α3 in the resting and active state. While in the resting ADP state, only the first part of α4, named here α4a, closes the 4-helical bundle letting the second α4b part disordered. In the active ATP state this disordered region of α4b freezes, forming the four helical bundle and letting α4a disordered. By contrast, in Tsw the predicted α4 is significantly shorter and continues with a highly variable segment with linker-like propensities such as in the Rx1 template.

The Tsw-NBS domain (aa 158-495) shares 30% identity to ZAR1 and it preserves all the functionally significant regions: the P-loop, Walker-B, RNBS motifs A through D, the GLPL motif, and the MHD motifs (Figure 4.5). By contrast, the N-terminal entry into the NBS diverges from the common <u>VG</u> motif (Wróblewski et al., 2018), which in Tsw, both in close and extended homologs groups, is of type "eeF<u>d</u>SR" (where lower case show the consensus amino acids with lower conservation). Not all, but only some of the intra- and inter-molecular contacts seen in ZAR1 cryo-Ems, are conserved in Tsw. As seen for other R proteins, for instance in Sw-5a and -b (De Oliveira et al., 2016), point mutations
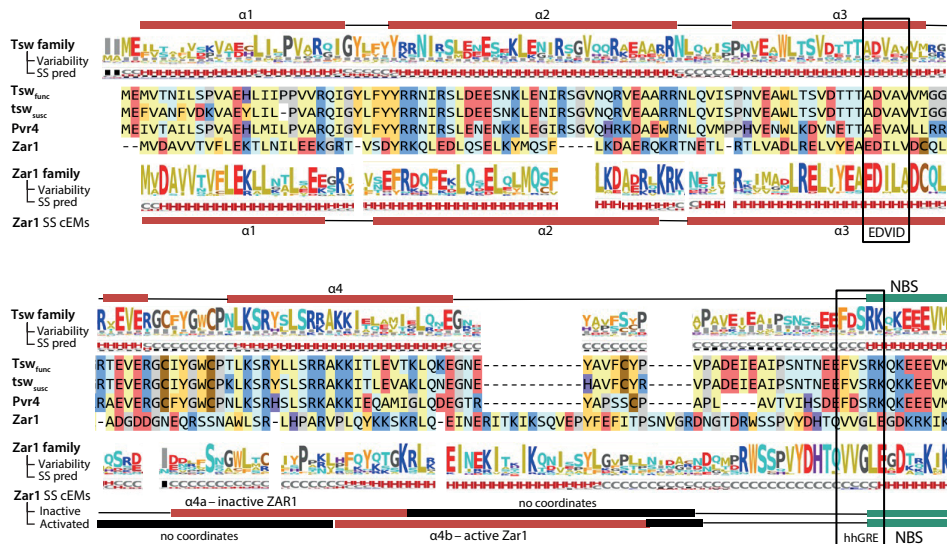
Figure 4.4 – Coiled Coil (CC) domain: sequence to structure mapping of *C. chinense* and *C. annuum* Tsw to the Arabidopsis thaliana ZAR1 template. The helical predicted structure is shown above the alignment. Variability and secondary structure (SS) prediction consensus lines are computed on the Tsw and ZAR1 homologous families at a 50% identity threshold using NLRscape (NLRscape, 2022). Variability is represented as relative entropy with respect to the background distribution - taller letters correspond to more conserved positions in the family. Secondary structure line reads: H – helix (red), E – extended (blue), C – coil (grey) with letter heights proportional to the percentage in the alignment. Helical structure assignments in cryo-EM inactive form (PDB: 6j5w) and oligomeric ZAR1 form (PDB: 6j5t) are shown below the alignment. Amino acids are colored according to their physical-chemical properties as follows: yellow - aliphatic hydrophobic; orange – aromatic; brown – cysteine; red - acidic negatively charged; blue - basic positively charged; light blue - polar neutral charge; purple - histidine; gray - glycine and proline.
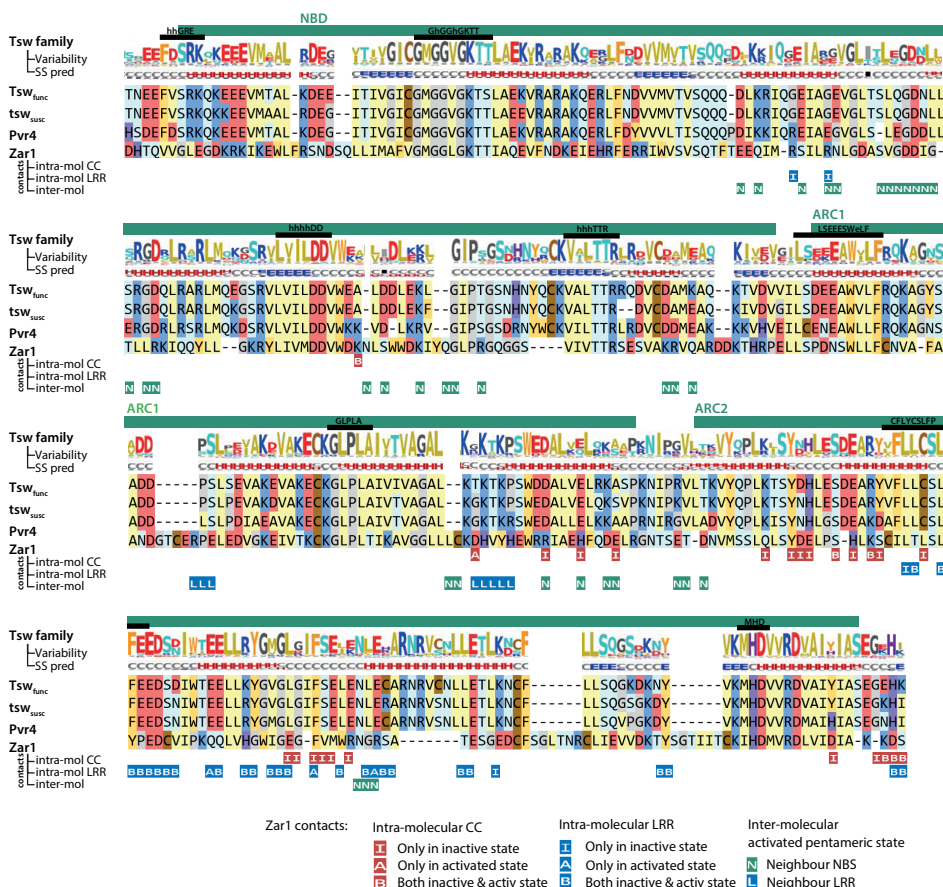
Figure 4.5 – Nucleotide binding domain (NBS): sequence to structure mapping of *C. chinense* and *C. annuum* Tsw to *A. thaliana* ZAR1 template. Sequence motifs and subdomains are annotated above the alignments. The variability and secondary structure (SS) prediction consensuses are computed and displayed as described in the legend included in the figure. Below the alignment are mapped the intra and inter-molecular contacts (under 5 Å) of ZAR1 cryo-EM structure in resting and activated conformations (6j5w, 6j5t).

between the functional and susceptible variants generate changes in salt bridge patterns at the NBS-LRR interface such as Tsw: D420-K540 changed to N417-K537 in tsw; or Tsw: C444-E660 changed to R441-E657 in tsw.

The LRR domain of the functional *C. chinense* Tsw variant spans over ~1,550 residues (aa 511-2,067). LRRpredictor delineates 57 LRR repeats, with several lower probability motifs discussed in detail below. Briefly, starting from the N-terminal side, the LRR domain is structured in a core section of 15 LRR repeats (aa 511-901), followed by seven consecutive quasi-identical blocks each consisting of six LRR repeats named A-F. The last repeat, 7-F, is incomplete and followed by a highly acidic tail with high propensity for an intrinsically disordered structure (aa 2,068-2,116) (Figure 4.6).
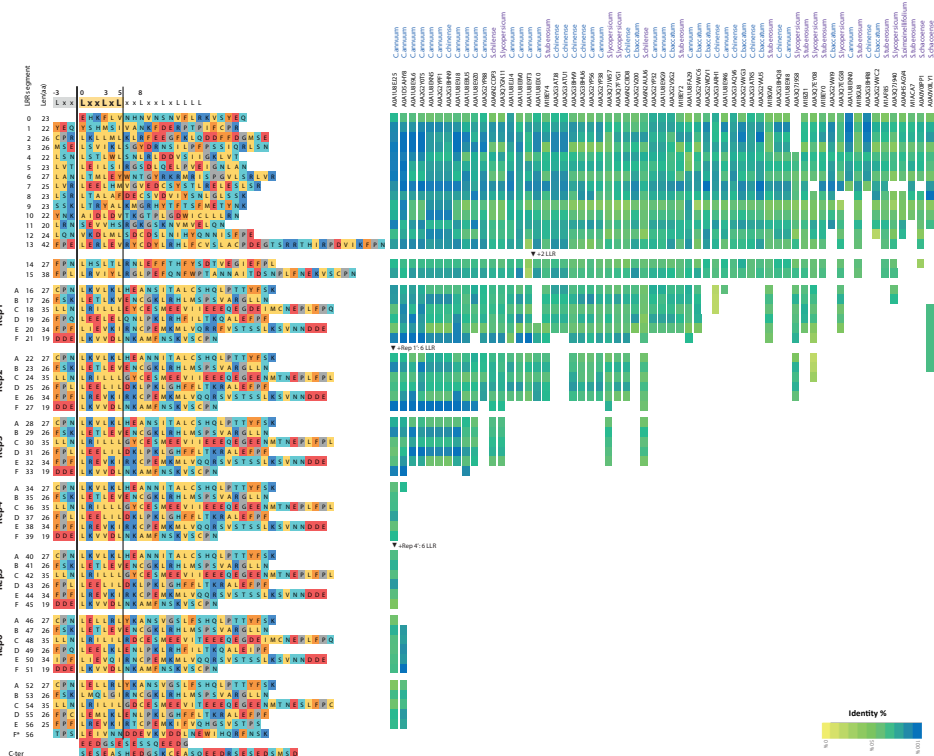
Figure 4.6 – LRR repeat delineation of Tsw from *C. chinense*. The right hand side of the figure shows Tsw's closest homologs from *Capsicum* (blue) and *Solanaceae* (purple) families alongside with the conservation variations across each LRR repeat.

In Tsw, the NB-ARC-LRR linker corresponding to the region between the end of the ARC2 subdomain (aa 495) and the first proper LRR motif (aa 512 — YSHMSI) is nearly double in size to its equivalent region in ZAR1: 17 aa vs. 9 aa. In the cryo-EM structure of ZAR1, this 9 aa linker folds as an improper LRR repeat that is loosely attached to the rest of the LRR domain. In Tsw, its 17 aa equivalent has notable similarities with proper LRR repeats, both in length and secondary structure propensity. It was therefore modeled as such and referred herein as the LRR 0 repeat (Figure 4.6).

In Tsw, the first seven LRR repeats, with which interactions with CC and NBS are expected to occur, show significant similarities with ZAR1. LRR motif predictions become problematic in the region aa 740-780. The motif probability (aa 761 — SEVVHS) is not very high here, thus this could be modeled either as a single long 44 aa repeat, or as two shorter repeats. Such a motif has a very low LRRpredictor probability. However, the propensity of this stretch for extended secondary structure and forming beta sheets is high. Hence, this was locked as a proper motif in both AlphaFold and adapted-JFRHM models.

The AlphaFold oddly models the orientation of LRR 9 in reverse, i.e. with its dorsal side oriented towards the inside of the horseshoe. None of the structurally solved LRR domains contains such an odd repeat orientation. Moreover, this orientation is inconsistent with overall NLR architecture in the resting ADP state, as such a reverse repeat would clash with the NBS domain. Therefore, most likely the local solution adopted by AlphaFold for LRR 9 is an artefact derived from fusing two templates and was corrected in the adapted-JFRHM model.

In each of the following seven blocks, LRRpredictor identifies six repeats termed herein A-F. This in contrast to AlphaFold, which merges repeats E and F in a single 5 aa strand. In our A-F delineation, all motifs on the ventral side of the horseshoe are of one of the following extended types: L-x-x-(L-x-L-x-x-L)-x-x-L-x-x/L-L-L-x/L, suggesting a structural reinforcement that was taken into consideration in the adapted models. Interestingly, all the A-F repeats display in addition a stretch of two to three consecutive hydrophobic residues on the dorsal side of the horseshoe in the 14-18 aa region of each repeat. This could potentially form a continuous hydrophobic 'band' over the entire structure.

The susceptible variant tsw from *C. annuum* lacks blocks 3-6, while the core section of block 1 and 7 follows the same LRR organization as the functional variant with only minor sequence variations (94.0% identity at nucleotide level). Apart from the missing LRR blocks, the last 23 aa at the C-terminal end – within the second half of the acidic tail – of tsw has a completely different, highly positively charged, aa sequence due to a frameshift resulting from a 1 nucleotide deletion. This makes a striking distinction between the functional vs susceptible variant, *i.e.* the C-terminal tail of Tsw being entirely acidic, while the tail of tsw is half acidic, half basic. Such differences could be responsible for the loss of its functionality as a resistance gene.

Interestingly, when inspecting close homologs of Tsw within the *Solanaceae* clade, the lengths of the LRR domains varies significantly, with most homologs not exceeding two or three repetitive blocks (Figure 4.6). Common to all, the core region of the LRR domain shows a higher degree of conservation within the N-terminal region (LRR 1-8), which is consistent with the region's expected involvement in NBS interaction.

Furthermore, the central C repeat of each repetitive block is highly negatively charged. This creates a noticeable repetitive pattern in 3D, and in addition makes each module slightly negatively charged. Overall, this adds up to an extreme -32 total net charge on the LRR domain span. When also considering the net charges of the CC (+6), CC-NBS linker (-7), NBS (-7), and the acidic C-terminal tail (-17), the total net charge of the overall protein becomes -57, which has to be compensated by the environment, or the folding of the protein. Tsw's total net charge is quite extreme in comparison to other R proteins, for which it generally ranges around -10 to +10. When compensated for by length, the total net charge remains below average.

While both the AlphaFold and the JFRHM adapted models display the LRR domain as a continuous solenoid supercoil, it cannot be excluded that the continuity of the LRR domain might be disrupted or altered at the less reinforced regions of the LRR pattern due to the molecular environment or on the large biological time scale.

Inter-domain interactions

AlphaFold only models Tsw starting from the ZAR1 ATP active state. Moreover, the five best AlphaFold models display the CC domain in a completely different configuration and opposite orientation to that seen in ZAR1-ATP active state (Figure 4.2). This local modeling solution adopted by AlphaFold results in a complete disruption of the interface between CC-NBS and CC-LRR as that seen in ZAR-ATP state (Figure 4.2), suggesting that the amino acids at the predicted interface are not compatible with a CC-NBS and CC-LRR contact state.

To further investigate this incompatibility, we analyzed the corresponding CC-NBS and CC-LRR interface on the JFRHM models generated starting from the inactive ADP conformation of ZAR1. As can be seen in Figure 7, indeed the amino acid properties at the CC-LRR interface do not match perfectly with those seen in the ZAR1 template, suggesting that the interaction between these two domains might be different in Tsw. In ZAR1, the CC-LRR interface is shared between the inactive and active conformations and has a strong electrostatic nature (Figure 4.7A, B) with the CC acidic EDVID motif area in α3
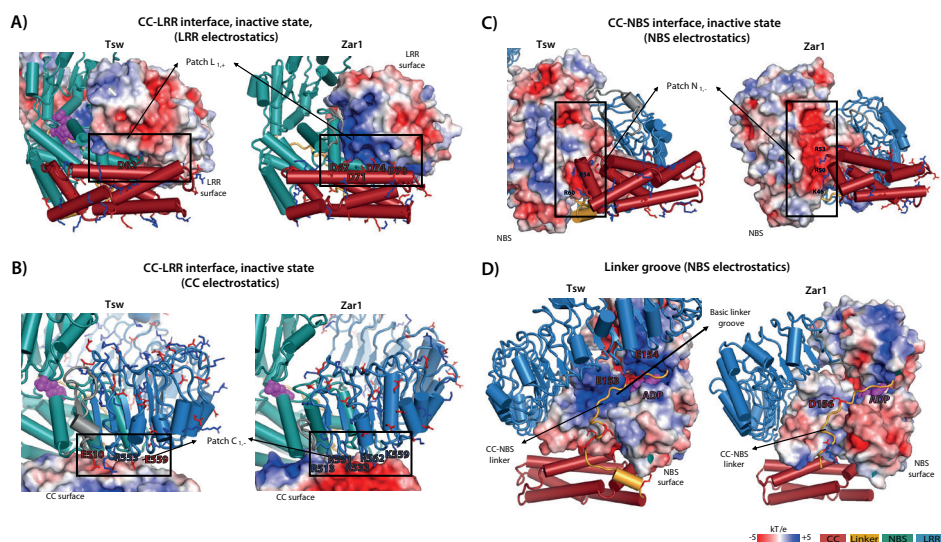
Figure 4.7 – Electrostatic properties of the interdomain contact surfaces mapped on the functional Tsw adapted JFRHM model and ZAR1 inactive ADP-binding cryo-EM structure (PDB: 6j5w). CC-LRR interface electrostatic potential mapped on **A)** the LRR and **B)** CC domains. Electrostatic potential surface of NBS domain at the contact interface with **C)** the CC domain and **D)** the CC-NBS linker. Negatively (red) and positively (blue) charged residues located at inter-domain interfaces (dotted line boxes) are labeled and displayed in stick representation. Electrostatic potential scale and domain color code are indicated in the figure legend.

being complemented by a proportionally basic patch on the lower side of the first four LRR repeats. On the other hand, in Tsw the interface is neutral, suggesting a far weaker affinity.

By contrast, the CC-NBS interface shares a very similar electrostatic profile with that seen in ZAR1 in its resting ADP state. The NBS part of the interface consists of a negatively charged patch $N_{1,-}$ (ARC1 end, ARC2 beginning) in contact with the positively charged region at the end of α3 and the outset of the α4 helix of the CC domain (Figure 4.7C).

Focusing now on the groove of formed across the NBS domain accommodating the CC-NBS linker in the JFRHM model of Tsw built starting from the ZAR1 inactive ADP template, the compatibility between the linker sequence and the NBS groove is even higher in Tsw than in ZAR1 due to a striking opposite charge matching (Figure 4.7D). In Tsw, the NBS groove is highly positively charged and the linker displays an increased net negative charge of -5 (aa 140-156), whereas the interface in ZAR1 has a more homogenous neutral charge. Together, these findings suggest that Tsw shares only partially interdomain interface profiles with ZAR1, implying that the interaction between domains might be different from the ones seen in ZAR1 during activation.

Conclusion and reflections

The activation of the 10% of NLRs with ≥1,000 aa-long LRRs have scarcely been elucidated. The modelled structures of the resistant and susceptible Tsw proteins provide a glimpse into the intricate resistance pathways in plants, although the exact function of the large LRRs remains a conundrum. Cryo-EM studies have only quite recently revealed multimeric resistosome formation from activated monomeric NLR proteins (Wang et al., 2019a; Ma et al., 2020; Martin et al., 2020b). It is unknown whether the formation of resistosomes is a general feature of NLRs, or an anomaly amongst NLRs. Their resemblance to apoptosomes and inflammosomes, however, provides support for these oligomeric structures as a generic feature of NLS. For quite a number of NLR proteins researchers have demonstrated self-association and the formation of dimers and oligomers (Bernoux et al., 2011; Maekawa et al., 2011; Casey et al., 2016; Zhao et al., 2021). From such configurations it is only a small step towards an activated pentameric resistosome. The question therefore arises whether these dimers and oligomers are involved in activities that are needed prior to further formation into a resistosome. All three currently confirmed resistosomes are formed by NLRs with an average LRR length, thus it will be interesting to see whether NLRs with long LRRs like Tsw are also capable of such formations. And if they are, whether they self-associate, or whether Tsw for instance requires assistance from helper NLRs for oligomerization and activation. A 3D protein model of Tsw has been proposed based on all currently available modelling tools, although these are not always suited for such large proteins. Without cryo-EM structures of Tsw, no definitive conclusions can be drawn. As it is likely that cryo-EM structures will be resolved for NLRs with long LRRs such as Tsw,

this subject might be resolved in the future. However, preliminary conclusions can still be drawn based on our proposed 3D model.

Upon seeing the highly unusual helical turn made by the LRR in functional Tsw, questions arise regarding the interaction of the LRR with other proteins. Are NLRs with long LRRs capable of directly interacting with effectors, or indirectly with a guardee/decoy protein? Is there a common denominator between these unusual R proteins, potentially indicating a shared protein interaction? Furthermore, one wonders whether the long LRR is crucial for the interaction, or whether as with typical NLRs the C-terminal end is the determinant for the interaction. The helical turn creates a large cavity in the LRR domain, which potentially may be indictive of the size and shape of its interactor, as is the case for CNL protein CYR1 from mungbean (normal length LRR), and its effector, the coat protein of *Mungbean yellow mosaic India virus* (MYMIV). *In silico* models revealed a niche formed by CYR1 in which MYMIV-CP fitted neatly upon their interaction (Maiti et al., 2012). Modeling attempts have been made for NSs (Olaya et al., 2019). However, when comparing Tsw and NSs, it is important to bear in mind that there is little to no resemblance in sequence between proteins with 3D structures confirmed by cryo-EM and TSWV NSs. The increasingly accurate protein structure prediction programs are steadily modeling many potential interaction partners and could in the future be used to support other findings.

The 3D model of Tsw allows for structure-informed decisions regarding future functionality experiments. The similarity between α1-α2 and α3-α4 in ZAR1 and Tsw might indicate a possible shared protein interaction spot, while the low conservation of α3 in the Tsw family could denote a potentially divergent interaction between CC and LRR in this family. Mutations in this region might therefore shed light on this possibility, while truncations in the LRR domain could aid in uncovering the function of the long LRR domain, and the need for the extra helical turn, as well as determining the necessity of the C-terminal tail in the recognition of NSs. Similar experiments have been performed for Pvr4, a homolog of Tsw from *C. annuum* that confers resistance against potyviruses (Kim et al., 2017a), where removal of three out of six LRR repeats triggered auto-immunity (Kim et al., 2018). Chimeras of the non-functional (susceptible) tsw and functional Tsw in which LRR domains are swapped/added/deleted would likewise allow for highly interesting functional studies. Similarly, a C-terminal end swap between Tsw and Pvr4 would determine whether this segment is responsible for pathogen recognition, much alike Gpa2 and Rx1 (Slootweg et al., 2017). Furthermore, as downstream signaling upon Pvr4 activation may function due to its $CC_{NO-EDVID}$ domain, a CC swap or mutation of the EDVID-like domain in Tsw would allow for potentially interesting observations. Incorporation of the hypotheses mentioned here in future research would potentially enhance our understanding of the roles of long LRRs, as well as other domains of Tsw, in both pathogen recognition and downstream immune pathways.

Supplemental data

A) **Taxanomic spread**

Extended Tsw homologs
> 30% identity

Close Tsw homologs
> 50% identity

200 nonredundant members
(90% identity)
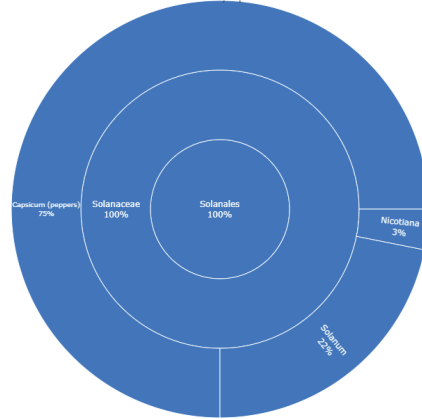
32 nonredundant members
(90% identity)

B)

Sequence variability
Close Tsw homologs (>50% identity)
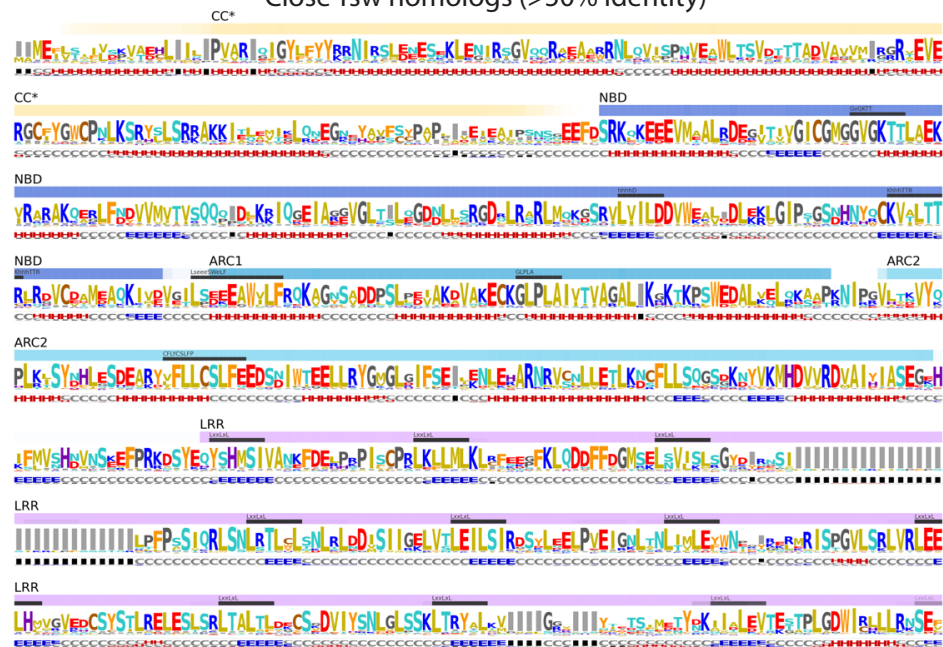
Distant Tsw homologs (>30% identity)
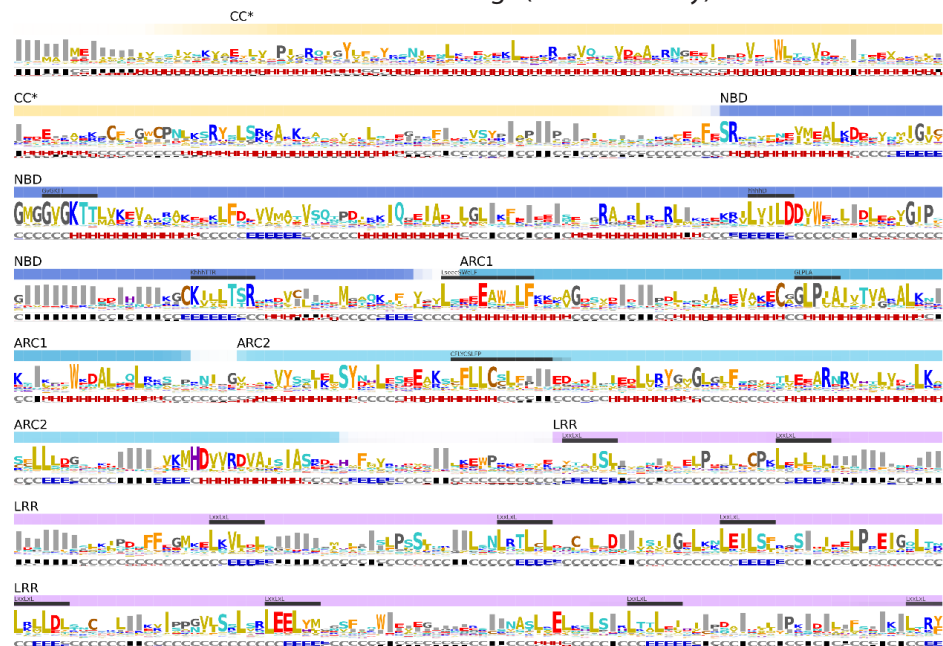
*Only the first 8 LRR repeats are displayed due to increased gap percentage over the rest of the LRR domain alignment.

C)

LRR motif variability

Close Tsw homologs
(>50% identity)

Distant Tsw homologs
(>30% identity)
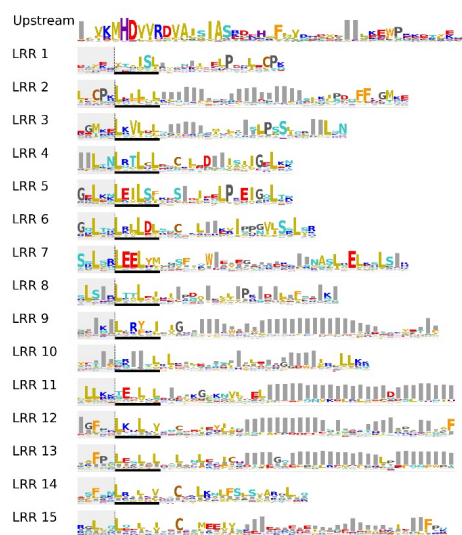
*Only the first 15 LRR repeats are displayed.

D)   Identity matrix - extended set of homologs
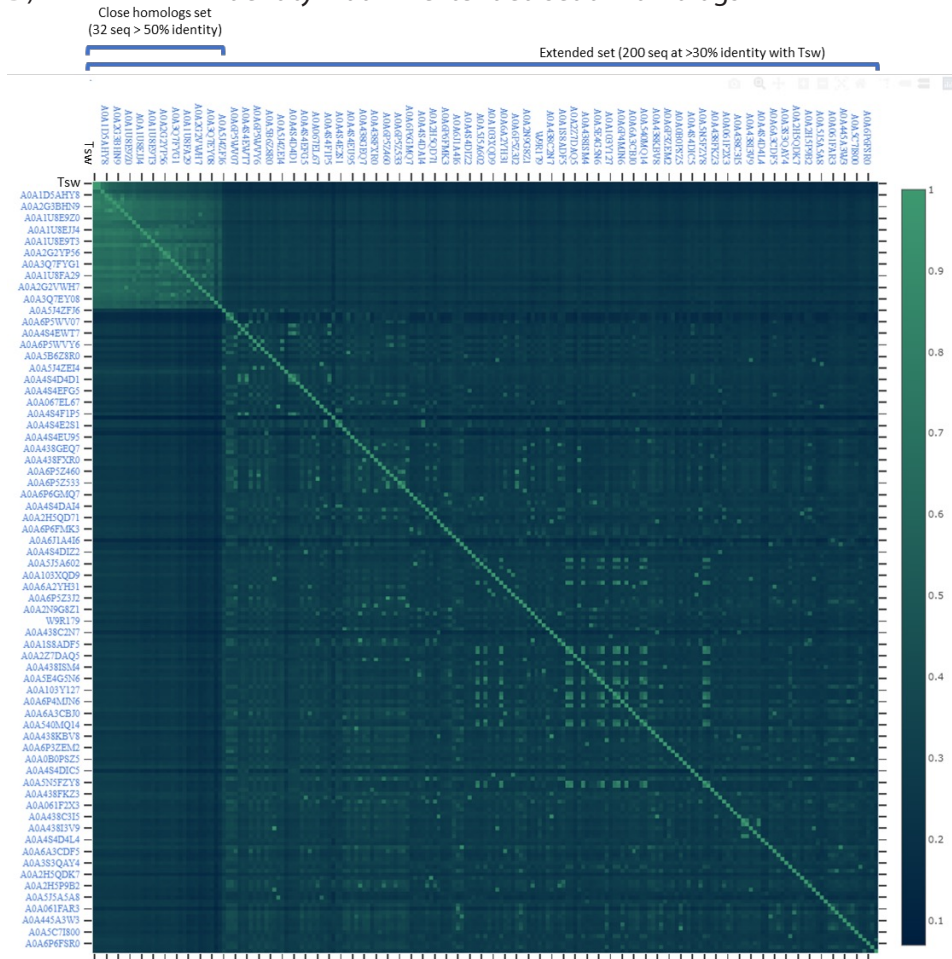
* Not all labels are visible (indicated by tickmarks) due to the large size of the matrix.

** Identity is computed as the ratio between the number of identical residues between Sequence 1 and Sequnce 2 over the length of Seq. 1 (lower diagonal side) or of Seq. 2 (upper diagonal side). Therefore the matrix is not symmetrical.

Figure S4.1 – Tsw variability analysis on two sets of homologs. An extended homologs group of 200 sequences sharing between 30-90% identity with Tsw and subset of it corresponding to 32 close homologs subset which share 50-90% identity with Tsw: **A)** Taxonomy distribution, **B)** Variability plots, **C)** LRR motif variability, and **D)** Identity matrix.

Figure S4.2 – The Tsw LRR repeat delineation used by AlphaFold and the adapted JFRHM models.
Differences are depicted by black boxes.

Chapter 5

In silico informed analyses of *Tsw* and related resistance genes

I.L. van Grinsven, I. Goet, M.M. van Oers, R. Kormelink

Abstract

The plant immune system is a multilayered system, from general defenses to the action of highly specific proteins providing resistance to single pathogens. The regulation of such resistance (R) proteins is also multilayered, in which transcriptional expression and protein functionalization is tightly controlled to prevent autoimmunity. The *R* gene product Tsw confers resistance against Tomato spotted wilt virus (TSWV) upon (in)direct detection of the viral RNA silencing suppressor NSs. Here, *in silico* analyses were performed to uncover potential regulatory elements in the full-length sequence of the *Tsw gene*, as well as the susceptible variant *tsw*, and the homolog *Pvr4*. Micro RNAs of the miR482 and miR6026 family were found to influence *Tsw* expression. Furthermore, a method was devised to enable easy and rapid domain exchange and mutagenesis in Tsw for functional gene studies in plants.

Introduction

Peppers (*Capsicum* sp.), similar to other members of the *Solanaceae* family, originate from South America but have become a commonly grown and used high-value vegetable crop all over the world. However, yields are jeopardized by harmful plant diseases, including Tomato spotted wilt virus (TSWV). TSWV is considered one of the most economically devastating plant viruses worldwide (Scholthof et al., 2011) and is mainly spread by western flower thrips (*Frankliniella occidentalis*). Attempts to combat thrips with insecticides have proven difficult, therefore methods to limit the spread of TSWV heavily lean on breeding for virus resistance traits (Gilbertson et al., 2015; Reitz et al., 2020).

Plants have a multilayered immune system that provides defense against various pathogens (de Ronde et al., 2014; Chapter 2/Zhu et al., 2019). A broad non-host resistance heavily relies on the recognition of conserved pathogen molecules, called Pathogen Associated Molecular Patterns (PAMPs). Downstream signaling pathways are triggered upon recognition of the PAMPs by transmembrane Pattern Recognition Receptors (PRRs), leading to PAMP-Triggered Immunity (PTI). While PRRs recognize widely conserved motifs, resistance conferred by resistance (R) proteins depends on their pathogen-specific recognition. The subsequent effect of extreme resistance (ER) or hypersensitive response (HR) allows the plant to combat and contain the intruder, preventing their (systemic) spread. While quick action against infection is necessary, constitutive high expression of *R* genes will lead to loss of fitness for the plant (Burdon and Thrall, 2003; Huot et al., 2014). Regulation of *R* gene transcription by microRNAs (miRNAs) is one method to allow for a fast response, without the costs of hypervigilance. Derived from host gene encoded hairpin RNA structures, single-stranded 21-22 nt miRNAs loaded into RISC complexes guide the targeting of complementary messenger RNA (mRNA) for degradation. This effectively halts translation, while still allowing for high levels of gene transcription. Effectors of invading viruses often suppress silencing, which also lifts the suppression of *R* gene translation. Various miRNAs and miRNA families are known to specifically target *R* gene transcripts, for instance the canonical miR482/2118 family that targets the sequence

encoding the conserved P-loop motif in the NB-ARC domains (Shivaprasad et al., 2012; de Vries et al., 2018).

The *R* gene *Tsw* originates from *Capsicum chinense* and confers resistance against TSWV. The effector recognized by *Tsw* is NSs, the non-structural protein from the S segment of TSWV that acts as an RNA silencing suppressor (RSS) (de Ronde et al., 2013). *Tsw* is a coiled coil (CC) nucleotide binding (NB) leucine-rich repeat (LRR) (CNL, NLR) receptor that has been cloned by Kim and colleagues, and whose amino acid (aa) sequence has been published in 2017 (Kim et al., 2017a). The central domain of NLRs acts as an ON/OFF switch, and is activated upon binding ATP rather than ADP (van Ooijen et al., 2008; Takken and Tameling, 2009; Maekawa et al., 2011; Williams et al., 2011). In canonical NLRs, both the CC domain and the LRR interact with the central domain and thereby suppress the ADP/ATP swap, keeping the NLR in an auto-inhibited state. (In)direct detection of the pathogen/effector will lift the inhibition, activate the *R* gene, which in turn will trigger downstream defense pathways. The study from Kim et al. (2017) revealed that *Tsw* resides in a gene cluster together with 13 highly similar (pseudo)genes in an area of 295 kbp, and furthermore discussed the evolution of this gene and its homolog *Pvr4* residing in *C. annuum*. Transient expression of a putative *Tsw* DNA copy from the transcriptionally expressed gene, in the additional presence of the viral NSs effector, confirmed its identity as the *Tsw* resistance gene (Kim et al., 2017a). However, the suppression and the activation mechanism of *Tsw* and the downstream HR signaling have not been elucidated. Here, the published amino acid sequence from Kim *et al.* (2017) and the encoding nucleotide sequence of *Tsw* obtained in Chapter 3 were utilized for further bioinformatical studies and the gene was tested for possible miRNA-mediated regulation. To enable future functional gene studies and the construction of chimeras of Tsw and the susceptible tsw, containing domain swaps between the two alleles, a synthetic gene approach is designed.

Materials and Methods

Virus, plants, and constructs

The two commercially available *Capsicum annuum* cultivars "Frazier" and "Mildred (Enza Zaden) were used in this study, respectively a TSWV susceptible and a resistant cultivar. The resistance inducing TSWV isolate Vir129 (Wageningen, NL (de Ronde et al., 2013)) was maintained on *Nicotiana benthamiana* by (maximally five) serial mechanical passages and stored as frozen leaves at -80 °C. Plants were grown and maintained under standard greenhouse conditions (22 °C; 16 h light/8 h dark cycle; 70% relative humidity).

Genome search

The amino acid and coding sequences of *Tsw* (A0A1C9TCM9_CAPCH, KT751527.1) published by Kim et al. (2017) were used to search the *Capsicum chinense* PI159236 scaffolds and assembled genome (Chinense.v.1.2 from the Pepper Genome Platform) for the full-length sequence of *Tsw*. The entire sequence of *Tsw* was also recovered

5

from a contig from a BAC library of *C. chinense* PI152225 (Chapter 3). The genome of the susceptible *C. annuum* cv. CM334 version 1.55 was searched for the full-length non-functional variant of *Tsw* using the amino acid sequence of *tsw* (Kim et al). All genome searches and subsequent analyses have been performed on Geneious Prime.

PCR

TRIzol extractions were performed to extract genomic DNA (gDNA) and RNA from leaf material of two genotypes of *C. annuum* (TSWV susceptible HK0004, and TSWV resistant YF0009), as well as from *C. chinense* PI 152225. First strand cDNA was synthesized using either using random hexamers (Roche) or sequence-specific primers (Table S5.2) and M-MLV reverse transcriptase according to the manufacturers' protocol (Promega). PCRs were performed with GoTaq G2 Flexi, Q5, Phire HotStart II, and Phusion High Fidelity polymerase according to manufactures specifications (Promega, ThermoFisher Scientific, New England Biolabs, New England Biolabs) with this genetic material in attempt to clone full-length and CDS versions of both *Tsw* and *tsw* PCR products were subsequently run on 1% agarose electrophoresis gel to determine presence or absence of fragments.

Construction of synthetic *Tsw* variants through Golden Gate assembly

Segments of *Tsw* and *tsw* sequences were ordered as gBlocks (IDT) flanked by *Bsa*I restriction sites, to enable construction of full-length *R* genes as well as truncations and chimeras through Golden Gate assembly (New England Biolabs). To allow for the first step of Golden Gate assembly, intermediate, or level 0, vectors harboring a *ccdB+* gene as well as appropriate *Bsa*I overhangs were made. A *ccdB+* gene was PCR amplified with Q5 polymerase and desired restriction sites were added. Fragments were run on an agarose gel, correctly sized fragments were excised and gel purified. Fragments and vector pGGG002 (Lampropoulos et al., 2013) were restriction enzyme digested and subsequently ligated. The internal *Bsa*I site in *ccdB+* was removed through Q5 site-directed mutagenesis (NEB), and the fragment was transformed to electrocompetent *Escherichia coli* strain DB3.1. The expression vector pGreen was adapted in a similar fashion using oligos with outward-facing *Bsa*I sites, thus allowing for scarless annealing in the destination vector.

gBlocks (IDT) were resuspended in TE buffer upon delivery and stored at -20 °C. Golden Gate assembly was performed by mixing 20 fmol of each fragment (resulting from either gBlocks or previously Golden Gate assembled level 0 vectors, to be inserted), 10 fmol of vector DNA, 10x T4 DNA ligase buffer (Promega), 3 Weiss Units T4 ligase (Promega) and 20 Weiss Units *Bsa*I-HFv2 (NEB), and subjecting the mixture to an assembly protocol (30 x (3 min 37 °C – 5 min 16 C) – 5 min 37 °C – 10 min 60 °C) followed by transformation to electrocompetent *E. coli* DH10β. Colony PCRs were used to screen for correctly assembled constructs and sequence verified prior to storage at -80 °C of a glycerol stock of fresh liquid culture.

In silico analysis of *Tsw* target prediction

The regions 2000 nt upstream of the translation start site of *Tsw*, *tsw*, *Pvr4*, *Sl5R-1*, *Sw-5b*, and *Rx1* were used for *in silico* analysis. The CpG island 1.1 identifier plugin in Geneious was used to search for CpG islands (Tobias Thierer and Biomatters Ltd). A search for RNA polymerase II promoters, or transcription start sites (TSS), was conducted using TSSplant (accessed April 8, 2022). The MEME motifs of the 279 transcription factors in *Capsicum annuum* (PlantTFDB (Tian et al., 2019), accessed April 8, 2022) were used to determine potential transcription factors with FIMO version 5.4.1 (MEME Suite, accessed April 8, 2022). PlantCARE (Lescot et al., 2002) was used for the prediction of cis-acting regulatory elements (CARE).

By combining several published *Capsicum* miRNA databases (Zhang et al., 2010; Kim et al., 2014b; Szcześniak and Makałowska, 2014; Seo et al., 2018; Taller et al., 2018), a group of 612 distinct miRNA sequences was constructed and used to find miRNA targets in the *Tsw* sequence using the psRNATarget tool (Dai et al., 2018). Potentially interesting miRNAs were selected based on expectation value (≤5.0), the miRNA target region in *Tsw*, and the reported regulatory function of the miRNA in literature.

Short tandem target mimic (STTM) vector construction

Based on a design of published STTM constructs (Sha et al., 2014), two miRNA silencing STTM constructs flanked by AttB1/2 sites were designed and constructed as gBlocksTM (IDT) (Figure S5.3). These were subsequently Gateway cloned into pDONR207 and transformed into heat-shock competent Escherichia coli DH10β. Colony PCR was performed to determine correctly sized inserts, and overnight liquid cultures were DNA purified with the GeneJET Plasmid Miniprep Kit (Thermo Scientific). Colonies containing the correct sequence were Gateway cloned into destination vector pTRV2 (Liu et al., 2002), electroporated into electro-competent *Agrobacterium tumefaciens* strain GV3101. pTRV1 and pTRV2-PDS silencing constructs were constructed previously (Senthil-Kumar and Mysore, 2014).

Virus induced gene silencing and viral infection

Proteins were transiently expressed via an Agrobacterium Transient Transformation Assay (ATTA). The *A. tumefaciens* GV3101 cells containing the desired constructs were grown for 18-36 hours at 27 °C in LB3 medium containing appropriate antibiotic selection pressure as well as 20 µg/ml rifampicin. The optical density at 600 nm (OD_{600}) of the cultures was determined, the cells were pelleted and resuspended in MMAi medium (Murashige-Skoog induction medium; 10 mM MES; 0.2 µM acetosyringone) to obtain the desired OD_{600} of 0.5. Equal volumes of pTRV1 and one pTRV2 culture were mixed and set in the dark for an hour. Pepper plants of 2-3 weeks old were brush inoculated on both the abaxial and adaxial side of the cotyledons, and across the hypocotyl with Agrobacteria mixtures to which 0.02% Silwett-77 had been added just prior to inoculation. Fourteen

5

days post agro-infiltration (dpa) with pTRV constructs, plants were challenged with TSWV by mechanical inoculation of TSWV-infected plant extracts (Bucher et al., 2003).

Determination of viral titers

Total RNA of *C. annuum* leaf material 14 dpa and 10 days post infection (dpi) was extracted with TRIzol reagent (Invitrogen) and Dnase treated with the DNA-*free*™ Kit (Ambion). First strand cDNA was synthesized from 2 ng of total RNA with oligo dT primers and M-MLV reverse transcriptase in a 20 μl reaction (Promega). The level of target gene mRNA was determined by performing an RT-qPCR with SYBR™ Select Master Mix (Applied Biosystems) on 10-fold diluted cDNA and gene-specific primers (Table S5.2) on a Bio-Rad CFX96. The efficiency of all primer pairs was determined, and melt curves were analyzed. The expression levels of ubiquitin-conjugating protein (UBI3) was used to normalize expression levels of target genes using the Pfaffl method (Vandesompele et al., 2002; Hellemans et al., 2007), and calibrated to the level of control plants (pTRV2-GUS infiltrated or mock-infected plants, 14 dpi and 14 dpa respectively). This pilot experiment was performed once with six plants and three technical replicates per plant. Data were analyzed with Prism 9.3.1.

Table 5.1 – Identity comparison of *Tsw* exons encoding for the leucine-rich repeat domain. Percentages of shared identity for nucleotides (left bottom, dark gray) and amino acid (right top, light gray) sequences. Exon 1 and exon 9 respectively encode the NB-ARC domain and the C-terminal tail, and are therefore left out of this table.

nt/aa	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Exon 8
Exon 2		89.8%	90.4%	89.8%	89.8%	88.0%	71.8%
Exon 3	94.4%		99.4%	100.0%	100.0%	83.2%	75.9%
Exon 4	94.6%	99.8%		99.4%	99.4%	83.8%	76.5%
Exon 5	94.4%	100.0%	99.8%		100.0%	83.2%	75.9%
Exon 6	94.4%	100.0%	99.8%	100.0%		83.2%	75.9%
Exon 7	94.0%	91.6%	91.8%	91.6%	91.6%		79.4%
Exon 8	84.0%	86.5%	86.7%	86.5%	86.5%	87.3%	

Table 5.2 – Percentage of shared identity of the *Tsw* introns flanking coding sequences of the leucine-rich repeat domain.

	Intron 1	Intron 2	Intron 3	Intron 4	Intron 5	Intron 6	Intron 7
Intron 1							
Intron 2	47.3%						
Intron 3	47.3%	99.8%					
Intron 4	47.2%	99.9%	99.8%				
Intron 5	47.3%	99.5%	99.6%	99.5%			
Intron 6	47.4%	98.9%	99.0%	99.0%	99.2%		
Intron 7	47.9%	97.7%	97.8%	97.7%	97.6%	97.5%	

Results

Comparison of resistance genes

Only the amino acid sequence of *Tsw* was included in the data when its discovery was published (Kim et al., 2017a). To uncover potentially crucial non-coding sequences and to facilitate comparison of nucleotide sequences, the aa sequence was back translated and used to search the *C. chinense* contigs indicated to contain the *Tsw* gene for the full sequence. The full sequence of the homologous *Pvr4* gene and *tsw* variant, leading to TSWV susceptibility instead of resistance, were uncovered in a similar fashion from the publicly available genome sequence of *C. annuum* 'CM344'.

The full-length *Tsw* gene is 18.6 kbp in size, with 9 exons and 8 introns (Figure 5.1A). The NB-ARC domain is located on exon 1, while the repeats of the LRR domain roughly span exon 2-8, with a C-terminal tail encoded by exon 9. As expected with LRR domain sequences, these exons contained a high number of (near-)identical repeats of approximately 501 bp, or 167 aa. These similarities were especially noticeable at nucleotide level (Table 5.1). With the exception of exon 8, the percentage of identical bases between exons was over 90%. Comparison of the introns flanking exons containing the LRR domain (intron 2-7) revealed near identical (97%) intron sequences (Table 5.2). Corresponding exons and introns spanning the LRR of susceptible *tsw* are somewhat less identical, with exons 2-4 at >80% and intron 2-3 at 88.3% identical (Table S5.1).

While the *Tsw* gene from *C. chinense* confers resistance to *Tsw*, its counterpart in *C. annuum* is not functional. The largest difference between the amino acid sequences of both genes is the size of the LRR domain. Whereas Tsw contains seven LRR repeats, tsw only has three (Chapter 4). Besides this difference, their C-terminal ends are also both highly different. Comparison of nucleotide sequences, rather than amino acid sequences, showed that their highly divergent C-terminal tails are caused by a one nucleotide deletion in exon 5 in tsw resulting in a frameshift upon translation of the gene (Figure 5.1B). Nucleotide sequences of *tsw* overlapping with those of *Tsw* match for 93.0%, their overlapping amino acid sequences have a shared identity of 83.6%.

The homologous protein Pvr4, with five LRR repeats, also has a shorter LRR domain than Tsw, as well as a different C-terminal tail. Comparison of nucleotide sequences showed that the C-terminal tail of *Pvr4* contains a 1 nucleotide insertion, again resulting in a frame shift, which in this case results in an earlier stop codon in comparison to *Tsw* (Figure 5.1B). The nucleotide sequence of *Pvr4* matches for 88.9% with that from *Tsw* gene, gaps excluded. Their shared amino acid sequences match for 79.3%.

5
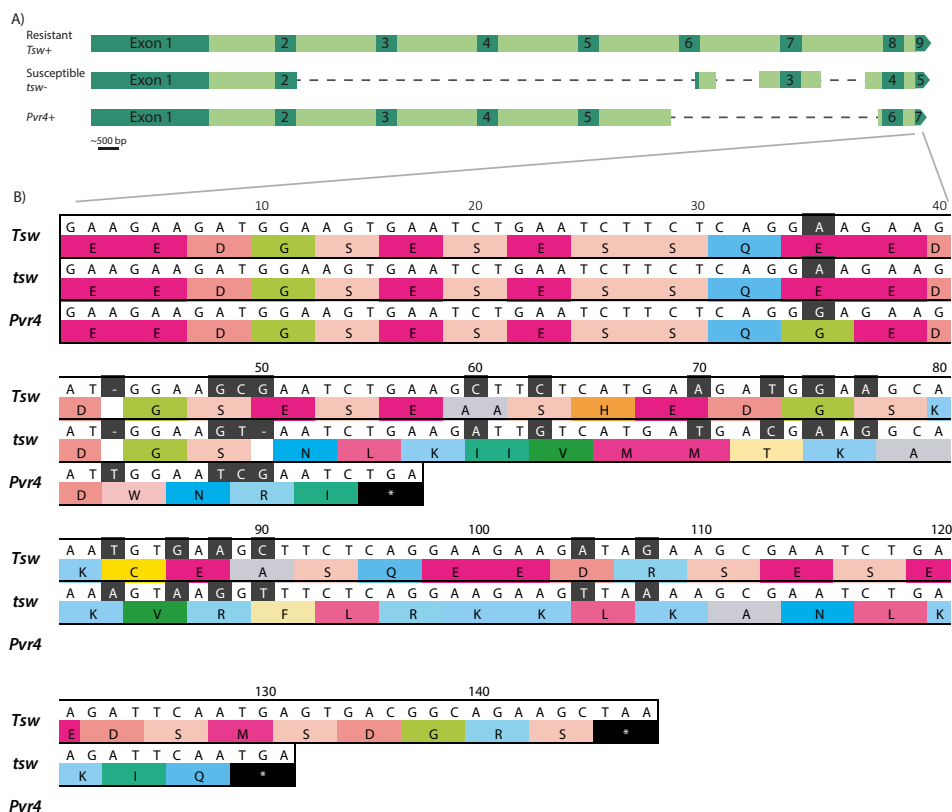
Figure 5.1 – Sequence alignments of the nucleotide sequences of the resistance genes *Tsw*, *tsw*, and *Pvr4*. A) Multiple sequence alignment of the nucleotide sequence *Tsw*, *tsw*, and *Pvr4*. Dark green indicates exons, light green depicts intron sequences. Dotted lines represent gaps. Illustration is roughly to scale. Aligned sequences do not indicate 100% identity match. **B)** Multiple sequence alignment of the 3' end nucleotide sequence of *Tsw*, *tsw*, and *Pvr4*, and the corresponding translation into amino acids below each sequence.

In silico predictions of transcriptional regulation targets of *Tsw* and (related) *R* genes

As non-coding sequences, i.e. UTRs and introns, are often targets for gene expression regulation, an *in silico* analysis was performed to identify such potential target sites. Without knowledge regarding the exact length of the promoter, sequences 2000 nt upstream of the *R* gene translation start site were used for the *in silico* search of regulatory elements: putative binding sites for transcription factors (TFs), *cis*-acting regulatory elements (CAREs), and miRNAs. Several possible transcription start sites (TSS) for *Tsw* and *Pvr4* were found using TSSplant. The TSS closest to the start codons, -253/-255 nt upstream, are most likely the true TSS. CpG islands, associated with gene silencing by DNA methylation, were not found. The MEME motifs targeted by known transcription factors in *Capsicum* were used to predict potential transcription factor binding sites 2000 nt upstream of the *Tsw* start site (Figure 5.2A, Figure S5.2). Only a few small differences were

found between the potential promoter sequences of *Tsw* and *Pvr4*. Target sites of the Dof transcription factor, involved in for instance seed maturation, light-mediated regulation, as well as responses to (a)biotic stress (Gupta et al., 2015), are very prevalent in both sequences (Figure 5.2A, Figure S5.2).

As expected from the alikeness of the 5' UTRs of *Tsw* and *Pvr4*, predicted transcription factor binding motifs in the 5' UTR of *Pvr4* largely overlapped with those predicted for *Tsw* (Figure 5.1B, Figure S5.2B). Further analysis with PlantCARE (Lescot et al., 2002) on the presence of *cis*-acting regulatory elements (CARE) in the regions upstream of *Pvr4* and *Tsw* revealed similarly located TATA-box elements approximately 225 nt upstream of the start codon, as well as several CAAT-box elements, both common *cis*-acting elements in promoter regions (Figure 5.2B, Figure S5.2B). Apart from the predicted (common) overlapping CARE motifs, *Tsw* contains a CGTCA-motif, putatively involved in a methyl jasmonate (MeJA) response (Rouster et al., 1997) (-1370 nt upstream), which is not found in the upstream sequence of *Pvr4*.

Predicted elements for the suspected promoter sequences of *Tsw, Pvr4,* and *tsw* (Figure 5.2) differ somewhat and several elements absent in *tsw* are related to defense responses. The region upstream of *tsw* is predicted to have has less Dof target sites and does not contain any of two the putative Auxin-response factor (ARF) binding motifs predicted for *Tsw* (Figure 5.2A). Furthermore, a putative binding site for the wound-inducible WUSCHEL related homeobox 13 (WOX13) is absent in the *tsw* promoter (Ikeuchi et al., 2022), but present in Tsw and Prv4. The ARF motifs are also absent in the sequence upstream of *Pvr4*, which does however contain a motif to which three different Ethylene Responsive Elements (ERFs) are predicted to bind.

In search of motifs and elements shared with *R* genes, a comparison was made to the 2000 nt region upstream of the ATG start codon from several other NLRs. Two, namely *Sl5R-1* and *Sw-5b*, are derived from *Solanum lycopersicum* and like *Tsw* confer resistance against TSWV, while the *R* gene *Rx1* found in *Solanum tuberosum* confers resistance to Potato Virus X (PVX) upon detection of its coat protein. Common factors found between all *R* gene promoter regions are the many CAAT-boxes, as well as several MYB and MYB-related motifs found in the region -650 to -1500 nt upstream of the ATG. *Sl5R-1* was the only gene to contain many (17) TC-rich repeats and CAT-boxes, while all other promoters were predicted to contain 0-1. For neither TFs nor for CAREs is a trend discernable between the (number of) motifs present and NLRs conferring resistance against TSWV or not.

All *R* genes have ≥22 predicted Dof factor target motifs in the 2000 nt region upstream of the start site, except for *Sw-5b* for which only 9 were predicted. The promoter region of *Rx1* contains relatively few target sites, although the second highest number of WRKY factors is predicted to bind to this region. Many WRKY target sites were also found in the
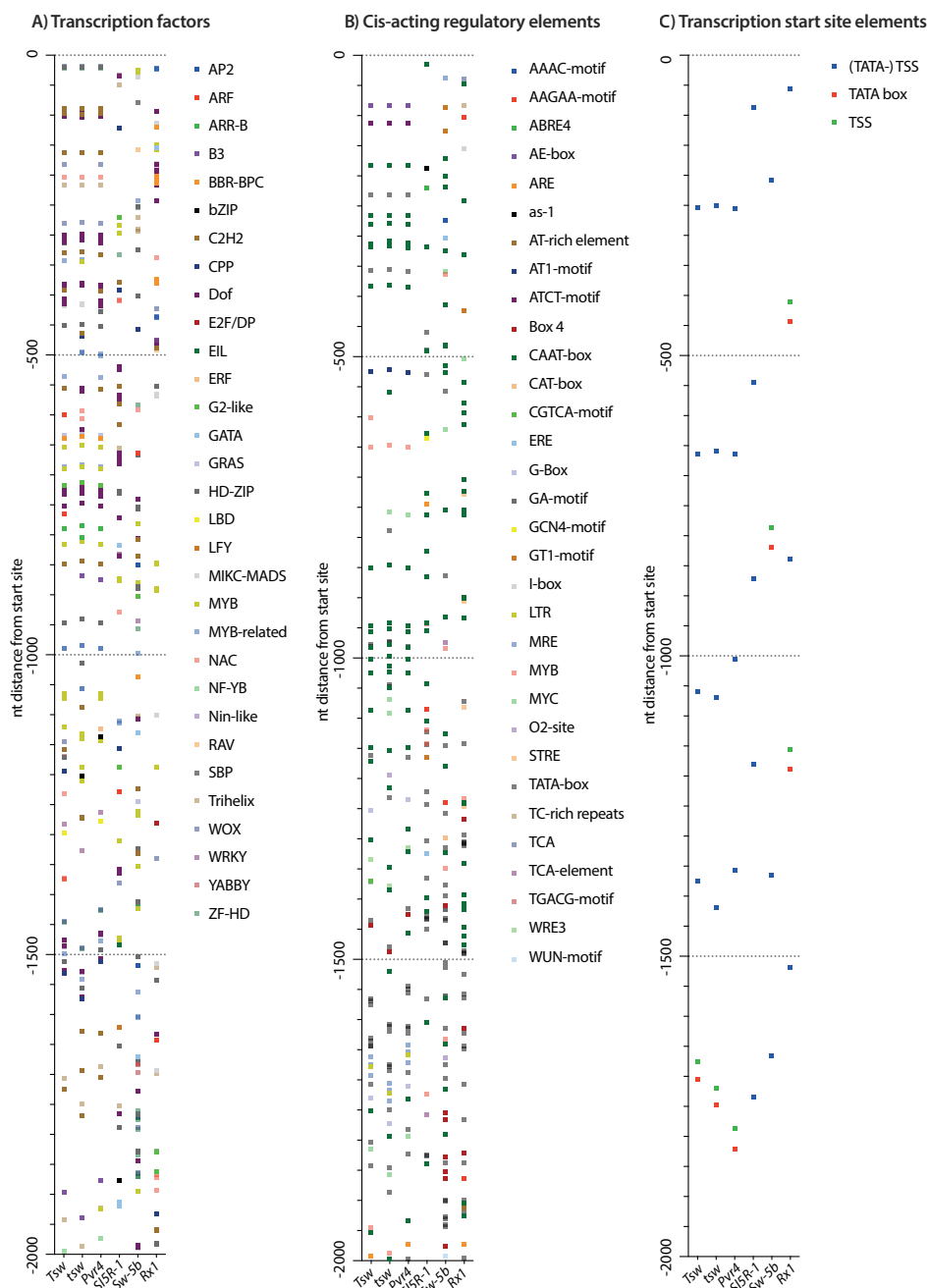
Figure 5.2 – Analysis of the estimated promoter sequences (2000 nt upstream of ATG site) of the resistance genes *Tsw*, *tsw*, *Pvr4*, *Sl5R-1*, *Sw-5b* and *Rx1*. The "ATG" start codon residue is positioned at location 0 on the y-axis. **A)** Location of transcription factor binding motifs predicted by PlantTFDB and FIMO 5.4.1. **B)** Location of cis-acting regulatory elements (CARE) predicted by PlantCARE. **C)** Location of transcription start site elements predicted by tssplant. *Abbreviations: A) AP2 – APETALA2; ARF – auxin response factor; ARR-B – type B authentic response regulator; B3 – highly conserved DNA binding domain ; BBR/ BPC – Barley B Recombinant/Basic Pentacysteine; bZIP – Basic Leucine Zipper Domain; (continues on next page)*

promoter region of *Sl5R-1*, while those of all other genes only contain one WRKY site. *Sl5R*-1 is also predicted to contain many (>20) E2F/DP target sites, while all other genes contain one or none at all in their promoter region. *Sw-5b* is predicted to be targeted by over 20 HD-ZIP and over 10 ZF-HD transcription factors, at least 2.5 times more than promotor regions of any other *R* gene analyzed.

In silico predictions of miRNA targets of *Tsw* and related *R* genes

Potential miRNA target sites in the sequence of *Tsw* were predicted using psRNATarget (Dai et al., 2018), generating a list of miRNAs and their target sequences in the mRNA of *Tsw* as shown in Figure 5.3. Numerous miRNAs of the miR482 family, earlier identified as being involved in *R* gene regulation (Shivaprasad et al., 2012), were predicted to target the domain encoding the P-loop of the NB-ARC. Several other miRNAs indicated before to be involved in plant defense responses, miR172, -6023 and -6026, also were predicted to target sites in *Tsw* (Li et al., 2012; Kravchik et al., 2014; Prigigallo et al., 2019; Zou et al., 2020). Target sites of pepper-specific miRNAs linked to defense responses included miR-n003, -n016a-b, and -n026 (Seo et al., 2018). Various miRNAs unrelated to *R* gene function were also predicted to bind to *Tsw,* including miR414, miR167, and the non-canonical miR-n012 and -n063. Few differences were found between the three gene sequences, except for miR167, only found in the Tsw sequence, and miR166 and -2026 only present in *Tsw* and *tsw*, and an extra miR6023 site found in the region encoding for the NB domain of *Pvr4*.

Effect of miRNA silencing on TSWV titers

To test and verify whether any of these predicted miRNAs genuinely target and regulate *Tsw* expression, their encoding host genes were silenced and the defense response to viral

C2H2 – zinc finger domain; CPP – Cysteine-rich Polycomb-like Protein; Dof – DNA binding with One Finger; E2F/DP – Adenovirus E2 promoter-binding Factor and E2F Dimerization Partner; EIL – Ethylene Insensitive 3-Like; ERF – Ethylene-Responsive-Element-binding Factor; G2-like – Golden2-like; GATA –Transcription factor binding to 'GATA' DNA sequence; GRAS – Gibberellic-acid insensitive, Repressor of gai, And Scarecrow; HD-ZIP – Homeodomain-Leucine Zipper; LBD – Lateral Organ Boundary Domain; LFY – LEAFY; MIKC_MADS – …; MYB – v myb avian myeloblastosis viral oncogene homolog; NAC – NAM, ATAF and CUC; NF-YB – Nuclear Factor YB; NIN-like – Nodule Inception-like; RAV – Related to ABI3/VP; SBP – Squamosa promoter binding; WOX – WUSCHEL-related homeobox; WRKY – Named after amino acid sequence (WRKY) it binds; YABBY – C terminal helix-loop-helix domains; ZF-HD – Zinc Finger-Homeodomain B) AAAC-motif – sequence conserved in alpha-amylase promoters; AAGAA-motif – Unknown function; ABRE4 – abscisic acid responsive; AE-box – Activation Element box; ARE – essential for the anaerobic induction; as-1 – response to auxin and salicylic acid; AT-rich element – binding site of AT-rich DNA binding protein; AT1-motif – part of a light responsive module; ATCT-motif – involved in light responsiveness; Box 4 – Part of a conserved DNA module involved in light responsiveness; CAAT-box – common cis-acting element in promoter and enhancer regions; CAT-box –Related to meristem expression; CGTCA-motif – Related to MeJA-responsiveness; ERE – ethylene-responsive element; G-Box –Involved in light responsiveness; GA-motif – Part of a light responsive element; GCN4-motif – involved in endosperm expression; GT1-motif – light responsive element; I-box – part of a light responsive element; LTR – low-temperature responsiveness; MRE – MYB binding site involved in light responsiveness; MYB – responsive to dehydration and water stress; MYC – Responsive to drought and salt stress; O2-site – zein metabolism regulation; STRE – Stress-responsive element; TATA-box – core promoter element; TC-rich repeats –involved in defense and stress responsiveness; TCA/TCA-element – salicylic acid responsive; TGACG-motif – involved in the MeJA-responsiveness; WRE3 – Wounding and pathogen response element; WUN-motif – wound-responsive element; C) (TATA-)TTS – TATA-less transcription start site; TSS – transcription start site; TATA box – core promoter motif
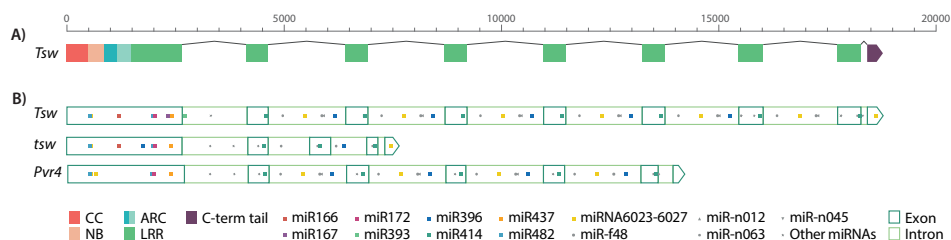
Figure 5.3 – Nucleotide sequences coding for protein domains of Tsw, and the miRNA target sites in *Tsw*, *tsw*, and *Pvr4*. A) A physical map of the domains locations on the coding sequence (CDS) of Tsw (CC – coiled coil; NB – nucleotide binding; ARC – nucleotide binding Apaf 1, R protein, and CED 4 domain; LRR – leucine rich repeat; C term – C terminal). **B)** Representation of miRNA target sites in the Tsw, tsw, and Pvr4 sequence. miRNAs predicted to target sequences with an expectation value (E. value) ≥ 5.0 are presented at their targeting sites in the transcript (dark green – exon; light green – intron). Different miRNAs/miRNA families are presented in different colors (legend included in the figure).
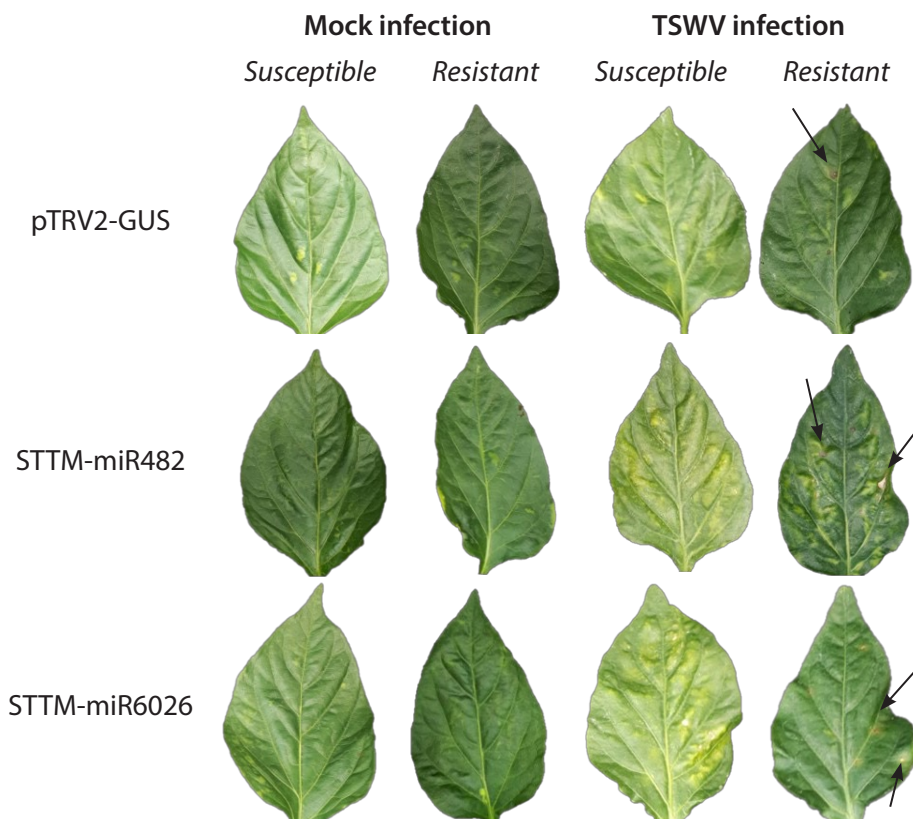
Figure 5.4 – Phenotypes of tomato spotted wilt virus (TSWV) infection on leaves of TSWV susceptible ('Frazier'; Enza Zaden, Enkhuizen, the Netherlands) and resistant ('Mildred'; Enza Zaden) *C. annuum* infiltrated with miRNA silencing constructs or pTRV2-GUS. Plants were agro-infiltrated with pTRV1 and pTRV2-GUS, STTM-miR482, or STTM-miR602 constructs. Silenced plants (mock) infected 14 dpa with TSWV. Images were taken 10 dpi, 24 dpa. Arrows indicate several necrotic lesions visible in the resistant, TSWV-challenged, pepper.

infection analyzed. To this end, two different short tandem target mimic (STTM) constructs were introduced to silence, the two most promising miRNAs (miR482 (light blue) and miR6026 (yellow), Figure 5.3) known from literature to target *R* genes for degradation, followed by TSWV infection of plants and determination of virus titers. To increase the chance of finding an effect of miRNA silencing in this pilot experiment, two related, but slightly different, miRNAs both predicted to silence Tsw were targeted simultaneously through one STTM construct. The sequence of miR482 was combined with miR482l, and miR6026-3p.1 with miR6026-3p.2. The key feature of the STTM constructs, a mismatched bulge in the middle of each miRNA sequence (Figure S5.3), ensures the STTM acts as a non-cleavable target mimic for the RISCs loaded with the targeted miRNAs. This blocks the cleavage activity of RISC and thus lifts silencing of genes targeted by these miRNAs.

As a first pilot, resistant and susceptible *C. chinense* plants were challenged with TSWV 14 days post agro-inoculation (dpa) with pTRV1 and pTRV2-STTM-miRNA or -GUS constructs. Susceptible plants infected with TSWV showed symptoms of the viral infection at the harvest of leaf material 10 dpi (Figure 5.4). Leaves of susceptible plants infiltrated with STTM-miRNA constructs seemed to present slightly more symptoms. Resistant plants infected with TSWV showed local lesions, indicating an HR response upon recognition of NSs by Tsw. While resistant plants infiltrated with the STTM-miR482 construct showed some local lesions, leaves were quite chlorotic and wrinkled compared to infection on resistant peppers without STTM-miR482, possibly indicating spread of the virus despite the *Tsw* gene. Resistant plants infiltrated with STTM-miR6026 showed local lesions, and some mild chlorosis. Silencing of the corresponding miRNAs was verified by RT-qPCR using miRNA stem-loop specific primers, but did not show any signal in the RT-qPCR. However, when a regular PCR was performed with these primers, targeted miRNAs were amplified in both susceptible and resistant pepper plants infiltrated with the pTRV2-STTM-miRNA constructs at similar or higher level than pTRV2-GUS infiltrated plants. Viral titers were determined 10 days post infection (dpi) and in resistant pepper plants silenced with either miR482 or miR6026 STTM constructs turned out to be significantly ($p<0.001$) higher than those in TRV2-GUS infiltrated plants (Figure 5.5). This in contrast to susceptible plants infiltrated with either pTRV2-GUS or -STTM-miRNA constructs, between which no difference in viral titer was found 10 dpi.

Attempts to clone the Tsw (coding) sequence from plant material

The *Tsw* gene recovered from the BAC library contig (Chapter 3), including all exons and introns, spans 18.6 kbp, while the coding sequence (CDS) only is 6.3 kbp in size. Cloning of the latter therefore would most likely enhance chances of successful cloning for the purpose of expression analysis. Using the sequence of *Tsw*, and of that of its surrounding UTRs, many primers were designed to attempt cloning of a full-length *Tsw* cDNA sequence from leaf material. Whereas fragments from the 5' part of *Tsw* up to the 1st exon were successfully cloned attempts to obtain the remaining part, and/or a complete sequence
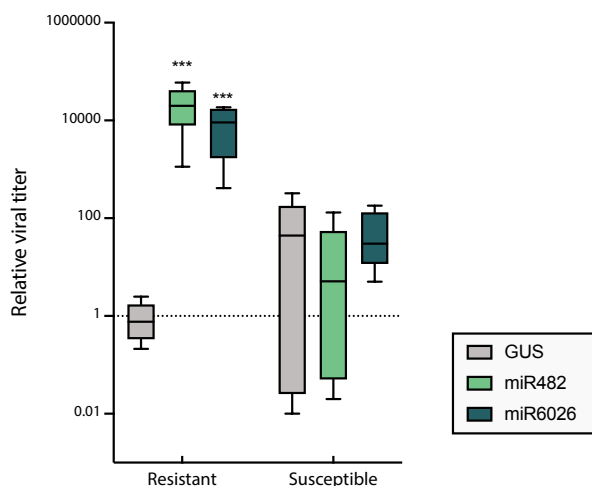
Figure 5.5 – Relative levels of TSWV N in resistant and susceptible peppers agro-infiltrated with miRNA silencing constructs. Relative level of TSWV N gene expression was determined by RT-qPCR in *C. annuum* leaf material (resistant (*Tsw*) 'Frazier' and susceptible (*tsw*) 'Mildred', Enza Zaden) 10 days post infection (dpi) with TSWV and 24 dpa with TRV1 and pTRV2-GUS, or STTM-miR482 or STTM-miR6026. Differences between groups were determined with a Two-Way ANOVA with a Tukey correction on Log transformed data, α=0.05. Error bars represent SE of 6 replicates per group.

up to the 3'end most often resulted in many fragments, each approximately 500 bp larger in size than the other, while the full-length CDS was never obtained.

Design and construction of a synthetic Tsw

Because recovery of a full-length *Tsw* clone from plant material was unsuccessful, and to facilitate a more rapid and easier construction of *Tsw* mutants and chimera in which domains are exchanged for the corresponding ones from the susceptible *tsw* allele, a synthetic gene approach was designed. Due to limitations in length and number of repetitions allowed in one synthetic gBlock, various gBlocks were designed to allow for the subsequent reconstruction of *Tsw* and *tsw* and chimeric derivatives. The recovered and published sequences of *Tsw* and *tsw* were used to design the gBlocks (IDT). As subsequent Golden Gate cloning would be performed using the type IIS restriction enzyme *Bsa*I, one nucleotide in the *Bsa*I restriction site found in exon 1 (104-109 bp) was exchanged to create a synonymous codon alternative ($AGG^{103-105}$>$AGA^{103-105}$), thereby removing the superfluous and inconvenient restriction site. To maximize the use of all synthetic fragments, care was taken to allow for maximum versatility in reassembly of the genes, which simultaneously would allow for future functionality assays of domains, as well as for various truncations and chimeras of the *Tsw* and *tsw* gene. Furthermore, considering that the success rate of Golden Gate drops with more sequences/fragments to ligate (Weber et al., 2011), gBlocks were designed to be used in a two-step Golden Gate assembly process. As correct production of gBlocks is hampered by repeats, care was also taken to not include (large) repeats within single gBlocks. Within the limits of the sequences of *Tsw* and *tsw*, attention was given to create as many different 4 bp overhangs as possible by choosing wisely where gBlocks ended, as (near) identical 4 bp overlaps may anneal, thereby creating incorrect combinations. To maintain options regarding expression vectors in the future, attL1/2 sites were added to gBlocks with 5' and 3' end sequences of *Tsw* and *tsw*, respectively. An overview of the synthetic gBlocks designed,
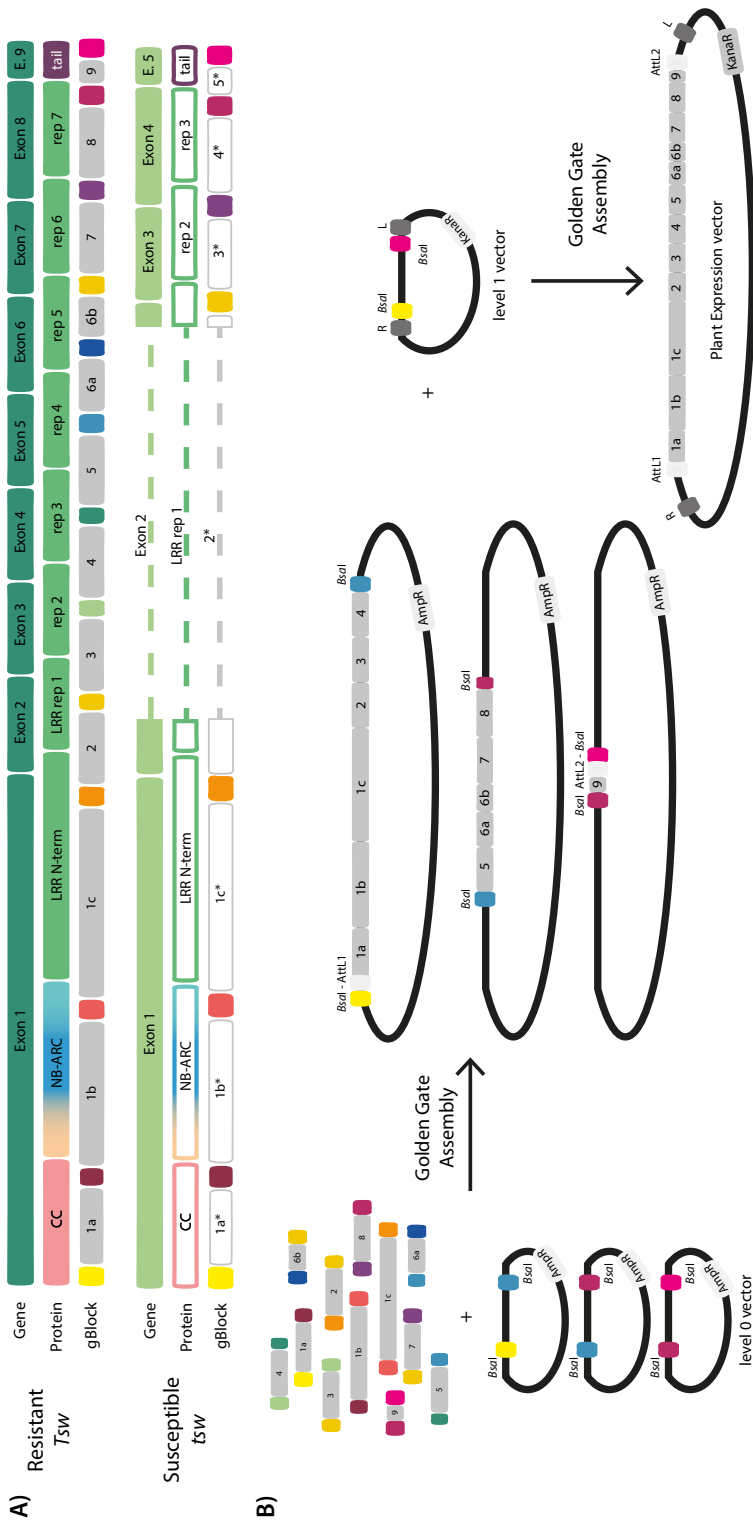
Figure 5.6 – Overview of the construction of a synthetic clone of the resistance gene *Tsw* from *C. chinense* and the non-susceptible version *tsw* from *C. annuum*. A) Location of gBlocks (grey, bottom tracks) mapped on the coding sequence of *Tsw* and *tsw* (dark and light green respectively, top tracks). The domain organization is depicted by the middle track (CC – coiled coil; NB ARC – nucleotide binding Apaf 1, R-protein, and CED 4 domain; LRR – leucine-rich repeat, N term – N terminal; rep – LRR domain repeat). Different colors between gBlocks indicate different overhangs at the 5' and 3' ends of the gBlocks created upon digestion with BsaI, and allowing for seamless annealing and accurate orientation of gBlocks. gBlocks originating from tsw are denoted with a *. **B)** A two-step Golden gate assembly workflow with gBlocks and matching level 0 and level 1 vectors. *Abbreviations: AmpR – Ampicillin resistance; KanaR – Kanamycin resistance; L – Left border; R – Right border.*

their approximate location within the genes, and their different overhang sequences (indicated as differently colored linkers) is shown in Figure 5.6. The C-terminal tail of Pvr4 was also used to construct a gBlock. Using this approach, the construction of the full-length CDS for *Tsw* and *tsw* will be enabled, as well as a total of at least 25 chimeras and truncations for follow up functional gene studies (Figure 5.7). Currently, the first gBlocks have been assembled correctly into level 0 vectors.

Discussion

Many pepper crops used in agriculture are resistant against TSWV due to the introgression of the resistance gene *Tsw*. Its effector NSs has been identified a decade ago (de Ronde et al., 2013), but the downstream pathways activated upon recognition by *Tsw* are unknown. The highly similar sequences of *Tsw*, the susceptible variant *tsw*, and the homolog *Pvr4* were used to find relevant differences and point to possible upstream and downstream pathways related to *Tsw*-mediated resistance. Apart from different LRR lengths, the largest difference between *Tsw*, *tsw*, and *Pvr4* is caused by frameshifts in the C-terminal tail, in the last ~100 amino acids of the genes. Interestingly, the promotor

1. *Tsw*
2. CC - NB-ARC - LRR - c-term tail
3. CC - NB-ARC - lrr - C-TERM TAIL
4. CC - NB-ARC - lrr - c-term tail
5. CC - nb-arc - LRR - C-TERM TAIL
6. CC - nb-arc - LRR - c-term tail
7. CC - nb-arc - lrr - C-TERM TAIL
8. CC - nb-arc - lrr - c-term tail
9. cc - NB-ARC - LRR - C-TERM TAIL
10. cc - NB-ARC - LRR - c-term tail
11. cc - NB-ARC - lrr - C-TERM TAIL
12. cc - NB-ARC - lrr - c-term tail
13. cc - nb-arc - LRR - C-TERM TAIL
14. cc - nb-arc - LRR - c-term tail
15. cc - nb-arc - lrr - C-TERM TAIL
16. *tsw*

17. CC - NB-ARC - Δ2-5 LRR - C-TERM TAIL
18. CC - NB-ARC - Δ2-5 LRR - c-term tail
19. CC - nb-arc - Δ2-5 LRR - C-TERM TAIL
20. CC - nb-arc - Δ2-5 LRR - c-term tail
21. cc - NB-ARC - Δ2-5 LRR - C-TERM TAIL
22. cc - NB-ARC - Δ2-5 LRR - c-term tail
23. cc - nb-arc - Δ2-5 LRR - C-TERM TAIL
24. cc - nb-arc - Δ2-5 LRR - c-term tail

25. CC - NB-ARC - LRR - *Pvr4 C-term tail*
26. CC - NB-ARC - Δ2-5 LRR - *Pvr4 C-term tail*
27. cc - nb-arc - lrr - *Pvr4 C-term tail*

Figure 5.7 – Schematic representation of possible combinations between domains of the resistance genes *Tsw*, *tsw*, and *Pvr4* through Golden Gate assembly with gBlocks. Capital letters and dark shades indicate domains resulting from *Tsw*, small letters and light shades indicate domains from *tsw*. Italics and grey block indicate the C-terminal tail of *Pvr4*. Truncations of the *Tsw* LRR domain in which the 2nd to 5th LRR repeat are removed are indicated by a short LRR (Δ2-5 LRR). *Abbreviations: CC – coiled coil; NB-ARC– nucleotide binding Apaf 1, R protein, and CED 4 domain; LRR – leucine-rich repeat.*

sequences of the functional *Pvr4* and *Tsw*, also exhibit a high similarity, indicating that both genes are controlled by similar transcription regulation mechanisms. The sequence of *Tsw* also revealed the presence of target sequences for miRNAs (miR482 and miR6026) demonstrated to transcriptionally regulate certain *R* genes (Shivaprasad et al., 2012; Wang et al., 2018), of which miR6026 was not found in the sequence of *Pvr4*. Upon their silencing in TSWV resistant pepper plants, viral titers increased and indicate that *Tsw* is also transcriptionally regulated by these miRNAs. Lastly, a method was designed to construct a synthetic *Tsw* CDS, as well as its susceptible variant, and at least 25 chimeric versions. This method does lift any regulations based on untranslated regions and introns.

An intra-gene comparison revealed a very high similarity between individual exons and introns of *Tsw*. Exons 3, 5, and 6 are completely identical, and only slightly differ with other LRR-encoding exons. Intron 2-7 are also highly identical (≥97.7% shared identity). PCR fragments that resulted from attempts to clone *Tsw* often showed a ladder-like distribution of fragments spaced ~500 bp apart, declining in concentration with increasing fragment size. Complete sequences were however never isolated. The similarity of these large 501 bp-sized exons, as well as similarity between the 1,775 bp-sized introns, possibly explains the difficulty of cloning a full-length or cDNA version of *Tsw*. Hairpin structures can be formed from repetitive sequences, which can hamper complete amplification. In addition, incomplete fragments can be inadvertently used as primers, resulting in amplification of unwanted sequences. In general, construction of shorter, incorrect, fragments is favored over the longer, correct, sequences. Despite the use of numerous different primer pairs, polymerases, and attempts to improve the PCR protocol, no full-length sequence, CDS, or LRR domain of *Tsw* was successfully amplified using PCR. In light thereof, the chosen gBlock approach to synthetically reconstruct *Tsw* and *tsw* might turn out to present a more feasible way to reconstruct the gene, while simultaneously providing maximal flexibility for the generation of a large series of variants for functional gene studies.

5

The non-coding sequences of *Tsw* found with the help of the BAC library (Chapter 3) allowed for a first *in silico* search of potential target sites of transcription factors (TFs), *cis*-acting regulatory elements (CAREs), and miRNA targets in the full gene. While *Tsw* resistance is overcome at temperatures above 32 °C (Ronde, 2013; Chung et al., 2018), this is not the case for its homolog *Pvr4* (Janzac et al., 2009). To exclude the influence of temperature-dependent regulatory elements, *Pvr4* was taken along in the *in silico* search for transcription regulatory sequences. As exact promoter regions are unknown, the region 2000 nt upstream of the start sites was used for further analysis. The location of predicted *cis*-acting regulatory elements and transcription start sites indicate that the promoter of *Tsw* and *Pvr4* possibly starts at 253 and 255 nt upstream of the start codon (Figure 5.8). With the exception of a C2H2 target site absent at nt position -139 in the *tsw* sequence, no large differences are found between *Tsw*, *tsw*, and *Pvr4*. No link was found between this TF (CA00g84240) and defense responses. Neither of the TSSs has a TATA-

box ~27 nt upstream but they do have an adjacent initiator (Inr) sequence, as well as a CAAT-box 65 nt upstream. While Downstream Promoter Elements (DPEs) are found in all proposed promoters, it may be too far from the TSS (+50 nt) to function as one. Further indications that this may indeed be the promoter sequence are the many Dof target sites located 50 nt upstream of the TSS, within the possible core promoter region. The Dof transcription factor family promotes expression and has been associated with plant defense responses. Promoter studies of the most recently discovered gene against TSWV, *SI5R-1* from tomato, also revealed numerous Dof target sites in its promoter sequence (Qi et al., 2022). Conversely, the promoter of the two other genes (Solyc05g009750 – *SI5R-2*, and Solyc05g009760 -*SL5R-3*) examined in search of the new NLR, hardly contained Dof target motifs. Dof has not been associated with NLR gene expression before. It could be the case that Dof is one of the transcription factors responsible for high continuous expression of NLR genes, however further research is required to test this hypothesis. Interestingly, in a recent study by Takahashi et al. (2022) on a persistent CMV infection in perennial *Arabidopsis halleri*, the promotor of the DOF1.7 gene was hypomethylated and the gene transcriptionally downregulated. If, and how, DOF1.7 downregulation is involved in the (establishment of a) persistent CMV infection in Arabidopsis is unclear, but it may lead to downregulation of a defense response gene.

Few noteworthy differences were found between the potential promoter sequences of *Tsw* and *Pvr4*. While the sequence upstream of *Tsw* was predicted to contain two ARF binding motifs, that of *Pvr4* contained three motifs for ERFs, both associated with gene regulation upon a pathogen attack. While many transcription factors have been predicted to bind to the upstream sequences of the *R* genes, relatively few are linked to gene regulation upon (biotic) stresses. However, there are some predicted TF targets that do provide promising leads to determine their effect on *Pvr4* and *Tsw* gene expression, such as the binding motif for a MYB transcription factor (CA03g30090). This TF is involved in defense responses, and a binding motif found at -1065 nt from the start codon of the *Tsw* gene. Other MYB TFs have also been shown to be involved in defense responses (Liu et al., 2004; Shi et al., 2021; Zhu et al., 2022). In the case of NLR gene *Sw5a* from
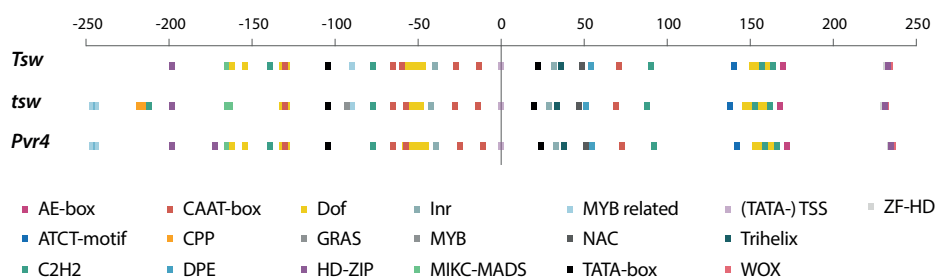
Figure 5.8 – Predicted elements in the proposed promoter sequence of *Tsw*, *tsw*, and *Pvr4*. The location of cis-acting regulating elements (CAREs), transcription factors (TF) and transcription start site (TSS) elements within a ~250 nt range flanking the proposed transcription start site and promoter region of *Tsw*, *tsw*, and *Pvr4*.

tomato, and conferring resistance to the Tomato yellow leaf curl begomovirus (TYLCV), its transcription is regulated by SlMyb33, which in turn, is regulated by miR159 (Sharma et al., 2021b), emphasizing the multilayered regulation of genes. Two binding motifs for a trihelix transcription factor (CA05g10030) are found in the 5' UTR sequence of *Tsw*. The Arabidopsis ortholog of this trihelix protein interestingly plays a role in the negative regulation of pattern-triggered immunity. A WRKY12 (CA01g28150) protein was predicted to bind the 5' UTR of *Tsw* (-1262 nt upstream). WRKY factors are known to play a role in pathogen defense responses. Overexpression of WRKY12 specifically in Arabidopsis and Chinese cabbage has been shown to reduce soft rot symptoms caused by *Pectobacterium carotovorum* (Kim et al., 2014a).

While *Tsw*-mediated resistance is temperature-sensitive, *Pvr4* is not (Janzac et al., 2009; Ronde, 2013; Chung et al., 2018). However, none of the regulatory elements found to differ between *Pvr4* and *Tsw* are known to be temperature dependent. Possibly folding of the encoded proteins is more error prone at higher temperatures, resulting in incorrectly folded R proteins. As the greatest difference between the two genes is the length of the LRR domain, in which for Tsw an extra helical turn is observed (Chapter 4), this might be a key factor. On the other hand, high temperatures are more often found to break resistance by NLRs (Venkatesh and Kang, 2019), most of which are shorter than the exceptionally long Tsw protein.

Several plant defense related miRNA target sites were predicted in the full-length *Tsw* sequence. Of those, the miR482-family is well known for its regulatory effect on *R* genes (Shivaprasad et al., 2012; Jiang et al., 2018), while the non-canonical miR6026 and miR6023 have been associated with *R* gene regulation in potato and tomato (Li et al., 2012; Kravchik et al., 2014; Prigigallo et al., 2019). Interestingly, several of the miRNAs targeting the sequence of *Tsw* were very recently also found to be up- or downregulated during TSWV infection in susceptible pepper (Tao et al., 2022), including the upregulated miR482.

To determine the effect of miRNAs on TSWV resistance, a pilot was set up to silence miRNAs using STTM constructs that have been shown to be very effective in miRNA silencing in *N. benthamiana* and tomato (Sha et al., 2014). Both miR482-family and miR6026 were indicated to target exon sequences, while miR6023 targets intron sequences. For this pilot, two pTRV2-STTM constructs were designed to silence miRNAs targeting coding sequences to allow for better comparison. Unexpectedly, viral titers in both variants of miRNA silenced plants were higher in resistant plants than susceptible plants, while no significant difference in virus titer was noticeable in susceptible plants infiltrated with pTRV2-STTM constructs. However, miRNA levels could not be quantified using RT-qPCR. A regular PCR with miRNA specific primers indicated a higher, rather than a lower, amplification in plants infiltrated with pTRV2-STTM-miRNA constructs. It might be possible that primers accidently anneal to the STTM-miRNA constructs themselves, rather

5

than the miRNAs, or the STTM-miRNA construct inadvertently acts as miRNA themselves. Without proper quantification of miRNAs, it is not possible to determine which of these scenarios has taken place. Nevertheless, infiltration with STTM-miRNA constructs led a change, i.e. high viral titers in *C chinense* plants harboring *Tsw*, which indicates that both miR482 and miR6026 influence transcription of *Tsw*. However, in contrast to what was expected, viral titers in these plants significantly increased relative to those observed in susceptible plants. If infiltration with pTRV2-STTM-miRNA constructs indeed lowered the amount of miRNAs present, *Tsw* would be constitutively highly expressed in the days before virus infection thereby possibly contributing to more protection against viral infection, but the opposite was observed. High expression levels of NLR are known to have a detrimental effect on fitness, and associate with auto-immunity (Xia et al., 2013). Whether high expression levels of *Tsw* might have inadvertently weakened the plant, thereby aiding the virus in its infection, remains to be investigated. Alternatively, when the STTM-miRNA constructs function as miRNAs instead of silencing them, miRNA-mediated regulation of *Tsw* would result in lowered amounts of *Tsw* transcripts and Tsw protein/translation, allowing virus replication to higher levels. To exclude either option the levels of miRNA and Tsw mRNA need to be properly determined. Aside from directly targeting *Tsw* transcripts, the miRNAs may also target other factors which in turn target Tsw transcription regulation and thereby strengthen the suppression of Tsw by miRNAs.

The most prominent differences between *Pvr4* and *Tsw* are the number of LRR repeats, and thus the size of this domain, and their different C-terminal ends caused by a frame shift. Domain and region swaps of *Rx1* and *Gpa2*, two highly similar potato resistance genes (88% amino acid identity) conferring resistance to two different pathogens (PVX and a potato cyst nematode, respectively), showed that the C-terminal end of the LRR determined their ability to recognize the effector (Slootweg et al., 2013). The C-terminal tails of functional *Tsw* and *Pvr4* are highly similar at nucleotide level but differ at amino acid level as a result of a single nucleotide insertion. This not only raises questions on the likelihood of such similar, yet different, C-terminal tails to recognize different pathogens, as well as how recognition takes place; indirectly through the same host-protein, or through, or in concert with, a different region in the NLR protein. Comparison of the non-functional *tsw* with *Tsw* does not provide conclusive information regarding this hypothesis, as both the C-terminal tail and the number of LRR repeats differ enormously.

To simultaneously overcome the difficulties of cloning *Tsw* and answer the questions regarding the functionality of its domains, a method was developed using synthetic gene fragments. As a result of smart design choices, at least 25 chimeras can be constructed besides a synthetic functional Tsw and a non-functional *tsw*. The expression and functionality of these constructs have not been tested yet. Based on domain functionality tests of other NLR genes, it is expected that the CC domain as well as the LRR domain inhibit the activation of the gene until (indirect) recognition of the pathogen. It is possible that *Tsw* is no longer able to (in)directly interact with NSs with a C-terminal tail derived

from tsw (Figure 5.7– 2). If so, a C-terminal tail swap between *Pvr4* and *Tsw* might be enough to exchange pathogen recognition (Figure 5.7– 25-27). Furthermore, truncations of the LRR domain of *Tsw* (Figure 5.7 – 17-24) might either lift its suppression of the central domain, or also render it unable to interact with NSs. The length of LRR domain in tsw is much more in line with that of other NLR genes, while that of Tsw is extraordinarily long and leads to an extra helical turn in the predicted 3D folding structure of its LRR domain (Chapter 4). Currently, it is not known what the (added) function is of the extra length in such NLRs. The inability of *tsw* to convey resistance might be caused by its shorter LRR domain, which can be determined with chimera 3 (Figure 5.7) for instance. If the LRR of tsw is not the root of its inability to convey resistance, the problem potentially lies in the other domains, which can be tested with for instance chimeras 5-15 (Figure 5.7).

The *in silico* analysis of *Tsw*, *tsw*, and *Pvr4* has found many potential leads to possibly uncover regulation, activation, and downstream pathways. *Tsw* has been proven to be a difficult gene to clone, while its synthetic counterpart enables the construction of (tagged) chimeras as well as options to introduce desired mutations without the difficulty of recovering the full-length product. During this approach, additional 3D structural folding predictions of the Tsw protein (Chapter 4) will assist in the design of chimeras that aim to functionally analyze its extraordinarily large LRR domain.

5

Supplemental data
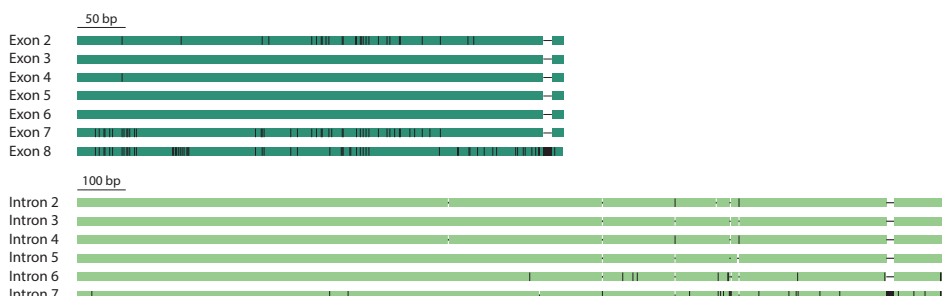
Figure S5.1 – Multiple alignment of *Tsw* exons (dark green) and introns (light green). Alignment of exon sequences 2, 4-8 to exon 3, and intron sequences 2, 4-7 to intron 3. One vertical black strip indicates a 1 nt difference, horizontal bars indicate a gap.

Table S5.1 – Identity comparison of *tsw* exons encoding for the leucine-rich repeat domain. Percentages of shared identity for nucleotides (left bottom, dark gray) and amino acid (right top, light gray) sequences. Exon 1 and 5 respectively encode the NB-ARC domain and the C-terminal tail, and are therefore excluded from this table.

nt\aa	Exon 2	Exon 3	Exon 4
Exon 2		77.2%	70.6
Exon 3	87.6%		76.5%
Exon 4	80.8%	87.5%	
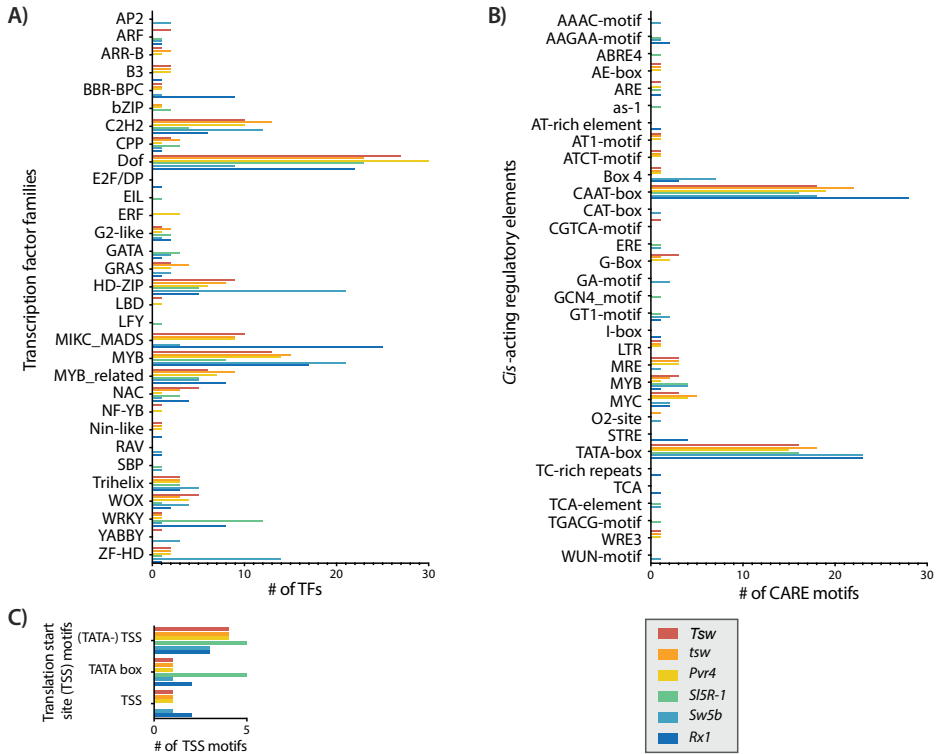
Figure S5.2 – Analysis of the estimated promoter sequences (2000 nt upstream of start site) of resistance genes *Tsw*, *tsw*, *Pvr4*, *Sl5R-1*, *Sw-5b*, and *Rx1*. A) Number of transcription factors (TFs) predicted to bind promoter regions of the R genes as predicted by PlantTFDB and FIMO 5.4.1. **B)** Number of cis-acting regulatory elements (CAREs) in the region 2000 nt upstream predicted by PlantCARE. **C)** Number of translation start site (TSS) motifs in the region upstream of the start sites as predicted by tssplant.
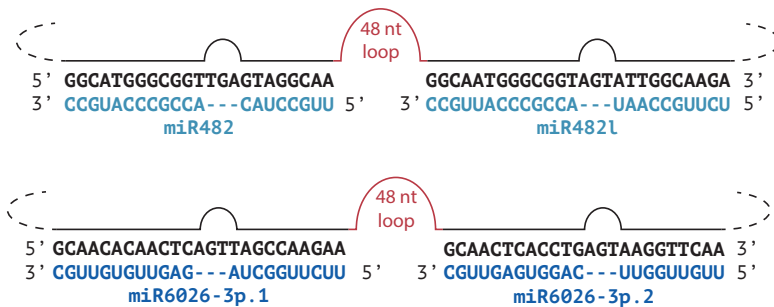
Figure S5.3 — TRV-based miRNA silencing constructs using short tandem target mimics (STTMs). One silencing construct contains two tandem target mimics for miRNAs separated by a 48-nt imperfect stem-loop linker. The two mimic small RNA target sequences are spaced by a mismatched 3-nt loop at the expected cleavage site to arrest its cleavage activity. pTRV2-STTM-miR482 contains miR482 and miR482l target sequences, pTRV2-STTM-miR6026 contains target sequences for miR6026-3p.1 and -3p.2.

Table S5.2 – Primers used in this research in attempts of cloning *Tsw*, and creation of all the different constructs.

Primer Name	Sequence (5'-3')	Purpose
miR482_Fw	TGCGCGGAGAACGGTT	RT-qPCR miR482
miR482_Fw-ext	TGCGCGGTCTTGCCAATAC	RT-qPCR miR482
miR482_RT	GTCTCCTCTGGTGCAGGGTCCGAGGTATTCGCAC-CAGAGGAGACAGGAATGGGCGGTA	For miR482 RT
miR482l_Fw	TGGCGCGGAACGGAT	RT-qPCR miR482l
miR482l_Fw-ext	GCGGTTGCCATCACCG	RT-qPCR miR482l
miR482l_RT	GTCTCCTCTGGTGCAGGGTCCGAGGTATTCGCAC-CAGAGGAGACAGGCATGGGCGGTG	For miR482l RT
miR6026_3p.1_Fw	TGGCGGCAAGAACCGA	RT-qPCR miR6026_3p.1
miR6026_3p.1_Fw-ext	CGCGGCTTCTTCCGTAGA	RT-qPCR miR6026_3p.1
miR6026_3p.1_RT	GTCTCCTCTGGTGCAGGGTCCGAGGTATTCGCAC-CAGAGGAGACAGCAACACAACTCT	For miR6026_3p.1 RT
miR6026_3p.2_Fw	CGGCCCGGAGACTTGG	RT-qPCR miR6026_3p.2
miR6026_3p.2_Fw-ext	AGCCCGGTTGAACCTTCAG	RT-qPCR miR6026_3p.2
miR6026_3p.2_RT	GTCTCCTCTGGTGCAGGGTCCGAGGTATTCGCAC-CAGAGGAGACAGCAACTCACCTGAA	For miR6026_3p.2 RT
pDONR207-Fw	CGCGTTAACGCTAGCATGGATC	Colony PCR BP
pDONR207-Rv	GTAACATCAGAGATTTTGAGACAC	Colony PCR BP
qPCR TSWV cp Fw	CTCTTGATGATGCAAAGTCTGTGA	RT-qPCR
qPCR TSWV cp Rv	TCTCAAAGCTATCAACTGAAGCAATAA	RT-qPCR
RG-ACT-qPCR-F	TGTTATGGTAGGGATGGGTC	RT-qPCR
RG-ACT-qPCR-R	TTCTCTCTATTTGCCTTGGG	RT-qPCR
RG-UBI3-qPCR-F	TGTCCATCTGCTCTCTGTTG	RT-qPCR
RG-UBI3-qPCR-R	CACCCCAAGCACAATAAGAC	RT-qPCR
Stem_Rv	GTGCAGGGTCCGAGGT	RT-qPCR
Stem-Rv-complement	CACGTCCCAGGCTCCA	RT-qPCR
TRV2-Fw-seq	CTCAAGGAAGCACGATGAGC	Colony PCR LR
TRV2-Rv-seq	TGAACCTAAAACTTCAGACACG	Colony PCR LR
Tsw 3'UTR - Rv1	AGGGTCATATTTATGTATTATGCCA	Tsw cloning
Tsw 3'UTR - Rv2	GGGTCATATTTATGTATTATGCCAA	Tsw cloning
Tsw 5'UTR - Fw1	GGAAAAGGGTTACATTCTTTTTAGA	Tsw cloning
Tsw 5'UTR - Fw2	AGAACAAATCAAAAAGAAAATAGCG	Tsw cloning
Tsw Exon 1 - Fw1	ATTCAAAGAGTGGAGGCTGC	Tsw cloning
Tsw Exon 1 - Fw2	ACACTTGATCATACCGCCAGT	Tsw cloning

Primer Name	Sequence (5'-3')	Purpose
Tsw Exon 1 - Fw3	CACTTGATCATACCGCCAG	Tsw cloning
Tsw Exon 1 - Fw4	CTATCACCGACTCAAATCCTCTTTTTAATGAAAAG	Tsw cloning
Tsw Exon 1 - Fw5	GGCCAACAGCCAACAATGCTATCACCGACT-CAAATCC	Tsw cloning
Tsw Exon 1 - Rv1	CTGGCGGTATGATCAAGTGT	Tsw cloning
Tsw Exon 1 - Rv2	AGCCTCAACATTGGGTGAAA	Tsw cloning
Tsw Exon 1 - Rv3	AACCGTAGATGCAACCCCTT	Tsw cloning
Tsw Exon 1 - Rv4	TAACCTTGCACGCAGCTGAT	Tsw cloning
Tsw Exon 1 - Rv5	CTTGTAAGTTTCTCCGGGC	Tsw cloning
Tsw Exon 1 (LRR) - Fw1	GTTTCTTGTCCCAACCTGAATTAGCTTCTGCCGT-CACTCA	Tsw cloning
Tsw Exon 1 (LRR) - Rv1	CTTAGAATTAAACATCGCTTTGTTGAG	Tsw cloning
Tsw Exon 1 (NB-ARC) - Rv1	CTTTTCATTAAAAAGAGGATTTGAGTCGG	Tsw cloning
Tsw Exon 1 (Start) - Fw1	ATGGAAATGGTAACTAATATTTTGAGCCCAGT	Tsw cloning
Tsw Exon 1 (Start) - F2	ATGGAAATGGTAACTAATATTTTGAGCCC	Tsw cloning
Tsw Exon 8 - Fw1	CCTACTTCAGCAAACTTATGCAACTAGG	Tsw cloning
Tsw Exon 8 - Fw2	CATTTAGCACCACAAGACAAACCTT	Tsw cloning
Tsw Exon 9 - Fw1	TCAATGAGTGACGGCAGAAG	Tsw cloning
Tsw Exon 9 - Fw2	ATTCAATGAGTGACGGCAGA	Tsw cloning
Tsw Exon 9 - Fw3	CCTGCTTGGAAATGCTGAAAC	Tsw cloning
Tsw Exon 9 - Fw4	GAGTCCTTATTTCCCTGCTTGG	Tsw cloning
Tsw Exon 9 - Fw5	GAAAATCTGCCAAAGCTGGGG	Tsw cloning
Tsw Exon 9 - Rv1	TTAGCTTCTGCCGTCACTCATT	Tsw cloning
Tsw Exon 9 - Rv2	CGAGACTTGGTGTACTCACAG	Tsw cloning
Tsw Exon 9 - Rv3	CGCTTCCATCTTCTTCCTGAG	Tsw cloning
Tsw Exon 9 - Rv4	TAGCTTCCTTTTACCTGCCCC	Tsw cloning
Tsw Exon 9 (Stop) - Rv1	TTAGCTTCTGCCGTCACTCATTG	Tsw cloning
Tsw Intron 8 - Fw1	TTAGCTTCCTTTTACCTGCCCC	Tsw cloning
Tsw Intron 8 - Rv1	AGCTTCTGCCGTCACTCATTG	Tsw cloning
Tsw-qPCR-Fw1	GGCCAACAGCCAACAATGC	RT-qPCR
Tsw-qPCR-Fw2	GGGCATTGAGTTCCCTCTGT	RT-qPCR
Tsw-qPCR-Rv1	TTGGTGAGAGCACAGAGCAG	RT-qPCR
Tsw-qPCR-Rv2	AGCACTTTCAGGTTGGGACA	RT-qPCR
UBI3-qPCR-Fw	CCTTATGCTGGAGGTGTA	RT-qPCR
UBI3-qPCR-Rw	GCTATTGATGTTAGGATGGAA	RT-qPCR

Chapter 6

Catalase 1 of *Nicotiana benthamiana* acts as a proviral factor for Tomato spotted wilt virus

I.L. van Grinsven, C.M. Voorburg, S. Boeren, A. Goverse, R. Kormelink

Abstract

Tomato spotted wilt virus (TSWV) encodes for a non-structural protein NSs, which acts as RNA silencing suppressor and effector of the dominant resistance gene *Tsw* from *Capsicum*. During an earlier study, mutation of a GW/WG motif at amino acid position 17/18 rendered NSs nonfunctional as both RNA silencing suppressor and effector, which indicated a putative Argonaut 1 interaction. In this study, potential host interaction partners of TSWV NSs in *Nicotiana benthamiana* have been investigated through co-immunoprecipitation and subsequent mass spectrometry. In total 69 and 68 proteins were found to interact with GFP-NSs and GFP-NSs[W17A/G18A] respectively, but Argonaut 1 was not among those. From the putative interacting proteins, Catalase 1, ef1α, and cytosolic glyceraldehyde-3-phosphate dehydrogenase were selected for further analyses. Bimolecular fluorescence complementation validated the interaction between Catalase 1 and NSs, but not with the other two proteins. Subsequent virus induced gene silencing, and hydrogen peroxide and catalase activity measurements indicated that the ability of Catalase 1 to catalyze hydrogen peroxide is promoted by its interaction with NSs. In another experiment, and not as part of this study, NSs was found to interact with the *Solanum tuberosum* SUMO E3 ligase SIZ1 in the nucleus. The role of Catalase 1 as a pro-viral factor, and of SIZ1, in light of the TSWV infection cycle is discussed.

Introduction

Plant viruses have caused devastation and crop losses for centuries (Peyambari et al., 2019). Of all known plant viruses, tomato spotted wilt virus (TSWV) is considered one of the most economically impacting plant viruses (Scholthof et al., 2011). It is able to infect over a thousand different plant species, many of which are important crop and ornamental species, such as pepper, tomato, lettuce, and chrysanthemum (Parrella et al., 2003). TSWV, the prototype of the *Tospoviridae/Orthotospovirus* genus, is a negative/ambisense RNA virus spread by arthropods. The main vector of orthotospoviruses is the western flower thrips (*Frankliniella occidentalis*), which transmits tospoviruses in a persistent and propagative manner (Wijkamp et al., 1993). Since the 1980s, thrips have spread from North America to all other continents (Gilbertson et al., 2015, CABI datasheet, accessed 2022), and thereby has contributed to the worldwide spread and re-emergence of this virus. To combat this virus there have been attempts to control thrips, which so far have proven difficult, and breeders and growers have therefore additionally invested in the search for and exploitation of host plant resistance. Few resistance traits have been identified for TSWV, and currently only two resistance (*R*) genes against this virus are available for commercial breeding, i.e. *Sw-5b* from tomato (*Solanum lycopersicum*) and *Tsw* from hot pepper (*Capsicum chinense*), both a coiled-coil nucleotide binding leucine rich repeat (CC-NB-LRR, CNL) immune receptor (Chapter 3 and 4). However, these *R* genes do not provide a durable resistance, as resistance breaking (RB) strains have emerged for both genes over the years (Ciuffo et al., 2005; Ferrand et al., 2015; Jiang et al., 2017b).

The TSWV virion contains three genomic RNA segments, named for their size: large (L), medium (M), and small (S). The viral RNA-dependent RNA polymerase (RdRp) is encoded by the L segment, the glycoprotein-precursor (GP) of the two glycoproteins (Gn and Gc) and the non-structural movement protein (NSm) by the ambisense M segment, and the nucleocapsid (N) protein and a second non-structural protein (NSs) by the ambisense S segment (de Haan et al., 1991; Kormelink et al., 1991, 1994, 2011; Adkins et al., 1995; van Knippenberg et al., 2002). The NSs protein is an RNA silencing suppressor (RSS) and able to suppress both local and systemic RNA silencing (RNA interference, RNAi) (Takeda et al., 2002; Bucher et al., 2003; Hedil et al., 2015). The protein binds long double stranded RNA (dsRNA) and small interfering and microRNAs (siRNA, miRNA), thereby respectively preventing the processing into small interfering RNA, and uploading and subsequent activation of the antiviral RNA induced silencing complex (RISC) (Schnettler et al., 2010; Hedil et al., 2017). TSWV NSs is furthermore able to trans-complement RNAi-suppressor deficient viruses (Garcia-Ruiz et al., 2018), and during mixed infections break RNAi-mediated resistance against the target virus (Hassani-Mehraban et al., 2009). NSs contains a typical GW/WG motif that has been associated with the binding of Argonaute (AGO) proteins by various other RSS proteins (Giner et al., 2010). Mutation of this motif in TSWV NSs (NSs[W17A/G18A], hereafter referred to as NSs[WG]) compromises its local, but not its systemic, RNA silencing ability. Whether NSs truly interacts with AGO1 is unknown. Alanine substitution analysis in NSs has indicated that the N-terminal part of NSs seems involved in RSS activity (de Ronde et al., 2014b). Apart from acting as RSS, the NSs protein also has been identified as the effector of *Tsw*-mediated resistance in *Capsicum* (de Ronde et al., 2013). The previously mentioned alanine substitution analysis has indicated that the N-terminal region of NSs also seems most relevant for its role as effector. Various *Tsw* resistance breaking isolates of TSWV have been reported in the field and collected, for which a single amino acid mutation in NSs can already be sufficient to break resistance (Almási et al., 2017).

6

NSs does not only play a role in the RNAi pathway, but is also involved in various other host processes, including its potential role in hijacking other defense pathways through its direct interaction with defense related proteins Calmodulin 3 and Importin subunit α (Zhai et al., 2021). Furthermore, NSs is necessary for persistent infection and transmission of TSWV by thrips (Margaria et al., 2014), and increases plant attractiveness for thrips by its influence on the secretion of volatiles and by suppressing JA-regulated plant defenses by direct interaction with the transcription factors MYC2, 3, and 4 (Wu et al., 2019; Du et al., 2020). The NSs protein from the orthotospovirus Tomato zonate spot virus was furthermore shown to interact with MYC2, suggesting this is a conserved pathway to prevent a defense response. Tomato zonate spot virus NSs also was recently linked to the hypersensitive-response associated protein Zingipain-2 like. Apart from the interaction with host proteins, the NSs proteins of Groundnut bud necrosis virus, watermelon silver mottle virus, and Capsicum chlorosis virus (CaCV) form dimers (Abraham and Savithri, 2016; Widana Gamage and Dietzgen, 2017; Huang et al., 2020). For Groundnut bud

necrosis virus NSs, authors proposed the requirement of dimerization to function as bidirectional DNA and RNA helicase (Bhushan et al., 2015; Abraham and Savithri, 2016).

Whereas the involvement of NSs and role in various host processes slowly becomes unveiled, many details of the underlying processes and the interactions of NSs with other host proteins have remained elusive. Here, in order to identify potential host interactors of NSs, including the potential interactor AGO1, affinity purification and subsequent mass spectrometry analysis have been used with NSs and the NSs[WG] mutant as bait. From a small set of selected candidates for further analysis, the interaction of Catalase 1 with NSs was validated by bimolecular fluorescence complementation (BiFC) and reverse affinity purification. Additional experiments indicated a pro-viral role of Catalase 1 during TSWV infection.

Materials and methods

Virus, plants, and constructs

Nicotiana benthamiana plants were grown and maintained under greenhouse conditions (22 °C; 16 h light/8 h dark cycle; 70% relative humidity). *N. benthamiana* 16C plants, constitutively expressing GFP, were grown likewise. TSWV isolate Vir129 (Wageningen, the Netherlands) was maintained on *N. benthamiana* by maximally five mechanical serial passages and stored as frozen leaves at -80 °C. Plants that suffered due to severe mishandling were removed from subsequent analysis.

Sequences of NSs and NSs[WG] previously obtained by De Ronde et al. 2013 were Gateway cloned into destination vectors pK7WGF2 and pK7FWG2 (Karimi et al., 2002), and transformed into *Agrobacterium tumefaciens* LBA4404 (Ooms et al., 1982).

Leaf material of *N. benthamiana* was used for TRIzol RNA isolation, of which 1-3 μg of RNA was used as a template for first strand complementary DNA (cDNA) synthesis with Superscript III One Stop (Invitrogen, Waltham, MA, USA). Gene specific primers for the potential interaction partners (Table S6.4) were designed based on DNA sequences available at NCBI. These primers were used for first strand synthesis and subsequent PCR amplification with Phusion high-fidelity *Taq* polymerase according to the manufacturer's specifications (New England Biolabs, Ipswich, MA, USA). Amplified PCR fragments were visualized on a 1% agarose gel. Fragments of the correct size were excised and gel-purified and used for PCR amplification with attB-containing gene-specific primers (Table S6.4). AttB-flanked fragments were used for Gateway Cloning into donor vector pDONR207 and subsequently into destination vectors pDEST-SCYNE(R)GW, pDEST-GWSCYNE, pDEST-SCYCE(R)GW, pDEST-GWSCYCE (Gehl et al., 2009), pGWB454 (NAKAGAWA et al., 2007), pK2GW7 (Karimi et al., 2002), and transformed into *Escherichia coli* and *A. tumefaciens* LBA4404 (Ooms et al., 1982).

Target sequences of approximately 300 bp for virus induced gene silencing (VIGS) were created using the cDNA of the genes of interest as a template for Phusion PCR amplification and gene-specific primers (Table S6.4). Amplified PCR fragments were visualized on a 1% agarose gel. Fragments of the correct size were excised and gel-purified and used for PCR amplification with AttB-containing gene-specific primers (Table S6.4). These were then used for Gateway Cloning into Tobacco rattle virus (TRV) 2 vectors (Liu et al., 2002) and transformed into *E. coli* and *A. tumefaciens* LBA4404 (Ooms et al., 1982). TRV1 and pTRV2-GUS and -PDS silencing constructs were constructed previously (Senthil-Kumar and Mysore, 2014), as well as the pBIN-HA-p19 (Lakatos et al., 2004), pBIN-GFP, pBIN-GUS, pBIN-p19 (Mlotshwa et al., 2002), and pBA-2b (Zhang et al., 2006).

A peroxisome marker was constructed by adding a PTS1 tag to mRFP from pGWB454 through PCR with mRFP specific primers with restriction sites and a C-terminal PTS1 tag with Q5 polymerase (New England Biolabs) following the manufacturer's procedure. The product was visualized on a 1% agarose gel and excised, gel-purified, restriction enzyme digested and ligated back into the pGWB454 vector, and subsequently transformed into *E. coli* and *A. tumefaciens* LBA4404.

Transient expression of proteins

Proteins were transiently expressed via an Agrobacterium Transient Transformation Assay (ATTA). The *A. tumefaciens* LBA4404 and GV3101 cells containing the desired constructs were grown according to Bucher et al. (2003), or after slight modifications. In brief, *A. tumefaciens* was grown for 18-36 hours at 27 °C in LB3 medium containing appropriate antibiotic selection pressure as well as 20 µg/ml rifampicin. The optical density at 600 nm (OD_{600}) of the cultures was determined, the cells were pelleted and resuspended in MMAi medium (Murashige-Skoog induction medium; 10 mM MES; 0.2 µM acetosyringone) to obtain the desired OD_{600} of 0.5. Plants were placed in a small layer of water and kept under dark conditions for 1 h before infiltration with a needleless syringe.

6

GFP silencing suppression assay

Leaves of *N. benthamiana* were agroinfiltrated as described above with LBA4404 containing a GFP-NSs[(WG)]/ NSs[(WG)]-GFP construct, with or without a LBA4404 harboring pBin-HA-p19, using a final OD_{600} of 0.5. Infiltrated leaves were monitored for GFP expression at 5 days post agro-infiltration (dpa) using a hand-held UV lamp. For the quantification of GFP fluorescence, leaf discs with a diameter of 1 cm were taken from the infiltrated leaf area 5 dpa and fluorescence units were measured using a Fluorstar Optima (BMG Labtech, Ortenberg, Germany).

GFP systemic silencing assays in *Nicotiana benthamiana* 16C plants

Seedlings of 1-2-week-old *N. benthamiana* 16C plants that constitutively express GFP (Voinnet and Baulcombe, 1997) were agro-infiltrated as described above with previously

mentioned TRV1 and TRV2-GUS, -PDS, -Cat1 constructs. To induce systemic silencing of GFP, the L3 and L4 leaves of these plants were agro-infiltrated with pBinGFP, and when testing for suppression of GFP silencing co-infiltrated with pBinGFP and pBinNSs, -GUS, or -2b 14 dpa with TRV constructs. Photographic images of GFP fluorescence of leaves L9 and L10 in UV light were made 17 days later with a Sony DSC-HX60V and scored using the Visual Systemic Silencing Index (Hedil et al., 2015).

Co-immunoprecipitation (co-IP)

Infiltrated leaves were harvested 4 dpa and snap frozen in liquid nitrogen. Plant tissue was ground in liquid nitrogen and extraction buffer (50 mM Tris, pH 8.0; 150 mM NaCl; 5 mM EDTA; 1% protease inhibitor (Sigma-Aldrich, Burlington, MA, USA)) with 1% NP-40 was added in 1 g sample:2 ml buffer ratio. After incubation on ice until completely thawed, the sample was centrifuged for 30 min at 34,000 x g at 4 °C. GFP-trap magnetic agarose beads (ChromoTek, Planegg, Germany), washed beforehand with extraction buffer, were added to the supernatant, and incubated end-over-end for 1 hour at 4 °C. Beads with samples destined for Western blot were washed three times with extraction buffer with 1% NP-40, resuspended in protein denaturation buffer and loaded on a 12% bisacrylamide gel. If destined for Mass Spectrometry, beads were washed two times with extraction buffer with 0.1% NP-40, subsequently washed once with extraction buffer without detergent, and then twice with 50 mM ammonium bicarbonate.

Mass spectrometry

For cysteine reduction and alkylation, the co-IP samples used for Mass spectrometry were incubated in 29 mM dithiothreitol in 50 mM ammonium bicarbonate for 1 hour at 60 °C, followed by incubation in 27 mM Iodoacetamide in 50 mM ammonium bicarbonate in the dark at room temperature for 30 minutes. L-cysteine dissolved in 50 mM ammonium bicarbonate was added to the sample to a final concentration of 32 mM to stop the alkylation. The proteins were enzymatically digested by adding 100 µl 5 ng/µl sequencing-grade bovine trypsin (Roche, Basel Switzerland)) in 50 mM ammonium bicarbonate and incubated overnight. The pH was brought down to 3 by adding 10% trifluoracetic acid in 50 mM ammonium bicarbonate. The samples were cleaned by running them over a home-made C18 microcolumn containing an C18 empore disk with 2 mg Lichroprep C18. The resulting peptides were analyzed by liquid chromatography-tandem mass spectrometry (LC-MSMS) using the label free quantification (LFQ) settings and peptides were identified based on the 1.0.1 *N. benthamiana* Solgenomics database (https://solgenomics.net/organism/Nicotiana_benthamiana/genome) using MaxQuant v2.1.1.0 and Perseus v2.0.5.0. T-tests and 2-way ANOVAs were performed to determine significant differences between the LFQ values of the different groups (Lu et al., 2011).

Confocal microscopy

Agro-infiltrated leaf material of *N. benthamiana* was analyzed 3 dpa. The signal of GFP, RFP, and SCFP in leaf epidermal cells was detected by confocal laser scanning microscope

(Zeiss LSM 510-META 18) using the Plan-Apochromat 63x oil immersion objective. The brightfield images were always collected via the Differential Interference Contrast channel and pinhole were set to 1. For GFP imaging, an argon laser was used with an excitation of 488 nm. The emitted light passing the dichroic beam splitter 488/543, was reflected by the secondary dichroic beam splitter 545 and then passed the band pass 505-530, upon which the signal was detected. For RPF imaging, a helium neon laser with an excitation of 543 nm was used. The emitted light passed a 488/543 dichroic beam splitter and next a secondary dichroic beam splitter 490, followed by a band pass 560-615, upon which the signal was detected. An argon laser with an excitation of 458 nm was used for SCFP imaging. The emitted light passed a dichroic beam splitter 458 and next was reflected by a secondary dichroic beam splitter 545, followed by a band pass 470-500, upon which the signal was detected. When multiple fluorophores were imaged in one sample, sequential scan setting was used. The obtained images were processed with the software program Fiji (Schindelin et al., 2012). A minimum of two independent *Agrobacterium* infiltrations and subsequent imaging was used for each observation.

Gene ontology analysis and protein-protein interaction networks

As there currently is no gene ontology data or analysis tool for *N. benthamiana*, the GO terms of its closest homologs in tobacco, tomato, and *Arabidopsis thaliana* were found. Solely the GO terms from *A. thaliana* were sufficient for comprehensive downstream GO term analysis, therefore only these were used for GO term overrepresentation and enrichment analysis with Panther17.0. Bar graphs were generated with Prism 9.3.1. Area-proportional Venn diagrams were made using a webtool (http://apps.bioinforx.com/bxaf7c/app/venn/index.php, accessed 22 March 2022). A full STRING network illustrating interactions between A. thaliana homologues of putatively interacting proteins of GFP-NSs was made using STRING Version 11.5, based on evidence, databases, and co-expression. The minimum interaction score was 0.700, the second shell shows up to 5 interactors.

Western blot analysis

Transient expression of proteins in *N. benthamiana* leaves was verified by SDS-PAGE and Western immunoblot analysis as described previously (de Ronde et al., 2013). Polyclonal rabbit antisera were used against TSWV NSs (Kormelink et al., 1991; de Ronde et al., 2013), catalase (Agrisera, Vännäs, Sweden), and GFP (Invitrogen).

Detection of H_2O_2 and catalase activity

Leaf material from TSWV-infected, mock-infected, and inoculated plants was collected and snap-frozen in liquid nitrogen. The H_2O_2 content was determined following the protocol of Mátai and Hideg (2017). Plant material was ground and homogenized in 0.1% trichloroacetic acid, followed by centrifugation at 12,000 x g for 15 minutes. The supernatant was mixed with 10 mM potassium phosphate buffer (pH 7.0) and 1 M potassium iodide in a 1:1:2 distribution. Absorption measurements were made in triplicate

at 390 nm on a Tecan Infinite M200 Pro. The calculations of the H_2O_2 concentration were based on a standard curve created with standard dilutions of 30% H_2O_2.

Catalase activity was determined using a modified protocol based on the method of Aebi (Aebi, 1984) by measuring the degradation of H_2O_2 by catalase over time. To this end, frozen plant material was ground and homogenized in extraction buffer (0.1 M phosphate buffer, pH 7.5; 0.5 mM EDTA) and centrifuged at 15,000 x g for 20 minutes at 4 °C. The supernatant was added to the reaction buffer (50 mM potassium phosphate buffer (pH 7.0); 20 mM H2O2; 2 mM EDTA Na2) in 1:29 ratio and the change in A_{240} was measured for 6.5 minutes. Equal amounts of proteins in reaction buffer without H_2O_2 was used as a blank.

To normalize the data of all assays, protein contents of the supernatants from the catalase activity assay and the H_2O_2 detection were quantified with the Pierce BCA Protein Assay kit (Thermo Scientific, Waltham, MA, USA).

Gene expression analysis

Total RNA of *N. benthamiana* leaf material 14 dpa or 14 days post infection (dpi) was extracted with TRIzol reagent (Invitrogen) and DNase treated with the DNA-*free*™ Kit (Ambion, Waltham, MA, USA). First strand cDNA was synthesized from 2 ng of total RNA with oligo dT primers and M-MLV reverse transcriptase in a 20 µl reaction (Promega, Madison, WI, USA). The level of target gene mRNA was determined by performing an RT-qPCR with SYBR™ Select Master Mix (Applied Biosystems, Waltham, MA, USA) on 10-fold diluted cDNA and gene-specific primers (Table S6.4) on a Bio-Rad CFX96. The efficiency of all primer pairs was determined, and melt curves were analyzed. The expression levels of NbPRR and NbEf1α (Liu et al., 2012) were used to normalize expression levels of target genes using the Pfaffl method adapted for multiple reference genes (Vandesompele et al., 2002; Hellemans et al., 2007), and calibrated to the level of control plants (pTRV2 GUS infiltrated or mock-infected plants, 14 dpa and 14 dpi respectively). Experiments were performed twice, with two technical replicates per sample. Data were analyzed with Microsoft 365 Excel and Prism 9.3.1. K-means clustering was performed using R studio.

SUMO predictions in NSs

Predictions of SUMO-attachment sites (SAS) and SUMO interaction motifs (SIM) were made on the sequences of the NSs proteins of TSWV (P26002) and CaCV (Q0P0H3) with the GPS sumoylation prediction server (high threshold for both sumoylation and SUMO interaction) and the Jassa server (only "high cut-off" results for predicted SASs).

Results

Construction and verification of NSs bait constructs

As a first step towards the extraction and identification of plant interaction partners of NSs and mutant NSs^WG, C- and N-terminal GFP-tagged bait constructs were made and cloned into the pK7WGF2 and pK7FWG2 binary expression vectors. Expression of the constructs was achieved in *N. benthamiana* leaves via an ATTA and subsequently confirmed by Western immunoblot analysis of infiltrated leaf material (Figure 6.1A, B). Next, the functionality of NSs tagged with GFP was tested with a local GFP silencing suppression assay. To this end, leaves of *N. benthamiana* were infiltrated with only a GFP-tagged NSs construct (Figure 6.1C), or co-infiltrated with p19 (RNA silencing suppressor of *Tombusvirus*, data not shown) as a positive control, and the relative fluorescence was measured 5 days post agro-infiltration (dpa) (Figure 6.1D). Although the C- and the N-terminally GFP tagged NSs expressed to different levels, they were both able to suppress silencing in the local leaf patch assay. A similar difference in the expression levels of GFP-tagged NSs^WG mutants was observed but, in contrast to the NSs constructs, both mutant constructs lacked the ability to suppress RNA silencing, in agreement with previous studies (de Ronde et al., 2014b).

NSs affinity purification and identification of interacting host proteins

As the yields of intact, C-terminally tagged, NSs-GFP and mutant NSs^WG-GFP were lower than their N-terminally GFP-tagged counterparts (Figure 6.1A, B), the latter were used for
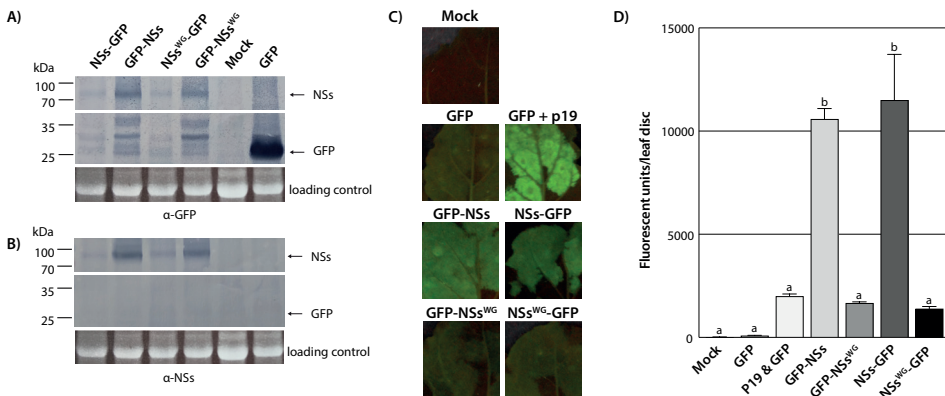
Figure 6.1 – Expression of N- and C terminally GFP-tagged NSs constructs in a pK7 vector in *A. tumefaciens* LBA4404. A) Western immunoblot detection of the GFP-tag 4 days post agro-infiltration (dpa) of NSs constructs in *N. benthamiana*. Free GFP is used as a positive control. **B)** Western immunoblot carried out as in panel A with a polyclonal antiserum against NSs. **C)** Fluorescence of GFP-tagged proteins 5 dpa in N. benthamiana. **D)** Quantification of fluorescence units measured in 1 cm² leaf disc collected from the agro-infiltrated leaf material shown in panel C. Error bars represent the standard error of the mean (SEM) of two replicates. Differences between groups were determined with a One-Way ANOVA test with Šídák correction for multiple comparisons, α=0.05. Groups with different letters (a, b) indicate a significant difference between them with a p-value <0.01.
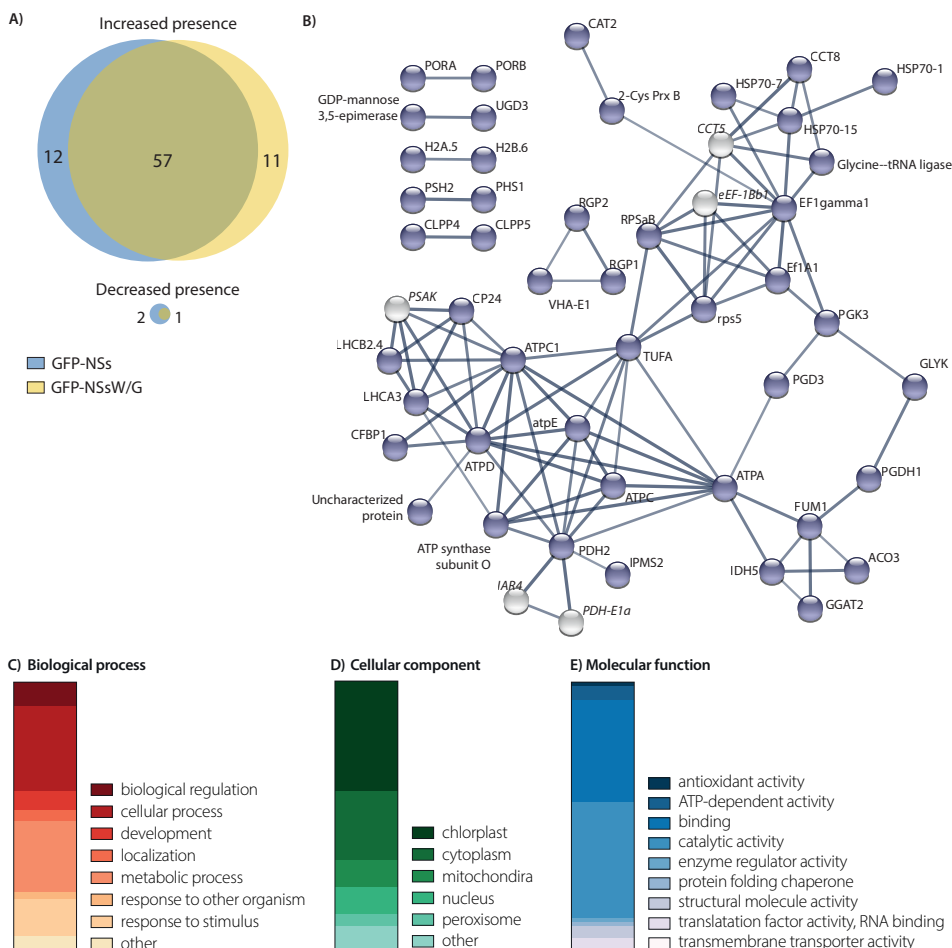
Figure 6.2 – Analysis of proteins putatively interacting with NSs(W/G). A) A Venn diagram illustrating the overlap between putatively interacting proteins found with LC-MSMS using GFP-tagged NSs or mutant NSs[WG] protein as bait. **B)** The network illustrating interactions between *A. thaliana* homologues of putatively interacting proteins of GFP-NSs (colored) using STRING Version 11.5, and up to five interactors in the second shell (white). **C), D), E)** Bar graphs displaying the relative distribution in GO terms for biological process, cellular component, and molecular function of the *A. thaliana* homologues of the proteins putatively interacting with NSs (Panther 17.0).

affinity purification and LC-MSMS as described in the Materials and methods. Unbound GFP was included as a negative control. All samples were prepared in triplicate. Proteins that did not significantly differ following a Two-Way ANOVA and Tukey correction ($\alpha=0.001$) in their abundance from the unbound GFP sample were seen as non-specific interactors and removed from the list of putative host protein interactors. In total 69 and 68 potential interaction partners of respectively GFP-NSs and GFP-NSs[WG] were identified (Table S6.1), of which 57 were observed in both experimental groups (Figure 6.2A). None of the proteins absent from the GFP-NSs samples in comparison to GFP-NSs[WG], or vice

versa, stood out as notable proteins. Consequently, downstream analyses were only performed on proteins found to potentially interact with GFP-NSs.

The closest *Arabidopsis thaliana* homologs of the 69 potential interaction partners of GFP-NSs were determined (Table S6.1) to facilitate GO term analyses. Enrichment analysis of the potentially interacting proteins indicated an overrepresentation of proteins localized in the mitochondrial proton-transporting ATP synthase complex and chloroplast stroma.

Validation and confirmation of genuine NSs-host protein interactions

From the list of candidate host proteins putatively interacting with NSs, a small set was selected for further analysis based on GO terms and literature research regarding involvement in other pathogen-host interactions. Catalase 1 and ef1α were selected for further analyses, as well as cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPC1, Niben101Scf01073g05003.1). GAPC1 is generally involved in HR-mediated cell death (Han et al., 2015), but was not significantly overrepresented among the list of host proteins interacting with either GFP-NSs or GFP-NSs[WG], and was therefore included in the BiFC assay as an independent control for the MS analysis. AGO1 did not show up in the list of candidate host proteins either, but, in light of the WG motif in NSs, was taken along for further analysis as well. As a first step to verify the putative interactions between NSs
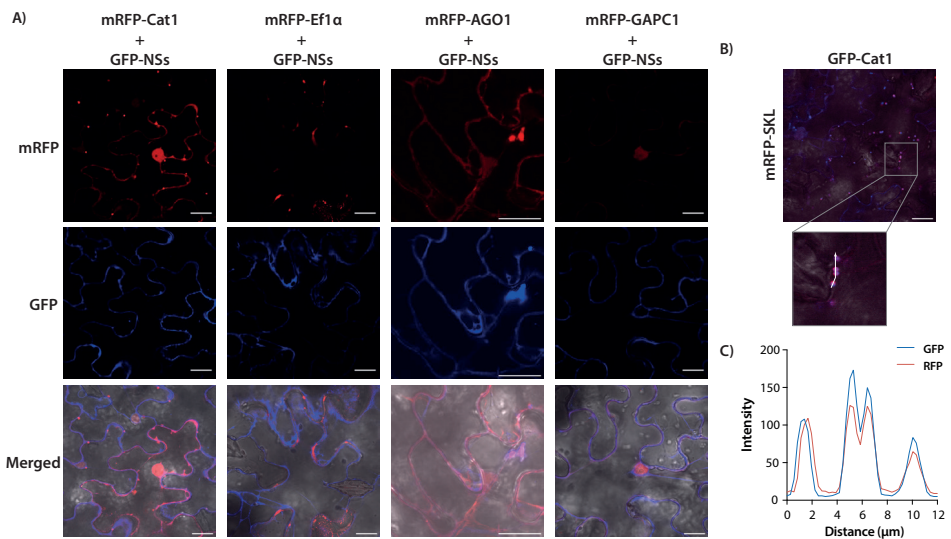
Figure 6.3 – Localization studies of NSs with four (putatively interacting) proteins. A) Confocal microscopy of co-infiltrations of GFP-NSs and three RFP-tagged potential interaction partners of NSs (RFP-Ef1α, RFP-Catalase 1, RFP-AGO1, RFP-GAPC1). Samples were harvested 3 days post agro-infiltration. The RFP fluorescent signal (red), GFP fluorescent signal (blue), and the two fluorescent signals merged with the brightfield image are shown. The scale bars indiacte 20 µm. **B)** Co-expression of GFP-Cat1 and the peroxisome marker mRFP-SKL. A merged image of the GFP (blue) and RFP (red) fluorescent signal, and the brightfield is shown, as well a fluorescence intensity plot of both fluorescent signals. The intesity of GFP and RFP along the arrow is depicted.

and the selected host proteins, an *in situ* co-localization analysis was performed with GFP-NSs and RFP-tagged host proteins following a co-agroinfiltration of the respective binary transient expression constructs (Figure 6.3A). As predicted based on literature and UniProt annotation, RFP-GAPC1 (red) indeed was found in the nucleus but not the nucleolus and did not co-localize with GFP-NSs (blue), which was expressed and found only in the cytoplasm. RFP-EF1α (red) localized to small condensations in the cytoplasm but did not show a clear co-localization with GFP-NSs (blue). RFP-AGO1 (red) localized to the cytoplasm, as well as small aggregations in the cytoplasm. Co-localization with NSs-GFP (blue) was seen in these aggregates. RFP-Catalase (red) was found in small bodies spread out over the cytoplasm, of which some co-localized with GFP-NSs (blue) (Figure 6.3A). Considering that catalases primarily localize in peroxisomes, a co-expression of GFP-Catalase (blue) and peroxisomal marker RFP-SKL (red) was performed that confirmed the identity of these Catalase bodies as peroxisomes (Figure 6.3B).

Next, to analyze for the occurrence of genuine interaction of the host proteins with NSs, all corresponding genes were cloned into the bi-molecular fluorescence complementation (BiFC) vectors SCYNE and SCYCE. Prior to the BiFC analysis, the *in planta* expression of all constructs was verified by Western Blot analyses of leaf samples collected from agro-infiltrated leaves, and the functionality of SCYCE-NSs as RSS was confirmed in a local GFP silencing leaf patch assay (Figure 6.4). For the BiFC analysis, SCYCE-NSs was co-infiltrated in an ATTA with one of the four potential interactors under investigation, all cloned into the SCYNE vector and therefore N-terminally tagged with an N-SCFP3A. Upon co-infiltration of the constructs encoding for either SCYNE-AGO1, SCYNE-GAPC1 or SCYNE-Ef1α with SCYCE-NSs no CPF signal was observed, indicating that the proteins did not interact (data
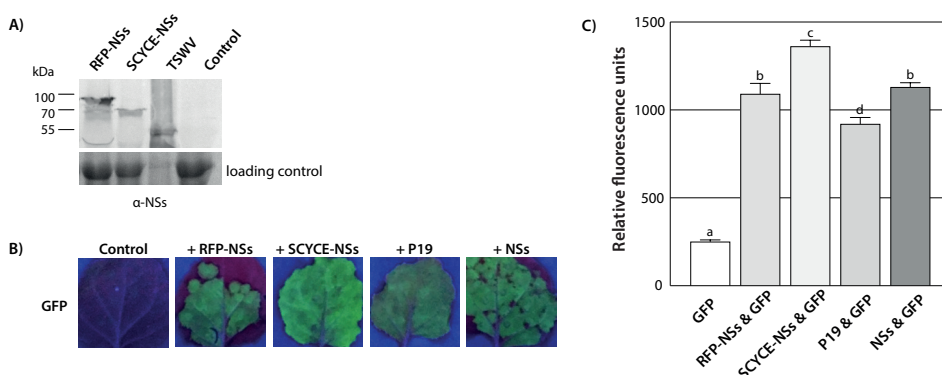
Figure 6.4 – Expression of C-terminally C-SCFP3A-tagged NSs construct in the pDEST-GWSCYCE vector (Gehl et al., 2009) by *A. tumefaciens* LBA4404 in *N. benthamiana* leaves. A) Western immuno-blot detection of NSs protein expression 3 dpa of NSs con-structs in *N. benthamiana.* **B)** Silencing of GFP in the absence and presence of wild type (untagged) and tagged versions of NSs 5 dpa in *N. benthamiana.* The RNA silencing suppressor P19 from Tombusvirus was included as a positive control. **C)** Quantification of fluorescence units from 1 cm^2 leaf discs collected from agro-infiltrated leaf material 5 dpa panels shown under B. Error bars represent the SEM of six replicates. Differences between groups were determined with a One-Way ANOVA test with Šídák correction for multiple comparisons, α=0.05. Different letters indicate a significant difference from the GFP control with a p-value <0.001.

Figure 6.5 – Confocal microscopy images of *N. benthamiana* harvested 3 days post co-agro-infiltration. All experiments were repeated at least three times and the scale bars indicate 20 µm. **A)** Co-expression of SCYCE-NSs (left panel) or SCYCE-NSsW/G (right panel) with Catalase 1 C-terminally tagged with N-SCFP3A in a pDEST-GWSCYNE. A merged image of the CFP (cyan) fluorescent signal and brightfield is shown. **B)** Co-expression of SCYCE-NSs and SCYCE-Cat1, in the additional presence of peroxisome marker mRFP-SKL. A merged image of the CFP (cyan) and RFP (red) fluorescent signal. **C)** A Fluorescence intensity plot of CFP and RFP intensity along the depicted arrow.

not shown). In contrast, a CFP signal was observed concentrated in cytoplasmic bodies when SCYNE-Cat1 and SCYCE-NSs were co-expressed, confirming the occurrence of a genuine interaction between Cat1 and NSs (Figure 6.5A). Co-infiltration of SCYNE-Cat1 with SCYCE-NSs^WG also showed CFP fluorescence in similar bodies, eliminating a role of the WG motif in the interaction between Cat1 and NSs. A co-infiltration of the two BiFC constructs in the additional presence of mRFP-SKL confirmed a peroxisomal localization of Cat1 and NSs interactions (Figure 6.5B).

ROS production during viral infection and in the presence of transiently expressed NSs

Considering that Catalase 1 interacts with TSWV NSs and is one of various plant enzymes involved in the conversion of the reactive oxygen species (ROS) hydrogen peroxide (H_2O_2) produced during a (hypersensitive) defense response, the accumulation of ROS in infected and uninfected leaves was determined. Staining with 3,3'-diaminobenzidine (DAB) confirmed an increased accumulation of ROS in TSWV infected *N. benthamiana* leaves in comparison to uninfected leaf material (data not shown). Subsequently, levels of H_2O_2 and catalase activity were determined in mock and TSWV infected plants. Quantitative measurements of H_2O_2 were made in ground-up leaves of both locally and systemically, or mock, infected leaf 3, 7, and 14 dpi. In both local and systemic leaf material of TSWV infected plants, an upward trend could be discerned regarding the H_2O_2 concentration over time (Figure 6.6A). In comparison, the H_2O_2 levels in mock infected local leaves only rose very slightly at 7 dpi compared to 3 dpi, but at 14 dpi had decreased back to levels seen 3 dpi. The concentration in systemically mock infected leaves only rose very slightly at 14 dpi compared to H_2O_2 levels measured 3 and 7 dpi. Only at 14 dpi did the H_2O_2 levels of systemic and local TSWV infected leaf material differ significantly ($p<0.05$) compared to respective leaf material of mock infected plants.

Catalase activity is indicative for the amount of catalase enzyme present and was likewise determined in ground-up leaves of both locally and systemically (mock) infected leaf 3,

7, and 14 dpi. The catalase activity in mock infected plants showed similar trends over time in comparison to the H_2O_2 levels in the same type of material. Unlike the H_2O_2 levels, the catalase activity in leaves systemically infected with TSWV did not change over time. Measurements on local leaves showed less catalase activity at 7 dpi, than 3 dpi, while activity increased at 14 dpi. However, these differences were not significant. Only for leaves locally infected with TSWV at 14 dpi was there a significant ($p < 0.05$) increase compared to mock infected leaves at the same timepoint (Figure 6.6B).

To determine involvement of NSs in the increase of H_2O_2, NSs was agro-infiltrated in *N. benthamiana*. As, besides Catalase 1, multiple different enzymes are able to convert H_2O_2 into water and oxygen, similar experiments were also performed in which Catalase 1 was either overexpressed or silenced simultaneously to the agro-infiltration with NSs. The virus-induced gene silencing (VIGS) approach was implemented to knock down Catalase and subsequently analyze the effect on virus titers after a challenge with TSWV. To this end, a 300 bp fragment of Cat1 was cloned into pTRV2, and subsequently co-infiltrated with pTRV1 in *N. benthamiana*. Another viral protein reported to interact with catalase, 2b of Cucumber mosaic virus (CMV), (Inaba et al., 2011; Masuta et al., 2012; Murota et al., 2017), was taken along as a positive control. In plants without concurrent overexpression or silencing of Cat1, no significant differences were seen 3 dpa in H_2O_2 levels or catalase activity in ground up *N. benthamiana* leaf material agro-infiltrated with NSs or 2b compared to the control and GUS infiltrated plants (Figure 6.7A, B). Plants overexpressed with Cat1 also did not show a difference in H_2O_2 levels compared to control plants (Figure
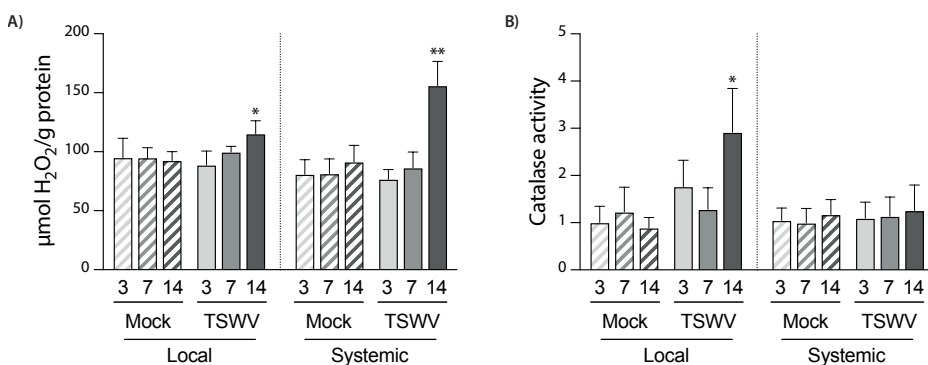
Figure 6.6 – Levels of H2O2 and catalase activity in leaf material locally or systemically infected with TSWV 3, 7, and 14 dpi. A) Quantitative measurements of hydrogen peroxide in local and systemic leaves of mock or TSWV infected 3, 7, and 14 days post infection. Hydrogen peroxide levels were measured as absorption at 390 nm in a spectrophotometer and calculated per g protein. **B)** Quantitative measurements of catalase activi-ty in local and systemic leaves of mock or TSWV infected 3, 7, and 14 dpi. Catalase activity was measured as decrease of absorption at 240 nm per minute per gram protein and normalized with the average of catalase activity at 3 dpi in mock local or systemic infected sample. Two-way ANOVAs with Šídák correction were performed for local and systemically infected leaves separately, α=0.05. Error bars depict SEM of nine samples from three independent experiments. * and ** indicates a significant difference (p<0.05 and p<0.01 respectively) from the mock sample at the same dpi in the same infected leaf type.

6.7B, C), but a significant (P<0.05) difference was observed when Cat1 was silenced (Figure 6.7B, D, E). However, when a Cat1-overexpression vector was co-infiltrated with NSs, 2b, or GUS these differences disappeared again (Figure 6.7B) Strikingly, in Cat1 silenced plants the H_2O_2 levels in leaves infiltrated with NSs or 2b did not rise in comparison to GUS silenced plants infiltrated with the same constructs (Figure 6.7D), while GUS or a mock infiltration in Cat1 silenced plants did lead to a H_2O_2 increase compared to GUS silenced or mock silenced plants.
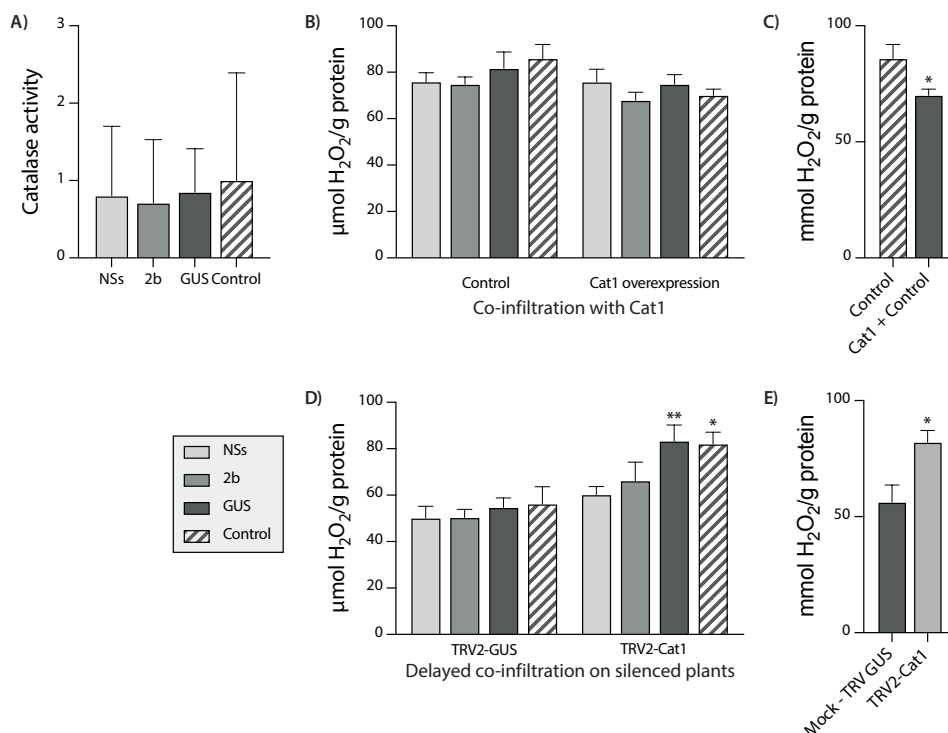
Figure 6.7 – Levels of H_2O_2 and catalase activity upon infiltration with RNA silencing suppressors in plants over-expressing Cat1, or with Cat1 silencing. A) Quantitative measurements of catalase activity in agro-infiltrated N. benthamiana leaves 3 dpa with NSs from TSWV, 2b from CMV, GUS and a control infiltration with MMAi. Catalase activity was measured as decrease of absorption at 240 nm per minute per gram protein and normalized with the average of catalase activity of the control sample. **B)** Quantitative measurements of hydrogen peroxide in co-agro-infiltrated *N. benthamiana* leaves 3 dpa with Cat1 over-expression or control infiltration and NSs from TSWV, 2b from CMV, GUS, control (MMAi). **C)** Quantitative measurements of hydrogen peroxide in agro-infiltrated Nicotiana benthamiana leaves 3 dpa with Cat1 overexpression or control infiltration. **D)** Quantitative measurements of hydrogen peroxide in agro-infil-trated N. benthamiana leaves 17 dpa with pTRV2-GUS or -Cat1 and 3 dpa with NSs from TSWV, 2b from CMV, GUS and a control infiltration with MMAi. Hydrogen peroxide levels were measured as absorption at 390 nm in a spectrophotometer and calculated per g protein. **E)** Quantitative measurements of hydrogen peroxide in agro-infiltrated *N. benthamiana* leaves 17 days post co-agro-infiltration with pTRV1 and pTRV2-GUS or -Cat1. One- and two-way ANOVA with Šídák correction were per-formed to compare between agro-constructs, α=0.05. Error bars depict SEM of twelve samples from two independent experiments. * indicates a significant difference (p<0.05) from the GUS sample.

Effect of Cat1 silencing on TSWV titers and systemic silencing by NSs in16C plants

Previous results indicated NSs and 2b play a role in lowering H_2O_2 levels, possibly through their respective interaction with Catalase. To further investigate the biological relevance of this interaction for the virus infection cycle, the aforementioned VIGS approach was implemented. Leaf samples were collected 14 dpa and assessed on Catalase transcript levels. Although Cat1 transcript levels were significantly lower ($p<0.001$) in TRV2-Cat1 infiltrated plants than in TRV2-GUS infiltrated plants, differences in the silencing of Catalase 1 were seen between experimental repeats. Upon closer examination, the fold change in Catalase 1 transcript levels in leaves collected was larger in leaves collected in summer (8.0x less Cat1 expression) compared to leaf material harvested in winter (1.7x less Cat1 expression) (Figure 6.8A).

Upon confirmation of the Catalase silencing, viral titers were determined in samples from the GUS and Cat1 silenced plants 14 days post infection (dpi) with TSWV, and thus 28 dpa with TRV constructs. Both Cat1 and GUS silenced plants challenged with TSWV showed a higher viral titer ($p<0.001$) than their respective control groups (Figure 6.8B). No significant difference was found in viral titers between TSWV challenged Cat1 and GUS silenced plants, nor was there a significant difference between unchallenged plants. However, when looking closer at the data, differences were found between samples depending on harvest seasons. Samples harvested in summer showed the same trends as described for the repetitions altogether. However, for material harvested in winter no difference was found in between infected and uninfected in Cat1 silenced, while virus titers of infected GUS silenced plants were still significantly ($p<0.01$) higher compared to uninfected plants. A k-means cluster analysis of the VIGS performed in winter showed that seven out of eleven TRV2-GUS silenced plants were considered infected (infected and starting infection), opposed to zero out of ten Cat1 silenced plants (Figure 6.8C, D). Altogether, these results indicated that Catalase 1 silencing seemed to slow down the local and/or systemic TSWV infection. Furthermore, stem length of Cat1 and GUS-silenced plants infected in summer (Figure 6.9) showed that Cat1 silenced plants were significantly taller ($p<0.05$) than GUS silenced plants for both TSWV infected and uninfected plants, although the slope of the difference between infected and uninfected plants was not affected by the silencing ($p>0.05$).

To determine whether the interaction between Cat1 and NSs influences systemic silencing, a pilot experiment was set up in constitutively GFP-expressing *N. benthamiana* 16C plants. To this end, these plants were first VIGS-silenced on Cat1 and 14 days later subjected to an infiltration with GFP on local leaves in the presence or absence of NSs constructs. As a negative control, 16C plants were VIGS silenced on GUS instead of Cat1. Systemic GFP silencing was determined by measuring GFP fluorescence in systemic leaves (L9, (Hedil et al., 2015)) 17 dpa. Infiltration with GUS serves as a negative control, and as expected did not show (much) systemic silencing of GFP in both GUS and Cat1
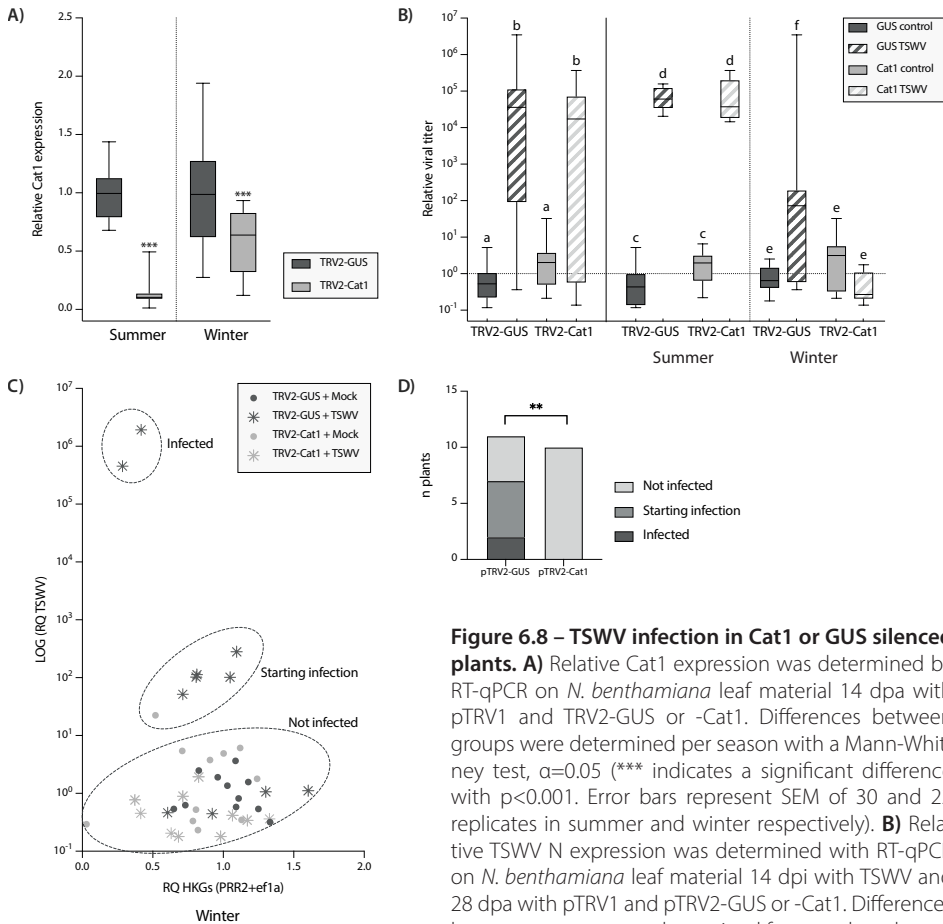
Figure 6.8 – TSWV infection in Cat1 or GUS silenced plants. A) Relative Cat1 expression was determined by RT-qPCR on *N. benthamiana* leaf material 14 dpa with pTRV1 and TRV2-GUS or -Cat1. Differences between groups were determined per season with a Mann-Whitney test, α=0.05 (*** indicates a significant difference with p<0.001. Error bars represent SEM of 30 and 22 replicates in summer and winter respectively). **B)** Relative TSWV N expression was determined with RT-qPCR on *N. benthamiana* leaf material 14 dpi with TSWV and 28 dpa with pTRV1 and pTRV2-GUS or -Cat1. Differences between groups were determined for complete dataset and per experimental repeat with a Two-Way ANOVA with a Tukey correction on Log transformed data, α=0.05 (different letters indicate a significant difference between groups (p < 0.001 for a-b, c-d; p < 0.05 for e-f). Error bars represent SEM of 26, 15, and 11 replicates in total, summer, and winter respectively). **C)** To break down and support the data seen in panel B), the Log expression of TSWV N against the combined expression of two housekeeping genes of the experiment performed in winter was used to differenti-ate these with K-means clustering into three different groups: not infected, starting infection and infected. **D)** A chi-share test was performed on VIGS plants infected with TSWV in winter, showing a significantly higher portion (** indicates p<0.01) of pTRV2-GUS infiltrated either infected or with a starting infection.
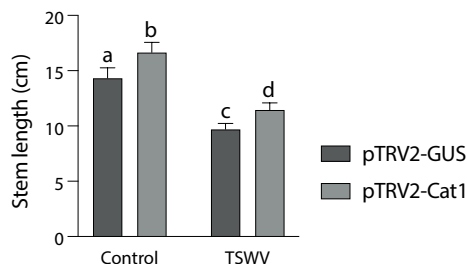
Figure 6.9 – Stem length in cm of *N. ben-thamiana* plants agro-infiltrated with pTRV2-GUS or -Cat1 and infected with TSWV. Stem length was determined at 28 dpa, 14 dpi (n=27). A two-way ANOVA with Tukey correction was performed, α=0.05. Differ-ent letters indi-cate a significant difference (p<0.05 for a-b, c-d; p<0.001 for a,b-c,d) between groups. Error bars indicate SEM.
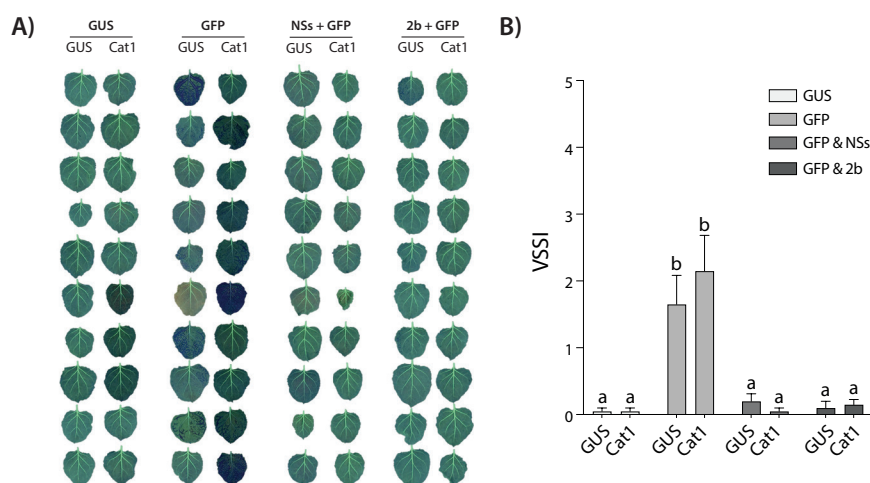
Figure 6.10 – Systemic GFP silencing in constitutively GFP-expressing *N. benthamiana* plants agro-infiltrated with pTRV2-GUS or -Cat1 and subsequently agro-infiltrated with pBIN-GFP, -NSs, -2b, and/or -GUS constructs. A) Images of the 9th true leaf (L9) taken 31 dpa with TRV constructs. **B)** Visually scored systemic silencing (VSSI) (Hedil et al., 2015) of GFP signal of aforementioned images. A two-way ANOVA with Tukey correction was performed, α=0.05. Different letters indicate a signifi-cant difference (p<0.01) between groups. Error bars indicate SEM.

silenced plants (Figure 6.10). Infiltration with GFP in GUS silenced plants triggered systemic silencing of GFP. Although not significant, compared to the GUS silenced plants, Cat1 silenced plants infiltrated with GFP seemed to show more systemic silencing. As expected, co-expression of GFP and RSS proteins NSs or 2b rescued GPF expression by blocking RNA silencing. No differences were found between GUS or Cat1 silenced plants.

Interaction of NSs with SIZ1, a nuclear residing SUMO-conjugating enzyme

Over the course of another research project, NSs was taken along as an assumed negative control during BiFC experiments to confirm interactions between a host gene and the *Solanum tuberosum* SUMO E3 ligase SIZ1. Surprisingly, two independent experiments consistently revealed an interaction between NSs and SIZ1 in the nucleus, where a CFP signal was observed dispersed in the nucleus and in nuclear bodies, but not the nucleolus (Figure 6.11). A nuclear localization of TSWV NSs has not been observed before and the relevance of this finding remains to be further investigated.

Discussion

Various attempts have been made to elucidate the molecular mechanisms behind TSWV pathogenicity by finding host factors necessary for TSWV infection, replication, or movement (Ramesh et al., 2017; Helderman et al., 2021, 2022; Zhai et al., 2021; Zhan et al., 2021). In this study, we aimed to identify host proteins interacting with NSs to shed more light on the interplay between TSWV NSs and its plant host, and the mechanistic details behind these interactions that might influence a viral infection. Considering the role of
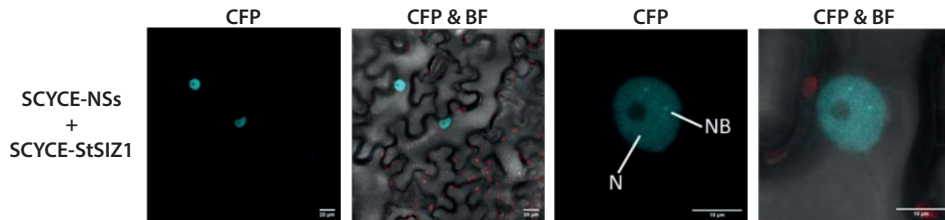
Figure 6.11 – Confocal microscopy images of *N. benthamiana* harvested 2 days post co-agro-infiltration. Co-expression of the C-terminal half of the super cyan fluorescent protein SCFP3A tagged NSs (SCYCE-NSs) with the N-terminal tagged Solanum tuberosum SIZ1 (SCYNE-StSIZ1). Scale bars indicate 10 or 20 µm.

TSWV NSs in the suppression of antiviral RNAi and as effector of *Tsw*-mediated resistance, host proteins identified to interact might add to our fundamental understanding of these processes. From the list of host proteins collected by a co-IP and subsequent mass spectrometry, several interesting candidate proteins were further investigated, from which Catalase was confirmed to genuinely interact with NSs in peroxisomes in the end. Unexpectedly, the list of interaction candidates of both NSs and NSs[WG] mutant did not include AGO proteins, despite earlier experimental evidence pointing to this (de Ronde et al., 2014b). Nor did SIZ1, a Sumo E3 ligase from *S. tuberosum* that unexpectedly and interestingly was observed to interact with NSs in the nucleus.

The absence of SIZ1 and AGO1 from the list of host proteins interacting with NSs supports the idea that the co-IP approach with NSs as bait is perhaps not sensitive enough. In recent years proximity labelling via biotin ligase ((mini)TurboID) has shown to be a more powerful and sensitive tool to identify weaker and transient interactions between proteins that do not show up via the custom co-IP and MS analysis approach. In addition, during another study to identify host interactors of TSWV NSs (Zhai et al., 2021), several potentially NSs-interacting host proteins were identified that did not show up in our study, or vice versa. This could suggest that timing and sampling of tissue may also be of importance and partially affect the outcome. This issue would also be circumvented when implementing proximity labeling, as any host protein in vicinity of NSs at any time point during the course of NSs expression will become biotinylated, and thereby should show up in a subsequent MSs analysis of biotinylated proteins.

The interaction between SIZ1 and NSs was not only unexpected but also very surprising and interesting. A nuclear localization for TSWV NSs has never been reported before, nor for any other tospovirus NSs protein except for Capsicum chlorosis virus (CaCV) NSs. In a study by Widana Gamaze and Dietzgen (2017) transient expression of a CaCV GFP-NSs fusion protein revealed a slightly punctate distribution in the periphery of cells (likely cytoplasmic aggregates as reported earlier (Kormelink et al., 1991)) and in the nucleus, but not the nucleolus. Further analysis with BiFC revealed that this CaCV NSs homodimerized, and that these self-interactions exclusively took place in the nucleus. The localization of

TSWV NSs in the nucleus, in addition to CaCV NSs, provides stronger support for a role of tospovirus NSs in the nucleus. The interaction of NSs with SIZ1 points towards a possible interaction of NSs with sumoylated proteins or whether NSs itself can be sumoylated. A prediction analysis was performed for the NSs proteins of TSWV and CaCV. Four SUMO attachment sites (SAS) and no SUMO interaction motif (SIM) were predicted for TSWV NSs, whereas for CaCV NSs only two SAS were predicted and one SIM with a low prediction score (Table S6.3). One of the SAS sites seems to be shared by both NSs variants. However, further investigation will be needed to determine the relevance of these sites. While not uncovered in our research, nor in that of Helderman et al. (2021), NSs has been found to interact with Importin subunit α (Zhai et al., 2021), which might facilitate the transport of NSs into the nucleus, enabling the nuclear-localized interaction with StSIZ1.

Although interaction of TSWV NSs and SIZ1 needs to be further analyzed and the functional relevance of tospovirus NSs in the nucleus is unclear and raises questions for further research, the NSs from several vertebrate-infecting bunyaviruses have been well studied on their Interferon-antagonistic activities in which the NSs protein from e.g. Bunyamwera, La Crosse and Rift Valley fever virus (RVFV) has been shown to inhibit the interferon-induced defense responses by blocking nuclear RNA polymerase II transcription (Weber et al., 2002; Billecocq et al., 2004; Thomas et al., 2004). While some of these NSs proteins were found to form cytoplasmic inclusions, numerous papers have reported and studied the nuclear localization of RVFV NSs (Swanepoel and Blackburn, 1977; Cifuentes-Muñoz et al., 2014; Cyr et al., 2015; Li et al., 2019b; Léger et al., 2020).

As the interaction between Cat1 and NSs or NSs[WG] was comparable, only NSs was used in further experiments. As one of three catalases present in *N. benthamiana*, Catalase 1, a class I catalase mainly expressed in photosynthetic tissue, catalyzes the reduction of hydrogen peroxide (H_2O_2) into water and oxygen. H_2O_2 is a damaging reactive oxygen species (ROS) that plays an important role in the response to pathogens and in cell metabolism. Our study showed that TSWV infection is associated with increased H_2O_2 levels and higher catalase activity in *N. benthamiana*. In low Cat1 expressing plants, transient expression of NSs leads to a reduced, baseline, level of H_2O_2. In contrast, reduction of Cat1 expression also leads to a slower infection rate in plants with a lowered metabolism. Altogether, these results support the idea of Cat1 as a potential susceptibility factor for TSWV.

The role of H_2O_2 in immunity is two-fold. High levels of H_2O_2 are known to trigger a hypersensitive response (HR) resulting in cell death and thus blocking viral spread. Furthermore, reactive oxygen species like H_2O_2 act as a signaling molecule for expression and activation of ROS scavenging enzymes, allowing for recovery after for instance a virus infection. Accumulation of H_2O_2 is undesirable for a virus, while ensuring that catalases continue to break down this component into water and oxygen is advantageous. Nevertheless, the effect of interactions between catalases and viral proteins varies between pathogens. In contrast to our findings regarding the interaction between Cat1

and NSs, the interaction between the RSS of Chilli veinal mottle virus, HCPro, with Cat1 seems beneficial for the plant, as overexpression of Cat1 led to a decrease of Chilli veinal mottle virus titers while a knockout of Cat1 resulted in increased titers in *N. tabacum* (Yang et al., 2020). Interaction with the precoat protein 2AV of the Gemini virus Tomato leaf curl palampur virus with NbCat2 (class II, mainly expressed in vascular tissue) is similarly beneficial for the plant (Roshan et al., 2018). On the other hand, interaction between the RSS protein 2b of CMV with *A. thaliana* Cat3 (class II) inactivated the catalase and induced a programmed cell death that, surprisingly, benefited viral infection (Inaba et al., 2011; Masuta et al., 2012; Murota et al., 2017). The interaction between possible movement protein TGBp1 (p26) of Pepino mosaic virus (PepMV) and tomato CAT1 (class II) most likely benefits the virus, whereas silencing led to decreased viral titers (Mathioudakis et al., 2013), much alike the effects of the Cat1-NSs interaction. Decreased Cat1 expression showed that the presence of NSs results in lower H_2O_2 levels, possibly by enhancing the reduction of H_2O_2 by catalase. "Neutral" levels of H_2O_2 will not trigger downstream immune pathways, thereby benefiting the virus. However, low catalase expression levels are also associated with no or lower TSWV titers 14 dpi in plants with slower metabolism. As viral accumulation leads to increased ROS production, the low amount of catalase proteins present is possibly no longer able to prevent H_2O_2 levels from rising to levels that trigger immune pathways. Unlike the programmed cell death triggered by 2b-AtCat3 interaction, the possible activation of the immune system seems to benefit the plant rather than the virus. Whether the mechanisms to slow down the TSWV infection rate in Cat1 silenced plants are also found in situations with normal metabolism rates is to be investigated by comparing viral titers at earlier time points in such a metabolic setting, nor is it known whether NSs is able to interact with other catalase variants similar to HCPro (Yang et al., 2020). Silencing experiments in constitutively GFP expressing 16C plants showed that the interaction with Cat1 does not seem to influence the suppression of systemic silencing by either 2b or NSs, although silencing of Cat1 by itself does seem to enhance GFP silencing in systemic leaves. Although speculative, whether Cat1 acts as an endogenous suppressor of RNAi and thereby relieves pressures on TSWV replication remains to be investigated.

6

The effect of lowered Cat1 expression in *N. benthamiana* on catalase activity is also unknown. The method used in this research to determine catalase activity has been used in many other papers, however results proved difficult to replicate and findings were accompanied by large standard deviations, causing questions to arise regarding the reliability of this method (Anjum et al., 2016). Rather than measuring absorption decrease over time for determination of catalase activity (Aebi, 1984), a method which employs a stop solution followed by calorimetric assessment, which allows for a more clean-cut comparison of catalase levels in different samples (Abcam, 2021), would be preferable.

The downstream consequences and the nature of the direct interaction between Cat1 and NSs requires more research, using for instance currently not yet uncovered mutants of NSs not able to interact with Catalase 1, as well as future studies in TSWV resistant

Capsicum sp. to uncover the potential role of *Tsw* in this interaction. This will possibly also indicate whether the ability of NSs to interact with Catalase is linked to some of its other functions, or regions associated with resistance breaking strains of NSs.

The LC-MS studies included proteins pulled down with a wildtype NSs as well as the NSs[WG] mutant. Although putative interactions between AGO1 and NSs had been indicated, in part by the presence of the WG motif (de Ronde et al., 2014b), no Argonaute proteins were found. The unintentional dissipation of weak and transiently associated proteins or low affinity binding is one of the drawbacks of co-IP (Yang et al., 2021), which could explain the absence of AGO1 among the putatively interacting proteins. Of these proteins, the bulk was enriched for both NSs and NSs[WG], including catalase. As localization and BiFC assays showed similar results for both RSS proteins, subsequent experiments only included wildtype NSs.

The list of proteins indicated to potentially interact with NSs through the LC-MS contains multiple potential leads to examine the interactome of TSWV NSs, in addition to the confirmed interaction with the probable pro-viral factor Catalase 1. One of the identified potential interaction proteins was eukaryotic translation elongation factor 1 alpha, earlier shown to be required for transcription and translation of TSWV (Komoda et al., 2014) and of which an isoform (eEF1A, NbS00023178g0001.1, Niben101Scf07423g04011.1)), found in viral RNPs, was demonstrated to have a pro-viral role (Helderman et al., 2021). In the present study, interaction of Ef1α isoform 2 with NSs was not confirmed via BiFC assay. In line with the data of Helderman et al. (2021), subsequent silencing of this Ef1α isoform resulted in such severe stunting and chlorosis (data not shown) that no plant material could be harvested 14 dpi and was therefore excluded from further investigation. Although a direct interaction between Ef1α isoform A1 and NSs could not be confirmed, it remains interesting to note that a previous study has indicated a translational enhancing role of TSWV NSs (Geerts-Dimatriadou et al., 2012). On this point, the NSs proteins of several animal-infecting bunyaviruses have also been implied to play a role in translation, in which for the Uukuniemi NSs protein even an association with 40S ribosomes was observed (Simons et al., 1992). Although speculative, whether a translational enhancement by TSWV NSs results from transient interactions with Ef1α, to support scanning and avoid premature transcription/translation, a phenomenon observed with several bunyaviruses, remains to be investigated.

Besides Ef1α and catalase, several other proteins found to putatively interact with NSs via our research overlap with those of two studies relatively similar to ours (Helderman et al., 2021; Zhai et al., 2021). These include various heat shock (related) proteins -70 and -90 (Niben101Scf00117g02019.1, Niben101Scf10036g00003.1, Niben101Scf12868g00008.1), chaperone proteins known to interact with and influence infection of several plant viruses (Moshe et al., 2016; Qian et al., 2018; Berka et al., 2022), chloroplastic carbonic anhydrase (Niben101Scf06349g00008.1) found to be involved with replication of

Bamboo mosaic virus in *N. benthamiana* (Chen et al., 2017), as well as different chlorophyll a-b binding proteins (Niben101Scf01328g01012.1, Niben101Scf04318g01016.1, Niben101Scf02513g08003.1). Furthermore, variants of tRNA ligase protein (Niben101Scf01818g05006.1) are found in all three studies, as well as proteins associated with the 40S ribosomal complex (Niben101Scf02353g02004.1) (Helderman et al., 2022). Both are protein groups involved in host RNA translation and discovered to be involved in virus translation as well (Nohales et al., 2012; Helderman et al., 2022). Apart from proteins overlapping with findings from other research, our list of putatively interacting proteins also holds unique proteins that provide interesting leads to determine their influence on TSWV infection, including poly(A)-binding protein (Niben101Scf04007g02013.1), another protein involved in translation which aids in the addition of poly(A) tail to create mature mRNA known to influence viral infection (Dufresne et al., 2008), as well as phosphoglycerate kinase (Niben101Scf02461g00003.1), an ATP generating kinase acting as a pro-viral factor for Tomato bushy stunt virus (Prasanth et al., 2017). Multiple other proteins are involved in management of (oxidative) stress responses, e.g. aconitase (Niben101Scf00859g01014.1), isocitrate dehydrogenase (Niben101Scf01789g02047.1), and Thioredoxin H3 (Niben101Scf00539g07015.1).

Agro-infiltration of gene constructs of viral proteins in search of its interacting proteins will never truly replicate the situation of a viral infection. While the use of for instance viral RNPs better reflects the natural infection state (Helderman et al., 2021), when looking at accessory, non-structural proteins such as NSs this method may not be entirely appropriate. After all, only expression of such a protein can lead to changes in the transcriptome, which may not reflect the change caused by a viral infection. In light of this, analysis of transcriptome changes upon transient expression of viral RSS using cDNA-amplified fragment length polymorphism (cDNA-AFLP) have earlier shown that expression of NSs highly increased transcription of HSP70 (Table S6.2, (Bucher, 2006)). Other proteins found using this technique also overlapped with the more recent studies, such as aconitase, thioredoxin, dnaJ (HSP40), and a calmodulin-binding protein. Apart from the interaction of *N benthamiana* Calmodulin 3 with NSs (Zhai et al., 2021), the calcium-signaling calmodulins have been shown to interact with other viral proteins as well, to both the detriment and the benefit of the virus (Nakahara et al., 2012; Li et al., 2014). Other studies have shown that thrips attacks on plants are associated with increased Ca^{2+}, which in turn also activates calmodulins, triggering cell death through the mitogen activated protein kinase (MAPK) pathways (Takahashi et al., 2011). Furthermore, a Calmodulin-binding protein in *A. thaliana* has been shown to interact with Catalase2, thereby promoting CAT2 enzyme activity, indirectly increasing JA content and enhancing resistance against *Botrytis cinerea* (Lv et al., 2019). All these studies indicate that the interaction of NSs with Cat1 and Calmodulin 3 could both possibly target the same MAPK signaling cascade towards apoptosis and thereby prevent programmed cell death, acting as susceptibility factors in TSWV infection. The ALFP (Table S6.2, (Bucher, 2006)) furthermore showed enrichment of an E3 ubiquitin protein ligase, in addition to

6

a polyubiquitin. Interestingly, ubiquitination of AtCAT3 by the host proteasome pathway has been shown to be (partly) responsible for degradation of AtCAT3 during CMV infection. The interaction with 2b of CMV removes AtCAT3 from the peroxisomes to the nucleus. While the interaction between Cat1 and NSs does not move Cat1 to the nucleus, it is unknown whether similar pathways are triggered for NSs, and whether this is related to the interaction of NSs with StSIZ1. Although this unexpected finding still needs further confirmation, this interaction not only opens new avenues for studies on a nuclear role of TSWV NSs, but its absence from the applied co-IP approach strengthens a repeated analysis by implementing the more powerful and sensitive proximity labelling.

Table S6.1 – Proteins from Nictotiana benthamiana potentially interacting with tomato spotted wilt virus (TSWV) NSs revealed by co-immunoprecipitation and mass spectrometry. The Gene ID was based on the N. benthamiana 1.0.1 database, its name, as well as the BLAST match in Solanum lycopersicum and Arabidopsis thaliana. "Up" indicates enriched proteins in the GFP-NSs or GFP-NSsWG samples, "Down" indicates enrichment in the GFP fraction (column 1 and 2). If a significant difference was found between NSs and NSsWG, the samples in which the protein was (more) enriched is indicated (column 3).

NSs	NSs^WG	NSs^WG vs NSs	N. benthamiana		S. lycopersicum BLAST match	A. thaliana BLAST match
			Database match	Name		
	Up	NSs^WG	Niben101Ctg12033g00001.1	Peroxiredoxin-2B	Solyc07g020860.2.1	AT1G65980
Up	Up		Niben101Scf02750g03012.1	ATP synthase gamma chain	Solyc03g115110.2.1	AT2G33040
Up	Up		Niben101Scf00117g02019.1	Chaperone protein dnaK	Solyc01g103450.2.1	AT5G49910
Up	Up		Niben101Scf16939g00010.1	Serine/threonine-protein phosphatase 4 regulatory subunit 1	Solyc05g009600.2.1	AT3G25800
Up		NSs	Niben101Scf00202g05007.1	Diaminopimelate decarboxylase	Solyc01g109850.2.1	AT3G14390
Up	Up		Niben101Scf04473g07001.1	ATP-dependent Clp protease proteolytic subunit	Solyc01g100520.2.1	AT1G02560
Up	Up		Niben101Scf02353g02004.1	40S ribosomal protein SA	Solyc06g072120.2.1	AT3G04770
Up	Up		Niben101Scf00367g02013.1	Plastid lipid-associated protein 2, chloroplastic	Solyc09g090330.2.1	AT3G23400
Up	Up		Niben101Scf00369g03018.1	Histone H2A	Solyc01g099410.2.1	AT5G27670
Up	Up		Niben101Scf02182g02007.1	Histone H2B	Solyc03g071620.1.1	AT3G45980
Up	Up		Niben101Scf00506g02003.1	Protochlorophyllide reductase	Solyc10g006900.2.1	AT4G27440
Up	Up		Niben101Scf00539g07015.1	Thioredoxin H3	Solyc04g081970.2.1	AT1G76080
	Up	NSs^WG	Niben101Scf00654g04006.1	Proteasome activator subunit 4	Solyc03g019780.2.1	AT3G13330
Up	Up	NSs^WG	Niben101Scf00712g15023.1	Catalase	Solyc04g082460.2.1	AT4G35090
Up	Up		Niben101Scf01817g02001.1	V-type proton atpase subunit E	Solyc08g081910.2.1	AT4G11150

NSs	NSs^WG	NSs^WG vs NSs	*N. benthamiana* Database match	Name	*S. lycopersicum* BLAST match	*A. thaliana* BLAST match
Up	Up		Niben101Scf00859g01014.1	3-isopropylmalate dehydratase large subunit	Solyc07g052350.2.1	AT2G05710
Up	Up		Niben101Scf00872g00007.1	Fumarate hydratase class II	Solyc09g075450.2.1	AT2G47510
Up	Up		Niben101Scf00978g08004.1	Glycogen phosphorylase 1	Solyc09g031970.2.1	AT3G46970
Up	Up		Niben101Scf01142g00013.1	Alpha/beta-Hydrolases superfamily protein	Solyc06g068220.2.1	AT1G52510
Up	Up		Niben101Scf01319g04016.1	Tubulin beta-6 chain	Solyc04g081490.2.1	AT1G75780
Up	Up		Niben101Scf01328g01012.1	Chlorophyll a-b binding protein 8	Solyc12g011280.1.1	AT1G61520
Up	Up		Niben101Scf01350g03009.1	D-3-phosphoglycerate dehydrogenase	Solyc03g112070.2.1	AT4G34200
Up			Niben101Scf01372g00003.1	Peroxiredoxin-2	Solyc01g007740.2.1	AT5G06290
Up	Up		Niben101Scf01399g00012.1	ATP synthase gamma chain	Solyc02g080540.1.1	AT4G04640
Up	Up		Niben101Scf01436g03010.1	30S ribosomal protein S5	Solyc02g077990.2.1	AT2G33800
Up	Up		Niben101Scf04318g01016.1	Chlorophyll a-b binding protein 37	Solyc07g047850.2.1	AT3G27690
Up	Up		Niben101Scf01552g05010.1	Elongation factor 1-gamma 2	Solyc11g028100.1	AT1G09640
	Up	NSs^WG	Niben101Scf07984g02015.1	Glucose-1-phosphate adenylyltransfer-ase family protein	Solyc07g056140.3	AT5G48300
Up		NSs	Niben101Scf01696g06050.1	ADP-L-glycero-D-manno-heptose-6-epimerase	Solyc01g097340.2.1	AT5G28840
Up	Up		Niben101Scf01789g02047.1	3-isopropylmalate dehydrogenase	Solyc08g077930.2.1	AT5G03290
Up	Up		Niben101Scf01818g05006.1	Glycine—tRNA ligase	Solyc11g039830.1.1	AT1G29880
Up	Up		Niben101Scf01822g16001.1	Conserved hypothetical protein	Solyc08g076450.2.1	AT1G32220

Up			Niben101Scf05414g00004.1	Alpha-1,4-glucan-protein synthase 1	Solyc05g012070.2.1	AT3G02230
Down	Down		Niben101Scf02240g02008.1	Acyl-coa-binding protein	Solyc08g075690.2.1	AT1G31812
	Up		Niben101Scf14799g02001.1	Glycogen phosphorylase 2	Solyc03g065340.2.1	AT3G29320
Up			Niben101Scf02346g06007.1	ATP-dependent Clp protease proteolytic subunit	Solyc08g075750.2.1	AT5G45390
Up	Up		Niben101Scf02461g00003.1	Phosphoglycerate kinase	Solyc07g066600.2	AT1G79550
Up	Up		Niben101Scf02479g02013.1	Fructose-1,6-bisphosphatase class 1	Solyc10g086730.1.1	AT3G54050
Up			Niben101Scf02513g08003.1	Chlorophyll a-b binding protein, chloroplastic	Solyc01g105030.2.1	AT1G15820
Up	Up		Niben101Scf02553g02002.1	Glutamate-1-semialdehyde 2,1-aminomutase 2	Solyc04g009200.2.1	AT3G48730
	Up		Niben101Scf02698g00008.1	Glutamate decarboxylase	Solyc03g098240.3	AT5G17330
	Up		Niben101Scf02824g03017.1	Chaperone protein htpg family protein	Solyc05g010670.2.1	AT2G04030
Up	Up		Niben101Scf02917g00004.1	Cell division protein ftsz	Solyc09g009430.2.1	AT3G52750
Up	Up		Niben101Scf03068g00011.1	ATP synthase subunit beta	Solyc01g007320.2	ATCG00470
Up	Up		Niben101Scf03070g08001.1	Galactokinase	Solyc01g058390.2.1	AT3G06580
Up	Up		Niben101Scf03455g05017.1	3-oxoacyl-[acyl-carrier-protein] synthase 2	Solyc02g070790.2.1	AT5G46290
Up	Up		Niben101Scf03579g02007.1	Delta subunit of Mt ATP synthase	Solyc01g087120.2.1	AT5G13450
	Up	NSs[WG]	Niben101Scf03500g02019.1	Fructose-1,6-bisphosphatase class 1 1	Solyc05g052600.2	AT3G55800
Up	Up		Niben101Scf03572g02005.1	T-complex protein 1 subunit theta	Solyc01g088080.2.1	AT3G03960
Up	Up		Niben101Scf03628g14016.1	Ketol-acid reductoisomerase	Solyc07g053280.2.1	AT3G58610
Up	Up		Niben101Scf03819g01016.1	UDP-glucose 6-dehydrogenase 3	Solyc02g067080.2.1	AT5G15490
Up	Up		Niben101Scf04007g02013.1	Polyadenylate-binding protein 4-like	Solyc12g088720.1.1	AT4G34110

NSs	NSs^{WG}	NSs^{WG} vs NSs	*N. benthamiana*		*S. lycopersicum* BLAST match	*A. thaliana* BLAST match
			Database match	Name		
Up	Up		Niben101Scf04133g01027.1	Cystathionine gamma-synthase	Solyc02g067180.2.1	AT3G01120
	Up	NSs^{WG}	Niben101Scf10834g00005.1	Asparagine—trna ligase	Solyc05g056250.2.1	AT4G31180
Up	Up		Niben101Scf04639g06007.1	Elongation factor 1-alpha	Solyc11g069700.1.1	AT1G07940
Up	Up		Niben101Scf04673g00004.1	6-phosphogluconate dehydrogenase	Solyc12g056120.1.1	AT5G41670
Up	Up		Niben101Scf04958g02014.1	UPF0061 protein ydiu	Solyc12g100200.1.1	AT5G13030
Up	Up		Niben101Scf05437g03003.1	Elongation factor Tu	Solyc03g112150.1.1	AT4G20360
Up	Up		Niben101Scf06117g01005.1	Protochlorophyllide reductase, chloroplastic	Solyc07g054210.2.1	AT5G54190
Up	Up		Niben101Scf06221g03002.1	ATP synthase subunit alpha	Solyc00g042130.1	ATMG01190
Up	Up		Niben101Scf06349g00008.1	Carbonic anhydrase	Solyc02g086820.2.1	AT3G01500
Up	Up		Niben101Scf06382g00006.1	Glutamate-1-semialdehyde 2,1-aminomutase 2	Solyc04g009200.3	AT3G48730
	Up		Niben101Scf06654g02002.1	Protein thf1	Solyc07g054820.2	AT2G20890
Down			Niben101Scf06825g04015.1	Alpha-1,4-glucan-protein synthase	Solyc03g019750.2.1	AT3G02230
Up	Up		Niben101Scf06914g00003.1	Pyruvate dehydrogenase E1 component subunit beta	Solyc06g072580.2.1	AT5G50850
Up	Up		Niben101Scf07182g04010.1	Formate—tetrahydrofolate ligase	Solyc01g006280.2.1	THFS
	Up		Niben101Scf08691g00013.1	RNA-binding protein 28	Solyc09g007850.2.1	AT2G37220
	Up		Niben101Scf08691g00013.1	RNA-binding protein 28	Solyc09g007850.2.1	AT2G37220
Up			Niben101Scf08721g01033.1	Phenazine biosynthesis phzc/phzf protein	Solyc09g064940.2.1	AT4G02860

Up	Up		Niben101Scf08892g00008.1	2-isopropylmalate synthase 1	Solyc06g053400.2.1	AT1G74040
Down		NSs[WG]	Niben101Scf09089g01026.1	Unknown chloroplastic protein	Solyc09g010120.2.1	AT2G36145
Up	Up		Niben101Scf09107g01004.1	Alpha-glucan phosphorylase 1	Solyc05g012510.2.1	AT3G29320
Up			Niben101Scf09883g01010.1	NAD(P)-binding Rossmann-fold super-family protein	Solyc04g082780.2.1	AT2G33600
Up	Up		Niben101Scf10036g00003.1	Heat shock 70 kda protein 2	Solyc12g043110.1.1	F4HQD4
Up			Niben101Scf10162g00004.1	Alanine aminotransferase 2	Solyc05g013380.2.1	AT1G70580
Up	Up		Niben101Scf10189g02009.1	3-ketoacyl-coa thiolase	Solyc09g091470.2.1	AT2G33150
Up			Niben101Scf11724g01006.1	Thiamine thiazole synthase	Solyc07g064160.2.1	AT5G54770
Up	Up		Niben101Scf12589g00003.1	Ubiquitin carboxyl-terminal hydrolase 34	Solyc10g081610.1.1	AT5G06600
Up			Niben101Scf12868g00008.1	Heat shock 70 kda protein 1A/1B	Solyc11g066060.1.1	AT5G02500
Up	Up		Niben101Scf15372g00003.1	ATP synthase delta-subunit gene	Solyc12g056830.1.1	AT4G09650
Up			Niben101Scf15752g00007.1	GTP binding Elongation factor Tu family protein	Solyc03g058190.2.1	AT5G13650
	Up		Niben101Scf16508g00001.1	Sulfate adenylyltransferase	Solyc03g005260.2.1	AT1G19920
Up	Up		Niben101Scf17815g00007.1	Alpha-1,4-glucan-protein synthase 2	Solyc04g005340.2	AT5G15650
Up	Up		Niben101Scf18107g00014.1	D-glycerate 3-kinase	Solyc03g120430.2.1	AT1G80380

Table S6.2 – Annotated summary of *Nicotiana benthamiana* cDNA-amplified fragment length polymorphism (cDNA-AFLP) fragments with upregulated expression due to the activity of NSs. Up-regulations higher then 500 times indicate that the band was over-exposed.

BLAST Hit	Upregulated
NSs	1000
Luminal binding protein (HSP70 family)	393
SAS protein 10	210
Mitochondrial import receptor subunit TOM6 homolog	42
GFP	22
Polyubiqitin	22
Translocase subunit SCY2	20
GFP	8
Exosome complex component RRP45A-like	6
Dynein light chain	5
deoxyribodipyrimidine photo-lyase	5
actin-depolymerizing factor	5
potassium transporter	4
e3 ubiquitin protein ligase RGLG2 like	4
TNF receptor-associated factor homolog 1b-like	4
dnaJ protein	4
chloroplast stem-loop binding protein of 41 kDa b	4
thioredoxin	3
aconitate hydratase	2
Bifunctional dihydrofolate reductase-thymidylate synthase-like	2
calmodulin-binding protein 60 D-like	2
DNA polymerase epsilon catalytic subunit A-like	2
aspartyl protease family protein 2-like	2
cell division cycle protein 27 homolog B	2
phosphoinositide phospholipase C 4-like	1
NbTPRb mRNA for nuclear pore complex protein TPRb	1
cyclin-dependent kinase inhibitor 7-like	1

Name	Sequence	Purpose
P1_Rv_NotI	GCGGCCGCATAGAATTGAATCTGCTTAAGTCT	Cloning GFP P1 construct
2B_Fw_B1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATG-GAATTGAACGAAGGCGC	Gateway sites 2b
2B_Rv_B2_	GGGGACCACTTTGTACAAGAAAGCTGGGTCT-CAAAACGACCCTTCCGCCC	Gateway sites 2b
2B_Rv_B2_ns	GGGGACCACTTTGTACAAGAAAGCTGGGTCAAAC-GACCCTTCCGCCCAC	Gateway sites 2b w/o stop codon
NSs_Fw_B1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAT-GTCTTCAAGTGTTTATGA	Gateway sites NSs
NSs_Rv_B2	GGGGACCACTTTGTACAAGAAAGCTGGGTCT-CATTTTGATCCTGAAGCATATG	Gateway sites NSs
NSs_Rv_B2_ns	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTTT-GATCCTGAAGCATATG	Gateway sites NSs w/o stop codon
P1_Fw_B1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGG-GGAAATCCAAACTCAC	Gateway sites P1
P1_Rv_B2	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTA-ATAGAATTGTATCTGTTTAAGTTTAC	Gateway sites P1
P1_Rv_B2_ns	GGGGACCACTTTGTACAAGAAAGCTGGGT-CATAGAATTGTATCTGTTTAAGTTTAC	Gateway sites P1 w/o stop codon
AGO1_AttB1B2_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGGT-GAGAAAGAGAAG	Gateway sites AGO1
AGO1_AttB1B2_R	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAG-CAGTAGAACATGAC	Gateway sites AGO1
Cat1_AttB1_Fw	GGGGACCACTTTGTACAAGAAAGCTGGGTCT	Gateway sites Cat1
Cat1_AttB2_Rv	TATGCAACAGAAGCCATATGG	Gateway sites Cat1
Cat1_Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCA	Cloning Cat1
Cat1_Rv	ATGGCATCTGACAAGAAGATC	Cloning Cat1
ef1a_AttB1_Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCA	Gateway sites Ef1a
ef1a_AttB2_Rv	ATGGGTAAAGAGAAGGTTCAC	Gateway sites Ef1a
ef1a_Fw	ATGGGTAAAGAGAAGGTTCACATCAATATCG	Cloning Ef1a
ef1a_Rv	TCACTTTTTCTTCTGGGCGGCC	Cloning Ef1a
GAPC1_AttB1_Fw	GGGGACCACTTTGTACAAGAAAGCTGGGTCT	Gateway sites GAPC1
GAPC1_AttB2_Rv	CACTTTTTCTTCTGGGCGGCC	Gateway sites GAPC1
GAPC1_Fw	ATGGCATCTGACAAGAAGATCAAGATCGG	Cloning GAPC1
GAPC1_Rv	TTATGCAACAGAAGCCATATGGCAG	Cloning GAPC1
pDONR207_Fw	CGCGTTAACGCTAGCATGGATC	Cloning into pDONR207
pDONR207_Rv	GTAACATCAGAGATTTTGAGACAC	Cloning into pDONR207

Name	Sequence	Purpose
Cat1_VIGS_AttB1_Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGAG-GAGATCGACTACTTCCCTTC	Cloning Cat1 TRV/VIGS
Cat1_VIGS_AttB2_Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTcTTAAG-CCTAGAAGCAAGCTTTTGACC	Cloning Cat1 TRV/VIGS
ef1a_VIGS_AttB1_Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAT-CACCTGATTTACAAGCTTGGTGG	Cloning Ef1a TRV/VIGS
ef1a_VIGS_AttB2_Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTc-GCTTCAAAACCACCAGTGGTGG	Cloning Ef1a TRV/VIGS
Cat1_qPCR_Fw	AGGTACCGCTCATTCACACC	RT-qPCR Ef1a
Cat1_qPCR_Rv	GATCAGACAAGGCCTCCACC	RT-qPCR Ef1a

6

Chapter 7

General discussion

I.L. van Grinsven

Introduction

My thesis focuses on closing the knowledge gap concerning the interaction between the pepper resistance (*R*) gene *Tsw* and the virus, Tomato spotted wilt virus (TSWV), to which this gene confers resistance. TSWV is considered one of the most economically devastating plant viruses, in part due to its broad host range (Parrella et al., 2003; Scholthof et al., 2011), as well as the difficulty to eradicate its insect vector *Frankliniella occidentalis* from fields and greenhouses (Gilbertson et al., 2015; Reitz et al., 2020). Tsw (in)directly recognizes the RNA silencing suppressor (RSS) from TSWV, NSs (in-depth overviews of plant immunity, and in resistance against TSWV are given in Chapter 1 and 2, respectively). However, resistance breaking strains are emerging at an increasing rate, with at times just a single point mutation in NSs (Almási et al., 2017). Global warming plays a role as well, high daytime temperatures lead to decreased effectiveness of Tsw and other Nod-like receptors (NLRs) (Chung et al., 2018; Ronde et al., 2019; Venkatesh and Kang, 2019) and allows for thrips to expand their geographical range (Krumov and Karadjova, 2012; Reitz et al., 2020).

To get closer to information on the interplay between NSs and Tsw, a two-sided track was followed (Figure 7.1). To allow investigation of Tsw, it was first necessary to obtain the sequence of *Tsw*. In Chapter 3, over 10 years after the construction of a 11x coverage BAC library of the full genome of *Capsicum chinense* PI152225, we have managed to isolate a contig containing Tsw (Chapter 3). Its translated sequence was highly similar to the by then published amino acid sequence of Tsw uncovered by Kim et al. (Kim et al., 2017a). In chapter 4, the amino acid sequences of Tsw and the susceptible variant tsw were used to compute 3D models. Apart from acknowledging the extraordinarily long leucine-rich repeat (LRR) domain found in Tsw, the 3D model raised questions regarding its function and mechanism to interact with NSs. A method was devised to create domain exchanges and mutations in Tsw easily and rapidly (Chapter 5), which will allow for functional gene studies in plants. Several microRNAs of the miR482 and -6026 family were found to influence *Tsw*-conferred resistance against TSWV. To close the gap from the other side (Figure 7.1), a search was conducted for host proteins interacting with NSs (Chapter 6). The interaction with Catalase 1 was shown to potentially benefit the virus, while no evidence was found for the hypothesized interaction between NSs and Argonaute 1 (AGO1). The different roles and locations of NSs and Tsw in the plant immune pathways have been illustrated in Figure 7.2, thereby linking back to the general overview provided in Chapter 1 and 2.

The exact modus operandi of Tsw in conferring resistance against TSWV has not yet been solved. Nevertheless, my thesis provides steppingstones for future research. In this final section, Chapter 7, I will reflect on how far this thesis has come in fulfilling the aim of unravelling the interaction between NSs and Tsw. I will raise many questions that remain and make a few recommendations to remedy the potential difficulties that answering.

The gap between Tsw and NSs

Many questions still need to be answered to close the current (knowledge) gap on the interaction between Tsw and NSs. The sequence of Tsw (Chapter 3) revealed its unusually long LRR domain, which puts the gene in 10% longest NLRs (Chapter 4). The length of the LRR defies the normal rules regarding domain sizes (Xu and Nussinov, 1998). It is therefore probable that the extra-long LRR does have a purpose. But what is it? Is the length linked to the size of the effector it recognizes, or the downstream pathway it uses to trigger an immune response? Are there specific motifs in these long NLRs that are absent in shorter, normal NLRs? Is there an overlap between the proteins with which the long NLRs Interact that would hint to a shared decoy/guardee protein, or to a shared downstream pathway for which the LRR-size is required? And what about their ability to multimerize into resistosome-like structures, or do they form other structures? The synthetically constructed chimera of Tsw (Chapter 5) can be used to attempt answer some of these questions, by for instance functionality testing after domain swaps, or

Figure 7.1 – An overview of the two-sided track in this thesis research to close the knowledge gap regarding the interaction between Tomato spotted wilt virus (TSWV) and the resistance gene *Tsw* conferring resistance against the virus, and the chapters in which these subjects are discussed. *Abbreviations: miRNA – microRNA; HF – host factor; Cat1 – Catalase 1; Siz1 – E3 SUMO-protein ligase; AGO1 – Argonaute 1.*

identifying interacting host proteins via a proximity labeling approach. As recognition of NSs by Tsw possibly could occur through an intermediate protein, proteins interacting with both NSs and Tsw (Chapter 6, (Helderman et al., 2021; Zhai et al., 2021)), will be most interesting to investigate further, while other proteins interacting with either NSs or Tsw may possibly be involved in subsequent downstream signaling pathways or as of yet unidentified functions of NSs.

While the use of chimeras for functional gene studies of Tsw has already been discussed in Chapter 5, which presents the most logical step from where many scientists embark to functionally analyze *R* genes and the activation of downstream signaling pathways, the high dissimilarity between the C-terminal tails of Tsw, tsw, and Pvr4 is especially interesting. C-terminal tails are often thought to be involved in pathogen recognition, and whether that is also the case for these three proteins will be a challenge to find out. For Rx1 and Gpa2, a sole swap of this final region switched the pathogen recognition of both genes. Whether this applies for Tsw as well, or whether it requires the rest of the LRR domain and/or interplay with other domains of Tsw remains to be investigated.

With the entire *Tsw* gene sequence elucidated, its transcriptional regulation can now also be further investigated. First experiments on silencing of miR482 and -6026 have indicated a putative role of miRNA-mediated transcriptional regulation of Tsw in a pilot experiment, although experiments need to be repeated, controls verified, and Tsw transcript levels quantified (Chapter 5). But if true, one would expect that, in line with other *R* genes regulated by miR482, *Tsw* messenger RNA (mRNA) is targeted for degradation, which leads to phased secondary small interfering RNA (phasiRNA) production and enhances its silencing. Interestingly, NSs is able to bind double stranded RNA independent of size, and binds any small interfering (si)RNA or miRNA. As such, NSs would relieve miRNA-mediated downregulation of Tsw transcripts and lead to a highly activate state of defense. As NSs protein from all tospoviruses exert RNA suppression activity in the same biochemical fashion (Hedil et al., 2017), any tospovirus able to infect *Capsicum* will activate this state of defense. Nevertheless, only TSWV NSs is recognized by Tsw and triggers activation of the downstream signaling pathway, leading to HR. This raises the question, which sequence or structure of NSs, in case of direct interaction, or which guardee/decoy is targeted only by TSWV NSs, and not by any other tospovirus NSs proteins? This will be an intriguing question to answer. Furthermore, and considering the increasing importance and emergence of other tospoviruses, can Tsw be subjected to an approach of artificial evolution to expand the resistance spectrum to perceive all tospoviruses? This approach has recently been applied successfully to generate Sw-5b variants that confer resistance to Sw-5b resistance breaking TSWV strains (Zhao et al., 2021). With this first glimpse of possible miRNA target sites in Tsw, it is also interesting to further look and study those that overlap with miRNAs that are up- and downregulated during TSWV infecting in pepper (Tao et al., 2022). And in this line of thinking, what about *Pvr4* transcription, is this regulated in the same manner as transcription of *Tsw*? Furthermore, the predicted

presence of MYC *cis*-acting regulatory elements in the 5' UTR of Tsw, also found for the tomato TSWV-resistance conferring gene *Sw-5b*, raises the possibility of a similar transcriptional regulation mechanisms. Interestingly, NSs was recently also shown to interact with MYC2, one of the jasmonate (JA) signaling regulators, and MYC3 and -4, to disable downstream terpene synthase genes and thereby attenuating host defense and increasing the attraction of thrips (Wu et al., 2019). Is there a link between the interaction of TSWV NSs and MYCs and transcription of *Sw-5b* or *Tsw*? There is evidently an intricate interplay between virus, plant host, and vector with a versatile role of NSs that is yet to be fully elucidated.

Apart from transcriptional regulation of Tsw-mediated defense responses, protein-protein interactions also play a role in the modulation of defense responses, either *R* gene-mediated, or antiviral RNAi based, with a major role for the NSs protein. While the protein triggers a Tsw-mediated HR response, the focus of this thesis research, NSs also suppresses the antiviral RNAi response. The identification of proteins interacting with NSs (Chapter 6) therefore provides leads to either defense responses. In this study, two proteins were found to interact with NSs, namely Catalase 1, which catalyzes the formation of hydrogen peroxide into water, and SIZ1, a E3 SUMO-protein ligase. Intriguingly, a hypothesized interaction between NSs and Argonaute 1 was not found. The presence of an AGO-binding WG/GW motif (Lakatos et al., 2004; Baumberger et al., 2007; Giner et al., 2010; Zhao et al., 2018) in the NSs sequence, and analysis of alanine substitution in NSs had hinted at a possible interaction between the two (de Ronde et al., 2014b). The results of neither the co-immunoprecipitation (co-IP) nor the bi-molecular fluorescent complementation (BiFC) assay used in Chapter 6 confirmed this, even though the two proteins do co-localize in cytoplasmic bodies (Figure 6.3). Do AGO1 and NSs truly not interact? Or is such an interaction very transient and therefore requires a more sensitive technique that prevent escapes, e.g. proximity labelling via biotinylation? This technique enables detection of weak spatial temporal interactions much better than BiFC or co-IP (Branon et al., 2018). The loss of RSS activity from NSs upon mutation of the WG/GW motif still provides strong support for the existence of this interaction (de Ronde et al., 2014b).

The interaction between an RSS such as NSs and a catalase is not entirely new. The RSS protein of Cucumber mosaic virus, 2b, is also able to interact with a catalase (Inaba et al., 2011; Masuta et al., 2012). However, the effect of the interaction is different as interaction with 2b moves AtCAT3 from its usual location in the peroxisomes in the cytoplasm to the nucleus, which is not the case for NSs-Cat1. What overlap is there between these two interactions? The interaction with CAT3 triggers necrosis and makes plants more tolerant for CMV, while knockout of CAT3 leads to higher viral titers. For TSWV, knockdown of Cat1 leads to slower viral infection. The opposite outcomes of both interactions indicate that NSs-NbCat1 interaction is not like that of 2b-AtCAT3. Interestingly, though, the presence of catalase seems to lower systemic silencing, a process in which both NSs and 2b play a similar modulating role. But how does it do that? Is it because lower H_2O_2 levels somehow

indicate a lower basal immune response in the cell, or does it specifically interfere and act as an endogenous suppressor of RNAi?

In addition to the differences in the outcome of the interaction between NSs and 2b with Catalase, the interaction between Cat1 and NSs does not take place in the nucleus, like that of 2b and CAT3. Based on previous literature, this would be a no-brainer, as there was no evidence for the presence of TSWV NSs in the nucleus, nor for any other tospovirus NSs protein with the exception of one report on Capsicum chlorosis virus (CaCV) NSs (Widana Gamage and Dietzgen, 2017). In the latter study, CaCV and TSWV NSs were predicted to have similar nuclear localization signals (NLSs), and further analyses on CaCV NSs showed that when this protein self-associated it localized to the nucleus. In this thesis study, an interaction between SIZ1 and TSWV NSs was observed, which took place in the nucleus (Chapter 6). Although still somewhat speculative, this unexpected finding, together with the earlier report on CaCV NSs, strengthens the idea of a, potentially transient, nuclear localization of NSs, which could be shared by all tospoviruses. But what are the implications of this interaction, and this nuclear localization? Whether nucleus entry and self-association are determined by two separate mechanisms or domains needs to be determined. However, a nuclear localization of NSs has been demonstrated previously for animal infecting orthobunyaviruses (Swanepoel and Blackburn, 1977; Gouzil et al., 2017). Could there be a shared feature or functionality between plant- and mammalian-infecting bunyavirid NSs? The NSs proteins of the animal-infecting viruses inhibit cellular transcription and act as a suppressor of the antiviral type I interferon response by targeting a subunit of RNA polymerase II for degradation through interaction with a transcription factor (Weber et al., 2002; Billecocq et al., 2004; Thomas et al., 2004). This blocks cellular mRNA transcription and halts the host response to the virus (Le May et al., 2004). Degradation of RNA polymerase-II (pol II) by La Cross Virus NSs partially depends on a cellular E3-ubiquitin ligase subunit (Mudhasani et al., 2016). Is the interaction between NSs and SIZ1 simply based in the possible sumoylation of NSs itself? TSWV NSs is predicted to contain two SUMO attachment sites (SAS) and one SUMO interaction motif (SIM). The NSs of CaCV is predicted to have no SIMS and two SAS, one of which seems to be conserved between the two NSs proteins. Further work is needed here, but if indeed true, what would be the relevance of sumoylation of NSs? Could the interaction between SIZ1 and NSs also result in suppression of transcription by pol II as is the case for RVFV NSs (Billecocq et al., 2004; Le May et al., 2004)? Are we looking at evolutionary shared properties of bunyavirus NSs proteins? Would TSWV NSs then also be able to block INF responses in mammalian cell lines? And what about the possible RNA silencing suppression properties of RVFV NSs in plants? How distant are these proteins from each other, evolutionary speaking? And if they still share these properties, one wonders about the selection pressures to maintain these functions. Or are host defense pathways conserved enough between these two very different hosts (plants vs. animals) that their corresponding infecting bunyaviruses can trigger a similar response in the other (non-)host? And, although speculative, do the structurally (partially) resembling Toll-like

receptors (TLRs) of Tsw in vertebrates recognize bunyaviral NSs in a way in which the NLR Tsw recognizes Tsw? It is clear that questions that aim to unravel the role and interplay of TSWV NSs with the host innate immunity system, will benefit from a cross pollination of the evolutionary related animal infecting bunyaviruses.

Previous studies have shown that the SIZ1 protein may act as a suppressor of pathogen related genes (Lee et al., 2007; Sharma et al., 2021a). SIZ1 was for instance shown to regulate protein levels of the NB-LRR protein SNC1 in *A. thaliana*, which was furthermore sumoylated in plants (Gou et al., 2017). SIZ1 also mediates regulation of plant immune responses and growth in response to temperature (Hammoudi et al., 2018). Sumoylation is known to play a role in the (re)localization and functioning of proteins. For several NLRs, nucleocytoplasmic shuttling has been observed, in which nuclear localization is required for a resistance response (Du et al., 2015; Richard et al., 2021). Whether SIZ1 plays a role in the regulation/activation of downstream responses of Tsw after interaction with NSs, by altering its nucleocytoplasmic distribution, remains an interesting idea to be tested.
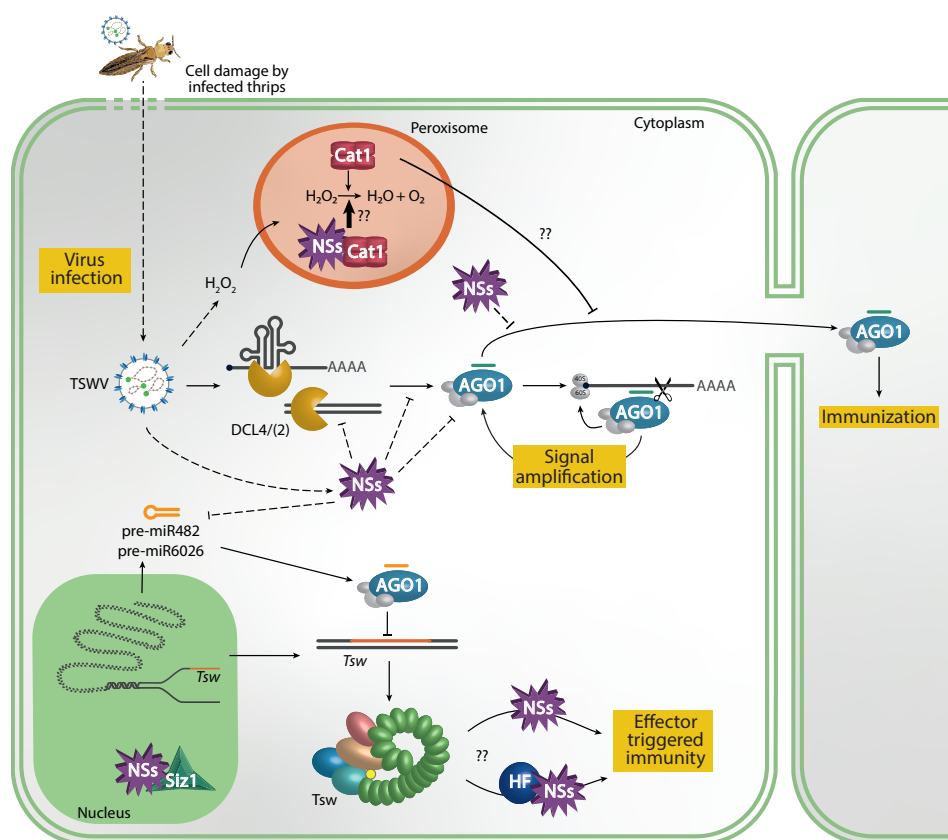
Figure 7.2 – A model of the roles of Nod-like receptor *Tsw* and the RNA silencing suppressor NSs in the multilayered plant immune system. *Abbreviations: AGO1- Argonaute 1, DCL – Dicer-like; Cat1 – Catalase 1; pre-miRNA – pre-mirco RNA; HF – host factor; TSWV – Tomato spotted wilt virus.*

149

Concluding remarks

Many questions have been raised that require answers to start closing our knowledge gap on NSs-triggering of Tsw-mediated defense responses. Some of these questions also touch on the evolutionary relation with the animal-infecting bunyaviral NSs proteins and their modulation of the vertebrate innate immunity response. The research presented in this thesis provides some increased insight of the intricacies of Tsw, as well as new insights into host proteins interacting with NSs (Figure 7.1). Based on the findings in this thesis, a model is postulated for the role of Tsw and NSs in the multilayered immune pathways that have been described in Chapter 1 and 2 (Figure 7.2). While there are various TSWV-resistant pepper variants, resistance in all varieties depends on the Tsw (Moury et al., 1997). This makes *Tsw* a valuable gene for breeders and growers despite the ever-growing number of resistant breaking strains (Ferrand et al., 2015; Jiang et al., 2017a; Macedo et al., 2019; Almási et al., 2020; Kwon et al., 2021). While the probable pro-viral factor Catalase 1 is partially essential in plants and can therefore not be removed to limit viral spread, its role in resistance and susceptibility is an interesting find. The nuclear localization and interaction of NSs with Siz1 provides highly interesting leads to further investigate NSs and a link with either the Tsw-mediated defense response or antiviral RNAi. Continuation of the efforts to further understand NSs will help us understand its multi-functionality as RNA Silencing suppressor and potential regulator of transcription-translation, as well as its detection by *Tsw*.

Appendices

References
Summary
Samenvatting
Word of thanks
Education statement
About the author

References

A

Abcam (2021). Catalase Activity Assay Kit (Colorimetric/Fluorometric). 1–3. Available at: https://www.abcam.com/ps/products/83/ab83464/documents/catalase-activity-assay-kit-protocol-book-v13-ab83464 (website).pdf.

Abraham, A., and Savithri, H. S. (2016). A novel viral RNA helicase with an independent translation enhancement activity. *FEBS Lett.* 590, 1187–1199. doi: 10.1002/1873-3468.12145.

Adachi, H., Derevnina, L., and Kamoun, S. (2019). NLR singletons, pairs, and networks: evolution, assembly, and regulation of the intracellular immunoreceptor circuitry of plants. *Curr. Opin. Plant Biol.* 50, 121–131. doi: 10.1016/j.pbi.2019.04.007.

Adkins, S., Quadt, R., Choi, T.-J., Ahlquist, P., and German, T. (1995). An RNA-Dependent RNA Polymerase Activity Associated with Virions of Tomato Spotted Wilt Virus, a Plant- and Insect-Infecting Bunyavirus. *Virology* 207, 308–311. doi: 10.1006/viro.1995.1083.

Aebi, H. (1984). "Catalase in vitro," in *Methods in enzymology*, 121–126. doi: 10.1016/S0076-6879(84)05016-3.

Almási, A., Csilléry, G., Csömör, Z., Nemes, K., Palkovics, L., Salánki, K., et al. (2015). Phylogenetic analysis of Tomato spotted wilt virus (TSWV) NSs protein demonstrates the isolated emergence of resistance-breaking strains in pepper. *Virus Genes* 50, 71–78. doi: 10.1007/s11262-014-1131-3.

Almási, A., Nemes, K., Csömör, Z., Tóbiás, I., Palkovics, L., and Salánki, K. (2017). A single point mutation in Tomato spotted wilt virus NSs protein is sufficient to overcome Tsw-gene-mediated resistance in pepper. *J. Gen. Virol.* 98, 1521–1525. doi: 10.1099/jgv.0.000798.

Almási, A., Nemes, K., and Salánki, K. (2020). Increasing diversity of resistance breaking pepper strains of tomato spotted wilt virus in the mediterranean region. *Phytopathol. Mediterr.* 59, 385–391. doi: 10.14601/Phyto-11346.

Anjum, N. A., Sharma, P., Gill, S. S., Hasanuzzaman, M., Khan, E. A., Kachhap, K., et al. (2016). Catalase and ascorbate peroxidase—representative H2O2-detoxifying heme enzymes in plants. *Environ. Sci. Pollut. Res.* 23, 19002–19029. doi: 10.1007/s11356-016-7309-6.

Aramburu, J., and Marti, M. (2003). The occurrence in north-east Spain of a variant of Tomato spotted wilt virus (TSWV) that breaks resistance in tomato (Lycopersicon esculentum) containing the Sw-5 gene. *Plant Pathol.* 52, 407–407. doi: 10.1046/j.1365-3059.2003.00829.x.

de Avila, A. C., Huguenot, C., Resende, R. d. O., Kitajima, E. W., Goldbach, R. W., and Peters, D. (1990). Serological Differentiation of 20 Isolates of Tomato Spotted Wilt Virus. *J. Gen. Virol.* 71, 2801–2807. doi: 10.1099/0022-1317-71-12-2801.

B

Bakker, E., van Vliet, J., Overmars, H., Smant, G., Sandbrink, H., Van der Vossen, E., et al. (2004). "R gene homologues in potato confer resistance to distinct pathogens: a virus and a nematode," in *Proceedings of the Fourth International Congress of Nematology, 8-13 June 2002, Tenerife, Spain* (BRILL), 359–365. doi: 10.1163/9789004475236_037.

Batuman, O., Turini, T. A., Oliveira, P. V., Rojas, M. R., Macedo, M., Mellinger, H. C., et al. (2017). First Report of a Resistance-Breaking Strain of Tomato spotted wilt virus Infecting Tomatoes With the Sw- 5 Tospovirus-Resistance Gene in California. *Plant Dis.* 101, 637–637. doi: 10.1094/PDIS-09-16-1371-PDN.

Baumberger, N., Tsai, C.-H. H., Lie, M., Havecker, E., and Baulcombe, D. C. C. (2007). The Polerovirus Silencing Suppressor P0 Targets ARGONAUTE Proteins for Degradation. *Curr. Biol.* 17, 1609–1614. doi: 10.1016/j.cub.2007.08.039.

Berka, M., Kopecká, R., Berková, V., Brzobohatý, B., and Černý, M. (2022). Regulation of heat shock proteins 70 and their role in plant immunity. *J. Exp. Bot.* 73, 1894–1909. doi: 10.1093/jxb/erab549.

Bernoux, M., Ve, T., Williams, S., Warren, C., Hatters, D., Valkov, E., et al. (2011). Structural and Functional Analysis of a Plant Resistance Protein TIR Domain Reveals Interfaces for Self-Association, Signaling, and Autoregulation. *Cell Host Microbe* 9, 200–211. doi: 10.1016/j.chom.2011.02.009.

Bezerra, I. C., de O. Resende, R., Pozzer, L., Nagata, T., Kormelink, R., and De Ávila, A. C. (1999). Increase of Tospoviral Diversity in Brazil with the Identification of Two New Tospovirus Species, One from Chrysanthemum and One from Zucchini. *Phytopathology®* 89, 823–830. doi: 10.1094/PHYTO.1999.89.9.823.

Bhushan, L., Abraham, A., Choudhury, N. R., Rana, V. S., Mukherjee, S. K., and Savithri, H. S. (2015). Demonstration of helicase activity in the nonstructural protein, NSs, of the negative-sense RNA virus, Groundnut bud necrosis virus. *Arch. Virol.* 160, 959–967. doi: 10.1007/s00705-014-2331-9.

Bi, G., Su, M., Li, N., Liang, Y., Dang, S., Xu, J., et al. (2021). The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. *Cell* 184, 3528-3541.e12. doi: 10.1016/j.cell.2021.05.003.

van der Biezen, E. A., and Jones, J. D. G. (1998). The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* 8, R226–R228. doi: 10.1016/S0960-9822(98)70145-9.

Billecocq, A., Spiegel, M., Vialat, P., Kohl, A., Weber, F., Bouloy, M., et al. (2004). NSs Protein of Rift Valley Fever Virus Blocks Interferon Production by Inhibiting Host Gene Transcription. *J. Virol.* 78, 9798–9806. doi: 10.1128/JVI.78.18.9798-9806.2004.

Black, L. L. (1991). Tomato Spotted Wilt Virus Resistance in Capsicum chinense PI 152225 and 159236. *Plant Dis.* 75, 863A. doi: 10.1094/PD-75-0863A.

Blakqori, G., Delhaye, S., Habjan, M., Blair, C. D., Sánchez-Vargas, I., Olson, K. E., et al. (2007). La Crosse Bunyavirus Nonstructural Protein NSs Serves To Suppress the Type I Interferon System of Mammalian Hosts. *J. Virol.* 81, 4991–4999. doi: 10.1128/JVI.01933-06.

Boevink, P., Oparka, K., Cruz, S. S., Martin, B., Betteridge, A., and Hawes, C. (1998). Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network †. *Plant J.* 15, 441–447. doi: 10.1046/j.1365-313X.1998.00208.x.

Boiteux, L. S. (1995). Allelic relationships between genes for resistance to tomato spotted wilt tospovirus in Capsicum chinense. *Theor. Appl. Genet.* 90, 146–149. doi: 10.1007/BF00221009.

Boiteux, L. S., and de Ávila, A. C. (1994). Inheritance of a resistance specific to tomato spotted wilt tospovirus in Capsicum chinense 'PI 159236.' *Euphytica* 75, 139–142. doi: 10.1007/BF00024541.

Bouloy, M., Plotch, S. J., and Krug, R. M. (1978). Globin mRNAs are primers for the transcription of influenza viral RNA in vitro. *Proc. Natl. Acad. Sci.* 75, 4886–4890. doi: 10.1073/pnas.75.10.4886.

Brandizzi, F., Snapp, E. L., Roberts, A. G., Lippincott-Schwartz, J., and Hawes, C. (2002). Membrane Protein Transport between the Endoplasmic Reticulum and the Golgi in Tobacco Leaves Is Energy Dependent but Cytoskeleton Independent. *Plant Cell* 14, 1293–1309. doi: 10.1105/tpc.001586.

Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., et al. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* 36, 880–887. doi: 10.1038/nbt.4201.

Brommonschenkel, S. H., Frary, A., Frary, A., and Tanksley, S. D. (2000). The Broad-Spectrum Tospovirus Resistance Gene Sw-5 of Tomato Is a Homolog of the Root-Knot Nematode Resistance Gene Mi. *Mol. Plant-Microbe Interact.* 13, 1130–1138. doi: 10.1094/MPMI.2000.13.10.1130.

Bucher, E. (2006). Antiviral RNA silencing and viral counter defense in plants.

Bucher, E., Sijen, T., de Haan, P., Goldbach, R., and Prins, M. (2003). Negative-Strand Tospoviruses and Tenuiviruses Carry a Gene for a Suppressor of Gene Silencing at Analogous Genomic Positions. *J. Virol.* 77, 1329–1336. doi: 10.1128/jvi.77.2.1329-1336.2003.

Burdon, J. J., and Thrall, P. H. (2003). The fitness costs to plants of resistance to pathogens. *Genome Biol.* 4, 1–3. doi: 10.1186/gb-2003-4-9-227.

C

CABI Datasheet Report for Frankliniella Occidentalis (Western Flower Thrips) Available online: https://www.cabi.org/isc/datasheetreport/24426 (accessed on 11 July 2022).

Canady, M. A., Stevens, M. R., Barineau, M. S., and Scott, J. W. (2001). Tomato spotted wilt virus (TSMV) resistance in tomato derived from Lycopersicon chilense Dun. LA 1938. *Euphytica*. doi: doi.org/10.1023/A:1004089504051.

Caplan, J. L., Mamillapalli, P., Burch-Smith, T. M., Czymmek, K., and Dinesh-Kumar, S. P. (2008a). Chloroplastic protein NRIP1 mediates innate immune receptor recognition of a viral effector. *Cell* 132, 449–462. doi: 10.1016/j.cell.2007.12.031.

Caplan, J., Padmanabhan, M., and Dinesh-Kumar, S. P. (2008b). Plant NB-LRR Immune Receptors: From Recognition to Transcriptional Reprogramming. *Cell Host Microbe* 3, 126–135. doi: 10.1016/j.chom.2008.02.010.

Casey, L. W., Lavrencic, P., Bentham, A. R., Cesari, S., Ericsson, D. J., Croll, T., et al. (2016). The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins. *Proc. Natl. Acad. Sci.* 113, 12856–12861. doi: 10.1073/pnas.1609922113.

Casteel, C. L., Walling, L. L., and Paine, T. D. (2006). Behavior and biology of the tomato psyllid, Bactericerca cockerelli, in response to the Mi-1.2 gene. *Entomol. Exp. Appl.* 121, 67–72. doi: 10.1111/j.1570-8703.2006.00458.x.

Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P. N. (2014). A novel conserved mechanism for plant NLR protein pairs: The "integrated decoy" hypothesis. *Front. Plant Sci.* doi: 10.3389/fpls.2014.00606.

Chan, A. Y., Vreede, F. T., Smith, M., Engelhardt, O. G., and Fodor, E. (2006). Influenza virus inhibits RNA polymerase II elongation. *Virology* 351, 210–217. doi: 10.1016/j.virol.2006.03.005.

Che, F.-S., Nakajima, Y., Tanaka, N., Iwano, M., Yoshida, T., Takayama, S., et al. (2000). Flagellin from an Incompatible Strain of Pseudomonas avenae Induces a Resistance Response in Cultured Rice Cells. *J. Biol. Chem.* 275, 32347–32356. doi: 10.1074/jbc.M004796200.

Chen, I.-H. H., Tsai, A. Y., Huang, Y.-P. P., Wu, I.-F. F., Cheng, S.-F. F., Hsu, Y.-H. H., et al. (2017). Nuclear-Encoded Plastidal Carbonic Anhydrase Is Involved in Replication of Bamboo mosaic virus RNA in Nicotiana benthamiana. *Front. Microbiol.* 8, 1–11. doi: 10.3389/fmicb.2017.02046.

Chen, X., Zhu, M., Jiang, L., Zhao, W., Li, J., Wu, J., et al. (2016). A multilayered regulatory mechanism for the autoinhibition and activation of a plant CC-NB-LRR resistance protein with an extra N-terminal domain. *New Phytol.* 212, 161–175. doi: 10.1111/nph.14013.

Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., and Felix, G. (2006). The Arabidopsis Receptor Kinase FLS2 Binds flg22 and Determines the Specificity of Flagellin Perception. *Plant Cell* 18, 465–476. doi: 10.1105/tpc.105.036574.

Chung, B. N., Choi, H. S., Yang, E. Y., Cho, J. D., Cho, I. S., Choi, G. S., et al. (2012). Tomato spotted wilt virus Isolates giving different infection in commercial Capsicum annuum cultivars. *Plant Pathol. J.* 28, 87–92. doi: 10.5423/PPJ.NT.09.2011.0169.

Chung, B. N., Lee, J. H., Kang, B. C., Koh, S. W., Joa, J. H., Choi, K. S., et al. (2018). HR-mediated defense response is overcome at high temperatures in capsicum species. *Plant Pathol. J.* 34, 71–77. doi: 10.5423/PPJ.NT.06.2017.0120.

Cifuentes-Muñoz, N., Salazar-Quiroz, N., and Tischler, N. D. (2014). Hantavirus Gn and Gc envelope glycoproteins: Key structural units for virus cell entry and virus assembly. *Viruses* 6, 1801–1822. doi: 10.3390/v6041801.

Ciuffo, M., Finetti-Sialer, M. M., Gallitelli, D., and Turina, M. (2005). First report in Italy of a resistance-breaking strain of Tomato spotted wilt virus infecting tomato cultivars carrying the Sw5 resistance gene. *Plant Pathol.* 54, 564–564. doi: 10.1111/j.1365-3059.2005.01203.x.

Collier, S. M., and Moffett, P. (2009a). NB-LRRs work a "bait and switch" on pathogens. *Trends Plant Sci.* 14, 521–529. doi: 10.1016/j.tplants.2009.08.001.

Collier, S. M., and Moffett, P. (2009b). NB-LRRs work a "bait and switch" on pathogens. *Trends Plant Sci.* 14, 521–529. doi: 10.1016/j.tplants.2009.08.001.

Cyr, N., de la Fuente, C., Lecoq, L., Guendel, I., Chabot, P. R., Kehn-Hall, K., et al. (2015). A ΩX a V motif in the Rift Valley fever virus NSs protein is essential for degrading p62, forming nuclear filaments and virulence. *Proc. Natl. Acad. Sci.* 112, 6021–6026. doi: 10.1073/pnas.1503688112.

D

Dahan-Meir, T., Filler-Hayut, S., Melamed-Bessudo, C., Bocobza, S., Czosnek, H., Aharoni, A., et al. (2018). Efficient in planta gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *Plant J.* 95, 5–16. doi: 10.1111/tpj.13932.

Dai, X., Zhuang, Z., and Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 46, W49–W54. doi: 10.1093/nar/gky316.

Dodds, P. N., Lawrence, G. J., and Ellis, J. G. (2001). Six Amino Acid Changes Confined to the Leucine-Rich Repeat β-Strand/β-Turn Motif Determine the Difference between the P and P2 Rust Resistance Specificities in Flax. *Plant Cell* 13, 163–178. doi: 10.1105/tpc.13.1.163.

Du, J., Song, X. Y., Shi, X. Bin, Tang, X., Chen, J. Bin, Zhang, Z. H. Z., et al. (2020). NSs, the Silencing Suppressor of Tomato Spotted Wilt Orthotospovirus, Interferes With JA-Regulated Host Terpenoids Expression to Attract Frankliniella occidentalis. *Front. Microbiol.* 11, 1–12. doi: 10.3389/fmicb.2020.590451.

Du, Y., Berg, J., Govers, F., and Bouwmeester, K. (2015). Immune activation mediated by the late blight resistance protein R1 requires nuclear localization of R1 and the effector <scp>AVR</scp> 1. *New Phytol.* 207, 735–747. doi: 10.1111/nph.13355.

Dufresne, P. J., Ubalijoro, E., Fortin, M. G., Laliberté, J.-F., and Laliberte, J. F. (2008). Arabidopsis thaliana class II poly(A)-binding proteins are required for efficient multiplication of turnip mosaic virus. *J. Gen. Virol.* 89, 2339–2348. doi: 10.1099/vir.0.2008/002139-0.

Duijsings, D., Kormelink, R., and Goldbach, R. (1999). Alfalfa Mosaic Virus RNAs Serve as Cap Donors for Tomato Spotted Wilt Virus Transcription during Coinfection of Nicotiana benthamiana. *J. Virol.* 73, 5172–5175. doi: 10.1128/JVI.73.6.5172-5175.1999.

Duxbury, Z., Wu, C.-H., and Ding, P. (2021). A Comparative Overview of the Intracellular Guardians of Plants and Animals: NLRs in Innate Immunity and Beyond. *Annu. Rev. Plant Biol.* 72, 155–184. doi: 10.1146/annurev-arplant-080620-104948.

F

Farnham, G., and Baulcombe, D. C. (2006). Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc. Natl. Acad. Sci.* 103, 18828–18833. doi: 10.1073/pnas.0605777103.

Felix, G., Duran, J. D., Volko, S., and Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J.* 18, 265–276. doi: 10.1046/j.1365-313X.1999.00265.x.

Feng, Z., Chen, X., Bao, Y., Dong, J., Zhang, Z., and Tao, X. (2013). Nucleocapsid of Tomato spotted wilt tospovirus forms mobile particles that traffic on an actin/endoplasmic reticulum network driven by myosin XI -K. *New Phytol.* 200, 1212–1224. doi: 10.1111/nph.12447.

Feng, Z., Xue, F., Xu, M., Chen, X., Zhao, W., Garcia-Murria, M. J., et al. (2016). The ER-Membrane Transport System Is Critical for Intercellular Trafficking of the NSm Movement Protein and Tomato Spotted Wilt Tospovirus. *PLOS Pathog.* 12, e1005443. doi: 10.1371/journal.ppat.1005443.

Ferrand, L., García, M. L., Resende, R. O., Balatti, P. A., and Dal Bó, E. (2015). First Report of a Resistance-breaking Isolate of Tomato spotted wilt virus Infecting Sweet Pepper Harboring the Tsw Gene in Argentina. *Plant Dis.* 99, 1869. doi: 10.1094/PDIS-02-15-0207-PDN.

Flor, H. H. (1971). Current Status of the Gene-For-Gene Concept. *Annu. Rev. Phytopathol.* 9, 275–296. doi: 10.1146/annurev.py.09.090171.001423.

Folkertsma, R. T., Spassova, M. I., Prins, M., Stevens, M. R., Hille, J., and Goldbach, R. W. (1999). Construction of a bacterial artificial chromosome (BAC) library of Lycopersicon esculentum cv. Stevens and its application to physically map the Sw-5 locus. *Mol. Breed.* 5, 197–207. doi: 10.1023/A:1009650424891.

Foster, S. J., Park, T. H., Pel, M., Brigneti, G., Sliwka, J., Jagger, L., et al. (2009). Rpi-vnt1.1, a Tm-22 homolog from Solanum venturii, confers resistance to potato late blight. *Mol. Plant-Microbe Interact.* doi: 10.1094/MPMI-22-5-0589.

Fraile, A., Pagan, I., Anastasio, G., Saez, E., and Garcia-Arenal, F. (2011). Rapid Genetic Diversification and High Fitness Penalties Associated with Pathogenicity Evolution in a Plant Virus. *Mol. Biol. Evol.* 28, 1425–1437. doi: 10.1093/molbev/msq327.

G

García-Cano, E., Resende, R. O., Fernández-Muñoz, R., and Moriones, E. (2006). Synergistic Interaction Between Tomato chlorosis virus and Tomato spotted wilt virus Results in Breakdown of Resistance in Tomato. *Phytopathology®* 96, 1263–1269. doi: 10.1094/PHYTO-96-1263.

Garcia-Ruiz, H. (2018). Susceptibility Genes to Plant Viruses. *Viruses* 10, 484. doi: 10.3390/v10090484.

Garcia-Ruiz, H., Gabriel Peralta, S. M., and Harte-Maxwell, P. A. (2018). Tomato Spotted Wilt Virus NSs Protein Supports Infection and Systemic Movement of a Potyvirus and Is a Symptom Determinant. *Viruses* 10, 129. doi: 10.3390/v10030129.

Geerts-Dimitriadou, C.; Lu, Y.-Y.; Geertsema, C.; Goldbach, R.; Kormelink, R. (2012) Analysis of the Tomato Spotted Wilt Virus Ambisense S RNA-Encoded Hairpin Structure in Translation. *PLoS ONE*, *7*, e31013, doi:10.1371/journal.pone.0031013.

Geerts-Dimitriadou, C., Goldbach, R., and Kormelink, R. (2011a). Preferential use of RNA leader sequences during influenza A transcription initiation in vivo. *Virology* 409, 27–32. doi: 10.1016/j.virol.2010.09.006.

Geerts-Dimitriadou, C., Zwart, M. P., Goldbach, R., and Kormelink, R. (2011b). Base-pairing promotes leader selection to prime in vitro influenza genome transcription. *Virology* 409, 17–26. doi: 10.1016/j.virol.2010.09.003.

Gehl, C., Waadt, R., Kudla, J., Mendel, R.-R. R., and Hänsch, R. (2009). New GATEWAY vectors for High Throughput Analyses of Protein–Protein Interactions by Bimolecular Fluorescence Complementation. *Mol. Plant* 2, 1051–1058. doi: 10.1093/mp/ssp040.

Gielen, J. J. L., de Haan, P., Kool, A. J., Peters, D., van Grinsven, M. Q. J. M., and Goldbach, R. W. (1991). Engineered Resistance to Tomato Spotted Wilt Virus, a Negative–Strand RNA Virus. *Bio/Technology* 9, 1363–1367. doi: 10.1038/nbt1291-1363.

Gilbertson, R. L., Batuman, O., Webster, C. G., and Adkins, S. (2015). Role of the Insect Supervectors Bemisia tabaci and Frankliniella occidentalis in the Emergence and Global Spread of Plant Viruses. *Annu. Rev. Virol.* 2, 67–93. doi: 10.1146/annurev-virology-031413-085410.

Giner, A., Lakatos, L., García-Chapa, M., López-Moya, J. J., and Burgyán, J. (2010). Viral protein inhibits RISC activity by argonaute binding through conserved WG/GW motifs. *PLoS Pathog.* 6, 1–13. doi: 10.1371/journal.ppat.1000996.

Goldbach, R., and Peters, D. (1994). Possible causes of the emergence of tospovirus diseases. *Semin. Virol.* 5, 113–120. doi: 10.1006/smvy.1994.1012.

Gordillo, L. F., Stevens, M. R., Millard, M. A., and Geary, B. (2008). Screening Two Lycopersicon peruvianum Collections for Resistance to Tomato spotted wilt virus. *Plant Dis.* 92, 694–704. doi: 10.1094/PDIS-92-5-0694.

Goto, K., Kobori, T., Kosaka, Y., Natsuaki, T., and Masuta, C. (2007). Characterization of Silencing Suppressor 2b of Cucumber Mosaic Virus Based on Examination of its Small RNA-Binding Abilities. *Plant Cell Physiol.* 48, 1050–1060. doi: 10.1093/pcp/pcm074.

Gou, M., Huang, Q., Qian, W., Zhang, Z., Jia, Z., and Hua, J. (2017). Sumoylation E3 Ligase SIZ1 Modulates Plant Immunity Partly through the Immune Receptor Gene SNC1 in Arabidopsis. *Mol. Plant-Microbe Interact.* 30, 334–342. doi: 10.1094/MPMI-02-17-0041-R.

Goulden, M. G., Köhm, B. A., Cruz, S. S., Kavanagh, T. A., and Baulcombe, D. C. (1993). A Feature of the Coat Protein of Potato Virus X Affects Both Induced Virus Resistance in Potato and Viral Fitness. *Virology* 197, 293–302. doi: 10.1006/viro.1993.1590.

Gouzil, J., Fablet, A., Lara, E., Caignard, G., Cochet, M., Kundlacz, C., et al. (2017). Nonstructural Protein NSs of Schmallenberg Virus Is Targeted to the Nucleolus and Induces Nucleolar Disorganization. *J. Virol.* 91, 1–15. doi: 10.1128/JVI.01263-16.

Guo, Y., Liu, B., Ding, Z., Li, G., Liu, M., Zhu, D., et al. (2017). Distinct Mechanism for the Formation of the Ribonucleoprotein Complex of Tomato Spotted Wilt Virus. *J. Virol.* 91. doi: 10.1128/JVI.00892-17.

Guo, Z., Li, Y., and Ding, S.-W. (2019). Small RNA-based antimicrobial immunity. *Nat. Rev. Immunol.* 19, 31–44. doi: 10.1038/s41577-018-0071-x.

Gupta, S., Malviya, N., Kushwaha, H., Nasim, J., Bisht, N. C., Singh, V. K., et al. (2015). Insights into structural and functional diversity of Dof (DNA binding with one finger) transcription factor. *Planta* 241, 549–562. doi: 10.1007/s00425-014-2239-3.

H

de Haan, P., Kormelink, R., de Oliveira Resende, R., van Poelwijk, F., Peters, D., and Goldbach, R. (1991). Tomato spotted wilt virus L RNA encodes a putative RNA polymerase. *J. Gen. Virol.* 72, 2207–2216. doi: 10.1099/0022-1317-72-9-2207.

de Haan, P., Gielen, J. J. L., Prins, M., Wijkamp, I. G., Schepen, A. van, Peters, D., et al. (1992). Characterization of RNA–Mediated Resistance to Tomato Spotted Wilt Virus in Transgenic Tobacco Plants. *Nat. Biotechnol.* 10, 1133–1137. doi: 10.1038/nbt1092-1133.

Hallwass, M., Oliveira, A. S., Campos Dianese, E., Lohuis, D., Boiteux, L. S., Inoue-Nagata, A. K., et al. (2014). The T omato spotted wilt virus cell-to-cell movement protein (NS M) triggers a hypersensitive response in S w-5 -containing resistant tomato lines and in N icotiana benthamiana transformed with the functional S w-5b resistance gene copy. *Mol. Plant Pathol.* 15, 871–880. doi: 10.1111/mpp.12144.

A

Hammoudi, V., Fokkens, L., Beerens, B., Vlachakis, G., Chatterjee, S., Arroyo-Mateos, M., et al. (2018). The Arabidopsis SUMO E3 ligase SIZ1 mediates the temperature dependent trade-off between plant immunity and growth. *PLOS Genet.* 14, e1007157. doi: 10.1371/journal.pgen.1007157.

Han, S., Wang, Y., Zheng, X., Jia, Q., Zhao, J., Bai, F., et al. (2015). Cytoplastic Glyceraldehyde-3-Phosphate Dehydrogenases Interact with ATG3 to Negatively Regulate Autophagy and Immunity in Nicotiana benthamiana. *Plant Cell* 27, 1316–1331. doi: 10.1105/tpc.114.134692.

Harris, C. J., Slootweg, E. J., Goverse, A., and Baulcombe, D. C. (2013). Stepwise artificial evolution of a plant disease resistance gene. *Proc. Natl. Acad. Sci.* 110, 21189–21194. doi: 10.1073/pnas.1311134110.

Hashimoto, M., Neriya, Y., Yamaji, Y., and Namba, S. (2016). Recessive Resistance to Plant Viruses: Potential Resistance Genes Beyond Translation Initiation Factors. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.01695.

Hassani-Mehraban, A., Brenkman, A. B., van den Broek, N. J. F., Goldbach, R., and Kormelink, R. (2009). RNAi-Mediated Transgenic Tospovirus Resistance Broken by IntraspeciesSilencing Suppressor Protein Complementation. *Mol. Plant-Microbe Interact.* 22, 1250–1257. doi: 10.1094/MPMI-22-10-1250.

Hatsugai, N., Igarashi, D., Mase, K., Lu, Y., Tsuda, Y., Chakravarthy, S., et al. (2017). A plant effector-triggered immunity signaling sector is inhibited by pattern-triggered immunity. *EMBO J.* 36, 2758–2769. doi: 10.15252/embj.201796529.

Hedil, M., de Ronde, D., and Kormelink, R. (2017). Biochemical analysis of NSs from different tospoviruses. *Virus Res.* 242, 149–155. doi: 10.1016/j.virusres.2017.09.020.

Hedil, M., Hassani-Mehraban, A., Lohuis, D., and Kormelink, R. (2014). Analysis of the A-U rich hairpin from the intergenic region of tospovirus S RNA as target and inducer of RNA silencing. *PLoS One.* doi: 10.1371/journal.pone.0106027.

Hedil, M., and Kormelink, R. (2016). Viral RNA silencing suppression: The enigma of bunyavirus NSs proteins. *Viruses* 8. doi: 10.3390/v8070208.

Hedil, M., Sterken, M. G., de Ronde, D., Lohuis, D., and Kormelink, R. (2015). Analysis of Tospovirus NSs Proteins in Suppression of Systemic Silencing. *PLoS One* 10, e0134517. doi: 10.1371/journal.pone.0134517.

Helderman, T. A., Deurhof, L., Bertran, A., Boeren, S., Fokkens, L., Kormelink, R., et al. (2021). An Isoform of the Eukaryotic Translation Elongation Factor 1A (eEF1a) Acts as a Pro-Viral Factor Required for Tomato Spotted Wilt Virus Disease in Nicotiana benthamiana. *Viruses* 13, 1–26. doi: 10.3390/v13112190.

Helderman, T. A., Deurhof, L., Bertran, A., Richard, M. M. S. S., Kormelink, R., Prins, M., et al. (2022). Members of the ribosomal protein S6 (RPS6) family act as pro-viral factor for tomato spotted wilt orthotospovirus infectivity in Nicotiana benthamiana. *Mol. Plant Pathol.* 23, 431–446. doi: 10.1111/mpp.13169.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* 8, R19. doi: 10.1186/gb-2007-8-2-r19.

Hemmes, H., Lakatos, L., Goldbach, R., Burgyán, J., and Prins, M. (2007). The NS3 protein of Rice hoja blanca tenuivirus suppresses RNA silencing in plant and insect hosts by efficiently binding both siRNAs and miRNAs. *RNA* 13, 1079–1089. doi: 10.1261/rna.444007.

Hohmann, U., Santiago, J., Nicolet, J., Olsson, V., Spiga, F. M., Hothorn, L. A., et al. (2018). Mechanistic basis for the activation of plant membrane receptor kinases by SERK-family coreceptors. *Proc. Natl. Acad. Sci. U. S. A.* 115, 3488–3493. doi: 10.1073/pnas.1714972115.

van der Hoorn, R. A. L., and Kamoun, S. (2008). From guard to decoy: A new model for perception of plant pathogen effectors. *Plant Cell.* doi: 10.1105/tpc.108.060194.

Huang, C.-H., Foo, M.-H., Raja, J. A. J., Tan, Y.-R., Lin, T.-T., Lin, S.-S., et al. (2020). A Conserved Helix in the C-Terminal Region of Watermelon Silver Mottle Virus Nonstructural Protein S Is Imperative For Protein Stability Affecting Self-Interaction, RNA Silencing Suppression, and Pathogenicity. *Mol. Plant-Microbe Interact.* 33, 637–652. doi: 10.1094/mpmi-10-19-0279-r.

Huang, C., Liu, Y., Yu, H., Yuan, C., Zeng, J., Zhao, L., et al. (2018). Non-Structural Protein NSm of Tomato Spotted Wilt Virus Is an Avirulence Factor Recognized by Resistance Genes of Tobacco and Tomato via Different Elicitor Active Sites. *Viruses* 10, 660. doi: 10.3390/v10110660.

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., et al. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* 14, 71–73. doi: 10.1038/nmeth.4067.

Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol. Biol. Evol.* 33, 1635–1638. doi: 10.1093/molbev/msw046.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. *J. Mol. Graph.* 14, 33–38. doi: 10.1016/0263-7855(96)00018-5.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9, 90–95. doi: 10.1109/MCSE.2007.55.

Huot, B., Yao, J., Montgomery, B. L., and He, S. Y. (2014). Growth–Defense Tradeoffs in Plants: A Balancing Act to Optimize Fitness. *Mol. Plant* 7, 1267–1287. doi: 10.1093/mp/ssu049.

I

Ikeuchi, M., Iwase, A., Ito, T., Tanaka, H., Favero, D. S., Kawamura, A., et al. (2022). Wound-inducible WUSCHEL-RELATED HOMEOBOX 13 is required for callus growth and organ reconnection. *Plant Physiol.* 188, 425–441. doi: 10.1093/plphys/kiab510.

Inaba, J., Kim, B. M., Shimura, H., and Masuta, C. (2011). Virus-Induced Necrosis Is a Consequence of Direct Protein-Protein Interaction between a Viral RNA-Silencing Suppressor and a Host Catalase. *Plant Physiol.* 156, 2026–2036. doi: 10.1104/pp.111.180042.

J

Jahn, M., Paran, I., Hoffmann, K., Radwanski, E. R., Livingstone, K. D., Grube, R. C., et al. (2000). Genetic Mapping of the Tsw Locus for Resistance to the Tospovirus Tomato spotted wilt virus in Capsicum spp. and Its Relationship to the Sw-5 Gene for Resistance to the Same Pathogen in Tomato. *Mol. Plant-Microbe Interact.* 13, 673–682. doi: 10.1094/MPMI.2000.13.6.673.

Jan, F.-J., Fagoaga, C., Pang, S.-Z., and Gonsalves, D. (2000). A minimum length of N gene sequence in transgenic plants is required for RNA-mediated tospovirus resistance. *Microbiology* 81, 235–242. doi: 10.1099/0022-1317-81-1-235.

Janzac, B., Fabre, M.-F., Palloix, A., and Moury, B. (2009). Phenotype and spectrum of action of the Pvr4 resistance in pepper against potyviruses, and selection for virulent variants. *Plant Pathol.* 58, 443–449. doi: 10.1111/j.1365-3059.2008.01992.x.

Janzac, B., Montarry, J., Palloix, A., Navaud, O., and Moury, B. (2010). A Point Mutation in the Polymerase of Potato virus Y Confers Virulence Toward the Pvr4 Resistance of Pepper and a High Competitiveness Cost in Susceptible Cultivar. *Mol. Plant-Microbe Interact.* 23, 823–830. doi: 10.1094/MPMI-23-6-0823.

Jiang, L., Huang, Y., Sun, L., Wang, B., Zhu, M., Li, J., et al. (2017a). Occurrence and diversity of Tomato spotted wilt virus isolates breaking the Tsw resistance gene of Capsicum chinense in Yunnan, southwest China. *Plant Pathol.* doi: 10.1111/ppa.12645.

Jiang, L., Huang, Y., Sun, L., Wang, B., Zhu, M., Li, J., et al. (2017b). Occurrence and diversity of Tomato spotted wilt virus isolates breaking the Tsw resistance gene of Capsicum chinense in Yunnan, southwest China. *Plant Pathol.* 66, 980–989. doi: 10.1111/ppa.12645.

Jiang, N., Meng, J., Cui, J., Sun, G., and Luan, Y. (2018). Function identification of miR482b, a negative regulator during tomato resistance to Phytophthora infestans. *Hortic. Res.* 5. doi: 10.1038/s41438-018-0017-2.

Jones, J. D. G. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286.

Jones, J. D. G., Vance, R. E., and Dangl, J. L. (2016). Intracellular innate immune surveillance devices in plants and animals. *Science (80-.).* 354. doi: 10.1126/science.aaf6395.

Jubic, L. M., Saile, S., Furzer, O. J., El Kasmi, F., and Dangl, J. L. (2019). Help wanted: helper NLRs and plant immune responses. *Curr. Opin. Plant Biol.* 50, 82–94. doi: 10.1016/j.pbi.2019.03.013.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nat. 2021 5967873* 596, 583–589. doi: 10.1038/s41586-021-03819-2.

K

Kajava, A. V. (1998). Structural diversity of leucine-rich repeat proteins 1 1Edited by F. Cohen. *J. Mol. Biol.* 277, 519–527. doi: 10.1006/jmbi.1998.1643.

Karasawa, A., Okada, I., Akashi, K., Chida, Y., Hase, S., Nakazawa-Nasu, Y., et al. (1999). One Amino Acid Change in Cucumber Mosaic Virus RNA Polymerase Determines Virulent/Avirulent Phenotypes on Cowpea. *Phytopathology®* 89, 1186–1192. doi: 10.1094/PHYTO.1999.89.12.1186.

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.* 7, 193–195. Available at: file:///C:/Users/Risa/Downloads/1-s2.0-S1360138502022513-main.pdf.

Kavanagh, T., Goulden, M., Santa Cruz, S., Chapman, S., Barker, I., and Baulcombe, D. (1992). Molecular analysis of a resistance-breaking strain of potato virus X. *Virology* 189, 609–617. doi: 10.1016/0042-6822(92)90584-C.

Kikkert, M., Van Lent, J., Storms, M., Bodegom, P., Kormelink, R., and Goldbach, R. (1999). Tomato Spotted Wilt Virus Particle Morphogenesis in Plant Cells. *J. Virol.* 73, 2288–2297. doi: 10.1128/JVI.73.3.2288-2297.1999.

Kim, H. S., Park, Y. H., Nam, H., Lee, Y. H. M., Song, K., Choi, C., et al. (2014a). Overexpression of the Brassica rapa transcription factor WRKY12 results in reduced soft rot symptoms caused by Pectobacterium carotovorum in Arabidopsis and Chinese cabbage. *Plant Biol.* 16, 973–981. doi: 10.1111/plb.12149.

Kim, S.-B. B., Lee, H.-Y. Y., Choi, E.-H. H., Park, E., Kim, J.-H. H., Moon, K.-B. B., et al. (2018). The Coiled-Coil and Leucine-Rich Repeat Domain of the Potyvirus Resistance Protein Pvr4 Has a Distinct Role in Signaling and Pathogen Recognition. *Mol. Plant-Microbe Interact.* 31, 906–913. doi: 10.1094/MPMI-12-17-0313-R.

Kim, S. H., Qi, D., Ashfield, T., Helm, M., and Innes, R. W. (2016). Using decoys to expand the recognition specificity of a plant disease resistance protein. *Science (80-.).* 351, 684–687. doi: 10.1126/science.aad3436.

Kim, S., Kang, W., Huy, H. N., Yeom, S., An, J., Kim, S., et al. (2017a). Divergent evolution of multiple virus-resistance genes from a progenitor in Capsicum spp. *New Phytol.* 213, 886–899. doi: 10.1111/nph.14177.

Kim, S., Park, J., Yeom, S.-I., Kim, Y.-M., Seo, E., Kim, K.-T., et al. (2017b). New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. *Genome Biol.* 18, 210. doi: 10.1186/s13059-017-1341-9.

Kim, S., Park, M., Yeom, S.-I., Kim, Y.-M., Lee, J. M., Lee, H.-A., et al. (2014b). Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nat. Genet.* 46, 270–278. doi: 10.1038/ng.2877.

A

Appendices

Kobe, B., and Kajava, A. V. (2001). The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* 11, 725–732. doi: 10.1016/S0959-440X(01)00266-4.

van Knippenberg, I., Goldbach, R., and Kormelink, R. (2002). Purified Tomato spotted wilt virus Particles Support Both Genome Replication and Transcription in Vitro. *Virology* 303, 278–286. doi: 10.1006/viro.2002.1632.

van Knippenberg, I., Goldbach, R., and Kormelink, R. (2004). In vitro transcription of Tomato spotted wilt virus is independent of translation. *J. Gen. Virol.* doi: 10.1099/vir.0.19767-0.

van Knippenberg, I., Goldbach, R., and Kormelink, R. (2005). Tomato spotted wilt virus S-segment mRNAs have overlapping 3′-ends containing a predicted stem-loop structure and conserved sequence motif. *Virus Res.* 110, 125–131. doi: 10.1016/j.virusres.2005.01.012.

Kohm, B. A., Goulden, M. G., Gilbert, J. E., Kavanagh, T. A., and Baulcombe, D. C. (1993). A Potato Virus X Resistance Gene Mediates an Induced, Nonspecific Resistance in Protoplasts. *Plant Cell* 5, 913. doi: 10.2307/3869659.

Komoda, K., Ishibashi, K., Kawamura-Nagaya, K., and Ishikawa, M. (2014). Possible involvement of eEF1A in Tomato spotted wilt virus RNA synthesis. *Virology* 468–470, 81–87. doi: 10.1016/j.virol.2014.07.053.

Komoda, K., Narita, M., Yamashita, K., Tanaka, I., and Yao, M. (2017). Asymmetric Trimeric Ring Structure of the Nucleocapsid Protein of Tospovirus. *J. Virol.* 91. doi: 10.1128/jvi.01002-17.

Kormelink, R. (2011). "The molecular biology of tospoviruses and resistance strategies," in *The Bunyaviridae: Molecular and Cellular Biology*, 163–191.

Kormelink, R., Garcia, M. L., Goodin, M., Sasaya, T., and Haenni, A.-L. (2011). Negative-strand RNA viruses: The plant-infecting counterparts. *Virus Res.* 162, 184–202. doi: 10.1016/j.virusres.2011.09.028.

Kormelink, R., Kitajima, E. W., Haan, P. De, Zuidema, D., Peters, D., and Goldbach, R. (1991). The nonstructural protein (NSS) encoded by the ambisense S RNA segment of tomao spotted wilt virus is associated with fibrous structures in infected plant cells. *Virology* 181, 459–468. doi: 10.1016/0042-6822(91)90878-F.

Kormelink, R., Storms, M., Van Lent, J., Peters, D., and Goldbach, R. (1994). Expression and Subcellular Location of the NSM Protein of Tomato Spotted Wilt Virus (TSWV), a Putative Viral Movement Protein. *Virology* 200, 56–65. doi: 10.1006/viro.1994.1162.

Kormelink, R., van Poelwijk, F., Peters, D., and Goldbach, R. (1992). Non-viral heterogeneous sequences at the 5′ ends of tomato spotted wilt virus mRNAs. *J. Gen. Virol.* 73, 2125–2128. doi: 10.1099/0022-1317-73-8-2125.

Korves, T., and Bergelson, J. (2004). A Novel Cost of R Gene Resistance in the Presence of Disease. *Am. Nat.* 163, 489–504. doi: 10.1086/382552.

Kozuki, T., Chikamori, K., Surleac, M. D., Micluta, M. A., Petrescu, A. J., Norris, E. J., et al. (2017). Roles of the C-terminal domains of topoisomerase IIα and topoisomerase IIβ in regulation of the decatenation checkpoint. *Nucleic Acids Res.* 45, 5995–6010. doi: 10.1093/nar/gkx325.

Kravchik, M., Sunkar, R., Damodharan, S., Stav, R., Zohar, M., Isaacson, T., et al. (2014). Global and local perturbation of the tomato microRNA pathway by a trans -activated DICER-LIKE 1 mutant. *J. Exp. Bot.* 65, 725–739. doi: 10.1093/jxb/ert428.

Krumov, V., and Karadjova, O. (2012). Influence of climate change on the potential for establishment of Frankliniella occidentalis (Thysanoptera: Thripidae) in Bulgaria. *Acta Phytopathol. Entomol. Hungarica* 47, 113–116. doi: 10.1556/APhyt.47.2012.1.14.

Kwon, S.-J., Cho, Y.-E., Kwon, O.-H., Kang, H.-G., and Seo, J.-K. (2021). Resistance-Breaking Tomato Spotted Wilt Virus Variant that Recently Occurred in Pepper in South Korea is a Genetic Reassortant. *Plant Dis.* 105, 2771–2775. doi: 10.1094/PDIS-01-21-0205-SC.

L

Laflamme, B., Dillon, M. M., Martel, A., Almeida, R. N. D., Desveaux, D., and Guttman, D. S. (2020). The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. *Science (80-.).* 367, 763–768. doi: 10.1126/science.aax4079.

Lakatos, L., Szittya, G., Silhavy, D., and Burgyán, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *EMBO J.* 23, 876–884. doi: 10.1038/sj.emboj.7600096.

Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J. U., and Forner, J. (2013). GreenGate - A novel, versatile, and efficient cloning system for plant transgenesis. *PLoS One* 8, e83043. doi: 10.1371/journal.pone.0083043.

Lan, H., Chen, H., Liu, Y., Jiang, C., Mao, Q., Jia, D., et al. (2016). Small Interfering RNA Pathway Modulates Initial Viral Infection in Midgut Epithelium of Insect after Ingestion of Virus. *J. Virol.* 90, 917–929. doi: 10.1128/JVI.01835-15.

Lanfermeijer, F. C., Warmink, J., and Hille, J. (2005). The products of the broken Tm-2 and the durable Tm-22 resistance genes from tomato differ in four amino acids. *J. Exp. Bot.* 56, 2925–2933. doi: 10.1093/jxb/eri288.

Latham, L. J., and Jones, R. A. C. (1998). Selection of resistance breaking strains of tomato spotted wilt tospovirus. *Ann. Appl. Biol.* 133, 385–402. doi: 10.1111/j.1744-7348.1998.tb05838.x.

Le May, N., Dubaele, S., De Santis, L. P., Billecocq, A., Bouloy, M., and Egly, J.-M. (2004). TFIIH Transcription Factor, a Target for the Rift Valley Hemorrhagic Fever Virus. *Cell* 116, 541–550. doi: 10.1016/S0092-8674(04)00132-1.

Leastro, M. O., De Oliveira, A. S., Pallás, V., Sánchez-Navarro, J. A., Kormelink, R., and Resende, R. O. (2017). The NSm proteins of phylogenetically related tospoviruses trigger Sw-5b–mediated resistance dissociated of their cell-to-cell movement function. *Virus Res.* 240, 25–34. doi: 10.1016/j.virusres.2017.07.019.

Lee, J., Miura, K., Bressan, R. A., Hasegawa, P. M., and Yun, D.-J. (2007). Regulation of Plant Innate Immunity by SUMO E3 Ligase. *Plant Signal. Behav.* 2, 253–254. doi: 10.4161/psb.2.4.3867.

Léger, P., Nachman, E., Richter, K., Tamietti, C., Koch, J., Burk, R., et al. (2020). NSs amyloid formation is associated with the virulence of Rift Valley fever virus in mice. *Nat. Commun.* 11, 3281. doi: 10.1038/s41467-020-17101-y.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van De Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325.

Lewandowski, D. J., and Adkins, S. (2005). The tubule-forming NSm protein from Tomato spotted wilt virus complements cell-to-cell and long-distance movement of Tobacco mosaic virus hybrids. *Virology* 342, 26–37. doi: 10.1016/j.virol.2005.06.050.

Li, F., Huang, C., Li, Z., and Zhou, X. (2014). Suppression of RNA Silencing by a Plant DNA Virus Satellite Requires a Host Calmodulin-Like Protein to Repress RDR6 Expression. *PLoS Pathog.* 10, e1003921. doi: 10.1371/journal.ppat.1003921.

Li, F., Pignatta, D., Bendix, C., Brunkard, J. O., Cohn, M. M., Tung, J., et al. (2012). MicroRNA regulation of plant innate immune receptors. *Proc. Natl. Acad. Sci.* 109, 1790–1795. doi: 10.1073/pnas.1118282109.

Li, J., Feng, Z., Wu, J., Huang, Y., Lu, G., Zhu, M., et al. (2015). Structure and Function Analysis of Nucleocapsid Protein of Tomato Spotted Wilt Virus Interacting with RNA Using Homology Modeling. *J. Biol. Chem.* 290, 3950–3961. doi: 10.1074/jbc.M114.604678.

Li, J., Huang, H., Zhu, M., Huang, S., Zhang, W., Dinesh-Kumar, S. P., et al. (2019a). A Plant Immune Receptor Adopts a Two-Step Recognition Mechanism to Enhance Viral Effector Perception. *Mol. Plant* 12, 248–262. doi: 10.1016/j.molp.2019.01.005.

Li, S., Zhu, X., Guan, Z., Huang, W., Zhang, Y., Kortekaas, J., et al. (2019b). NSs Filament Formation Is Important but Not Sufficient for RVFV Virulence In Vivo. *Viruses* 11, 834. doi: 10.3390/v11090834.

Li, W., Lewandowski, D. J., Hilf, M. E., and Adkins, S. (2009). Identification of domains of the Tomato spotted wilt virus NSm protein involved in tubule formation, movement and symptomatology. *Virology* 390, 110–121. doi: 10.1016/j.virol.2009.04.027.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., et al. (2014). Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* 46, 1220–1226. doi: 10.1038/ng.3117.

Liu, D., Shi, L., Han, C., Yu, J., Li, D., and Zhang, Y. (2012). Validation of Reference Genes for Gene Expression Studies in Virus-Infected Nicotiana benthamiana Using Quantitative Real-Time PCR. *PLoS One* 7, e46451. doi: 10.1371/journal.pone.0046451.

Liu, Y., Schiff, M., and Dinesh-Kumar, S. P. (2002). Virus-induced gene silencing in tomato. *Plant J.* 31, 777–786. doi: 10.1046/j.1365-313X.2002.01394.x.

Liu, Y., Schiff, M., and Dinesh-Kumar, S. P. (2004). Involvement of MEK1 MAPKK, NTF6 MAPK, WRKY/MYB transcription factors, COI1 and CTR1 in N -mediated resistance to tobacco mosaic virus. *Plant J.* 38, 800–809. doi: 10.1111/j.1365-313X.2004.02085.x.

Lopez, C., Aramburu, J., Galipienso, L., Soler, S., Nuez, F., and Rubio, L. (2011). Evolutionary analysis of tomato Sw-5 resistance-breaking isolates of Tomato spotted wilt virus. *J. Gen. Virol.* 92, 210–215. doi: 10.1099/vir.0.026708-0.

Lovato, F. A., Inoue-Nagata, A. K., Nagata, T., de Ávila, A. C., Pereira, L. A. R., and Resende, R. O. (2008). The N protein of Tomato spotted wilt virus (TSWV) is associated with the induction of programmed cell death (PCD) in Capsicum chinense plants, a hypersensitive host to TSWV infection. *Virus Res.* 137, 245–252. doi: 10.1016/j.virusres.2008.07.020.

Lu, J., Boeren, S., de Vries, S. C., van Valenberg, H. J. F., Vervoort, J., and Hettinga, K. (2011). Filter-aided sample preparation with dimethyl labeling to identify and quantify milk fat globule membrane proteins. *J. Proteomics* 75, 34–43. doi: 10.1016/j.jprot.2011.07.031.

Lv, T., Li, X., Fan, T., Luo, H., Xie, C., Zhou, Y., et al. (2019). The Calmodulin-Binding Protein IQM1 Interacts with CATALASE2 to Affect Pathogen Defense. *Plant Physiol.* 181, 1314–1327. doi: 10.1104/pp.19.01060.

M

Ma, S., Lapin, D., Liu, L., Sun, Y., Song, W., Zhang, X., et al. (2020). Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. Science (80-.). 370. doi: 10.1126/science.abe3069.

Macedo, M. A., Rojas, M. R., and Gilbertson, R. L. (2019). First Report of a Resistance-Breaking Strain of Tomato Spotted Wilt Orthotospovirus Infecting Sweet Pepper with the Tsw Resistance Gene in California, U.S.A. Plant Dis. 103, 1048. doi: 10.1094/PDIS-07-18-1239-PDN.

MacKenzie, D. J. (1992). Resistance to Tomato Spotted Wilt Virus Infection in Transgenic Tobacco Expressing the Viral Nucleocapsid Gene. *Mol. Plant-Microbe Interact.* 5, 34. doi: 10.1094/MPMI-5-034.

Maekawa, T., Cheng, W., Spiridon, L. N., Töller, A., Lukasik, E., Saijo, Y., et al. (2011). Coiled-Coil Domain-Dependent Homodimerization of Intracellular Barley Immune Receptors Defines a Minimal Functional Module for Triggering Cell Death. *Cell Host Microbe* 9, 187–199. doi: 10.1016/j.chom.2011.02.008.

Maes, P., Alkhovsky, S. V., Bào, Y., Beer, M., Birkhead, M., Briese, T., et al. (2018). Taxonomy of the family Arenaviridae and the order Bunyavirales: update 2018. *Arch. Virol.* 163, 2295–2310. doi: 10.1007/s00705-018-3843-5.

Maiti, S., Paul, S., and Pal, A. (2012). Isolation, Characterization, and Structure Analysis of a Non-TIR-NBS-LRR Encoding Candidate Gene from MYMIV-Resistant Vigna mungo. *Mol. Biotechnol.* 52, 217–233. doi: 10.1007/s12033-011-9488-1.

A

Margaria, P., Bosco, L., Vallino, M., Ciuffo, M., Mautino, G. C., Tavella, L., et al. (2014). The NSs Protein of Tomato spotted wilt virus Is Required for Persistent Infection and Transmission by Frankliniella occidentalis. *J. Virol.* 88, 5788–5802. doi: 10.1128/JVI.00079-14.

Margaria, P., Ciuffo, M., Pacifico, D., and Turina, M. (2007). Evidence That the Nonstructural Protein of Tomato spotted wilt virus Is the Avirulence Determinant in the Interaction with Resistant Pepper Carrying the Tsw Gene. *Mol. Plant-Microbe Interact.* 20, 547–558. doi: 10.1094/MPMI-20-5-0547.

Margaria, P., Ciuffo, M., Rosa, C., and Turina, M. (2015). Evidence of a tomato spotted wilt virus resistance-breaking strain originated through natural reassortment between two evolutionary-distinct isolates. *Virus Res.* 196, 157–161. doi: 10.1016/j.virusres.2014.11.012.

Martin, E. C., Sukarta, O. C. A. A., Spiridon, L., Grigore, L. G., Constantinescu, V., Tacutu, R., et al. (2020a). LRRpredictor—A New LRR Motif Detection Method for Irregular Motifs of Plant NLR Proteins Using an Ensemble of Classifiers. *Genes (Basel).* 11, 286. doi: 10.3390/genes11030286.

Martin, R., Qi, T., Zhang, H., Liu, F., King, M., Toth, C., et al. (2020b). Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. *Science (80-.).* 370, 1–21. doi: 10.1126/science.abd9993.

Masuta, C., Inaba, J. ichi, and Shimura, H. (2012). The 2b proteins of Cucumber mosaic virus generally have the potential to differentially induce necrosis on Arabidopsis. *Plant Signal. Behav.* 7, 43–45. doi: 10.4161/psb.7.1.18526.

Mátai, A., and Hideg, É. (2017). A comparison of colorimetric assays detecting hydrogen peroxide in leaf extracts. *Anal. Methods* 9, 2357–2360. doi: 10.1039/C7AY00126F.

Mathioudakis, M. M., Veiga, R. S. L. L., Canto, T., Medina, V., Mossialos, D., Makris, A. M., et al. (2013). Pepino mosaic virus triple gene block protein 1 (TGBp1) interacts with and increases tomato catalase 1 activity to enhance virus accumulation. *Mol. Plant Pathol.* 14, 589–601. doi: 10.1111/mpp.12034.

de Medeiros, R. B., Figueiredo, J., Resende, R. D. O., and De Avila, A. C. (2005). Expression of a viral polymerase-bound host factor turns human cell lines permissive to a plant- and insect-infecting virus. *Proc. Natl. Acad. Sci.* 102, 1175–1180. doi: 10.1073/pnas.0406668102.

Meyers, B. C., Dickerman, A. W., Michelmore, R. W., Sivaramakrishnan, S., Sobral, B. W., and Young, N. D. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.* 20, 317–332. doi: 10.1046/j.1365-313X.1999.t01-1-00606.x.

Meyers, B. C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R. W. (2003). Genome-Wide Analysis of NBS-LRR–Encoding Genes in Arabidopsis. *Plant Cell* 15, 809–834. doi: 10.1105/tpc.009308.

Milligan, S. B., Bodeau, J., Yaghoobi, J., Kaloshian, I., Zabel, P., and Williamson, V. M. (1998). The Root Knot Nematode Resistance Gene Mi from Tomato Is a Member of the Leucine Zipper, Nucleotide Binding, Leucine-Rich Repeat Family of Plant Genes. *Plant Cell* 10, 1307–1319. doi: 10.1105/tpc.10.8.1307.

Mir, M. A., Duran, W. A., Hjelle, B. L., Ye, C., and Panganiban, A. T. (2008). Storage of cellular 5′ mRNA caps in P bodies for viral cap-snatching. *Proc. Natl. Acad. Sci.* 105, 19294–19299. doi: 10.1073/pnas.0807211105.

Mitter, N., Koundal, V., Williams, S., and Pappu, H. (2013). Differential Expression of Tomato Spotted Wilt Virus-Derived Viral Small RNAs in Infected Commercial and Experimental Host Plants. *PLoS One* 8, e76276. doi: 10.1371/journal.pone.0076276.

Mlotshwa, S., Verver, J., Sithole-Niang, I., Prins, M., Kammen, A. V., and Wellink, J. (2002). Transgenic plants expressing HC-Pro show enhanced virus sensitivity while silencing of the transgene results in resistance. *Virus Genes* 25, 45–57. doi: 10.1023/A:1020170024713.

Moshe, A., Gorovits, R., Liu, Y., and Czosnek, H. (2016). Tomato plant cell death induced by inhibition of HSP90 is alleviated by Tomato yellow leaf curl virus infection. *Mol. Plant Pathol.* 17, 247–260. doi: 10.1111/mpp.12275.

Moury, B., Palloix, A., Gebre Selassie, K., and Marchoux, G. (1997). Hypersensitive resistance to tomato spotted wilt virus in three capsicum chinense accessions is controlled by a single gene and is overcome by virulent strains. *Euphytica* 94, 45–52. doi: 10.1023/A:1002997522379.

Mucyn, T. S., Clemente, A., Andriotis, V. M. E., Balmuth, A. L., Oldroyd, G. E. D., Staskawicz, B. J., et al. (2006). The Tomato NBARC-LRR Protein Prf Interacts with Pto Kinase in Vivo to Regulate Specific Plant Immunity. *Plant Cell* 18, 2792–2806. doi: 10.1105/tpc.106.044016.

Mudhasani, R., Tran, J. P., Retterer, C., Kota, K. P., Whitehouse, C. A., and Bavari, S. (2016). Protein Kinase R Degradation Is Essential for Rift Valley Fever Virus Infection and Is Regulated by SKP1-CUL1-F-box (SCF)FBXW11-NSs E3 Ligase. *PLOS Pathog.* 12, e1005437. doi: 10.1371/journal.ppat.1005437.

Murota, K., Shimura, H., Takeshita, M., and Masuta, C. (2017). Interaction between Cucumber mosaic virus 2b protein and plant catalase induces a specific necrosis in association with proteasome activity. *Plant Cell Rep.* 36, 37–47. doi: 10.1007/s00299-016-2055-2.

Mysore, K. S., and Ryu, C.-M. M. (2004). Nonhost resistance: how much do we know? *Trends Plant Sci.* 9, 97–104. doi: 10.1016/j.tplants.2003.12.005.

N

Nakagawa, T., Suziki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., et al. (2007). Improved Gateway Binary Vectors: High-Performance Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. *Biosci. Biotechnol. Biochem.* 71, 2095–2100. doi: 10.1271/bbb.70216.

Nakahara, K. S., Masuta, C., Yamada, S., Shimura, H., Kashihara, Y., Wada, T. S., et al. (2012). Tobacco calmodulin-like protein provides secondary defense by binding to and directing degradation of virus RNA silencing suppressors. *Proc. Natl. Acad. Sci.* 109, 10113–10118. doi: 10.1073/pnas.1201628109.

Nebenfür, A., Gallagher, L. A., Dunahay, T. G., Frohlick, J. A., Mazurkiewicz, A. M., Meehl, J. B., et al. (1999). Stop-and-Go Movements of Plant Golgi Stacks Are Mediated by the Acto-Myosin System. *Plant Physiol.* 121, 1127–1141. doi: 10.1104/pp.121.4.1127.

Ngou, B. P. M., Ahn, H.-K., Ding, P., and Jones, J. D. G. (2021). Mutual potentiation of plant immunity by cell-surface and intracellular receptors. *Nature* 592, 110–115. doi: 10.1038/s41586-021-03315-7.

Nicaise, V., Gallois, J.-L., Chafiai, F., Allen, L. M., Schurdi-Levraud, V., Browning, K. S., et al. (2007). Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in Arabidopsis thaliana. *FEBS Lett.* 581, 1041–1046. doi: 10.1016/j.febslet.2007.02.007.

NLRscape (2022). NLRscape - A Plant NLR Atlas.

Nohales, M.-Á., Molina-Serrano, D., Flores, R., and Daròs, J.-A. (2012). Involvement of the Chloroplastic Isoform of tRNA Ligase in the Replication of Viroids Belonging to the Family Avsunviroidae. *J. Virol.* 86, 8269–8276. doi: 10.1128/JVI.00629-12.

Nombela, G., Williamson, V. M., and Muñiz, M. (2003). The Root-Knot Nematode Resistance Gene Mi-1.2 of Tomato Is Responsible for Resistance Against the Whitefly Bemisia tabaci. *Mol. Plant-Microbe Interact.* 16, 645–649. doi: 10.1094/MPMI.2003.16.7.645.

O

Okuda, S., Fujita, S., Moretti, A., Hohmann, U., Doblas, V. G., Ma, Y., et al. (2020). Molecular mechanism for the recognition of sequence-divergent CIF peptides by the plant receptor kinases GSO1/SGN3 and GSO2. *Proc. Natl. Acad. Sci.* 117, 2693–2703. doi: 10.1073/pnas.1911553117.

Olaya, C., Adhikari, B., Raikhy, G., Cheng, J., and Pappu, H. R. (2019). Identification and localization of Tospovirus genus-wide conserved residues in 3D models of the nucleocapsid and the silencing suppressor proteins. *Virol. J.* 16, 7. doi: 10.1186/s12985-018-1106-4.

de Oliveira, V. C., da Silva Morgado, F., Ardisson-Araújo, D. M. P., Resende, R. O., and Ribeiro, B. M. (2015). The silencing suppressor (NSs) protein of the plant virus Tomato spotted wilt virus enhances heterologous protein expression and baculovirus pathogenicity in cells and lepidopteran insects. *Arch. Virol.* 160, 2873–2879. doi: 10.1007/s00705-015-2580-2.

de Oliveira, A. S., Koolhaas, I., Boiteux, L. S., Caldararu, O. F., Petrescu, A.-J., Oliveira Resende, R., et al. (2016). Cell death triggering and effector recognition by Sw-5 SD-CNL proteins from resistant and susceptible tomato isolines to Tomato spotted wilt virus. *Mol. Plant Pathol.* 17, 1442–1454. doi: 10.1111/mpp.12439.

Oliveira, V. C., Bartasson, L., de Castro, M. E. B., Corrêa, J. R., Ribeiro, B. M., and Resende, R. O. (2011). A silencing suppressor protein (NSs) of a tospovirus enhances baculovirus replication in permissive and semipermissive insect cell lines. *Virus Res.* 155, 259–267. doi: 10.1016/j.virusres.2010.10.019.

de Oliveira, A. S., Boiteux, L. S., Kormelink, R., and Resende, R. O. (2018). The Sw-5 Gene Cluster: Tomato Breeding and Research Toward Orthotospovirus Disease Control. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01055.

Oliver, J. E., and Whitfield, A. E. (2016). The Genus Tospovirus : Emerging Bunyaviruses that Threaten Food Security. *Annu. Rev. Virol.* 3, 101–124. doi: 10.1146/annurev-virology-100114-055036.

van Ooijen, G., Mayr, G., Kasiem, M. M. A., Albrecht, M., Cornelissen, B. J. C., and Takken, F. L. W. (2008). Structure–function analysis of the NB-ARC domain of plant disease resistance proteins. *J. Exp. Bot.* 59, 1383–1397. doi: 10.1093/jxb/ern045.

Ooms, G.; Hooykaas, P.J.J.; Van Veen, R.J.M.; Van Beelen, P.; Regensburg-Tuïnk, T.J.G.; Schilperoort, R.A. (1982) Octopine Ti-Plasmid Deletion Mutants of Agrobacterium Tumefaciens with Emphasis on the Right Side of the T-Region. *Plasmid*, 7, 15–29, doi:10.1016/0147-619X(82)90023-3.

P

Pang, S.-Z., Slightom, J. L., and Gonsalves, D. (1993). Different Mechanisms Protect Transgenic Tobacco Against Tomato Spotted Wilt and Impatiens Necrotic Spot Tospoviruses. *Nat. Biotechnol.* 11, 819–824. doi: 10.1038/nbt0793-819.

Parrella, G., Gognalons, P., Gebre-Selassiè, K., Vovlas, C., and Marchoux, G. (2003). An update of the host range of Tomato spotted wilt virus. *J. Plant Pathol.* 85, 227–264.

Peiró, A., Cañizares, M. C., Rubio, L., López, C., Moriones, E., Aramburu, J., et al. (2014). The movement protein (NSm) of Tomato spotted wilt virus is the avirulence determinant in the tomato Sw-5 gene-based resistance. *Mol. Plant Pathol.* 15, 802–813. doi: 10.1111/mpp.12142.

Peng, J.-C., Chen, T.-C., Raja, J. A. J., Yang, C.-F., Chien, W.-C., Lin, C.-H., et al. (2014). Broad-Spectrum Transgenic Resistance against Distinct Tospovirus Species at the Genus Level. *PLoS One* 9, e96073. doi: 10.1371/journal.pone.0096073.

Peyambari, M., Warner, S., Stoler, N., Rainer, D., and Roossinck, M. J. (2019). A 1,000-Year-Old RNA Virus. *J. Virol.* 93. doi: 10.1128/JVI.01188-18.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., et al. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802. doi: 10.1002/jcc.20289.

A

Appendices

Phillips, J. C., Hardy, D. J., Maia, J. D. C., Stone, J. E., Ribeiro, J. V., Bernardi, R. C., et al. (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* 153, 044130. doi: 10.1063/5.0014475.

Plotly Technologies Inc. (2015). Collaborative data science. *Plotly Technol. Inc., Montr. QC*.

van Poelwijk, F., Kolkman, J., and Goldbach, R. (1996). Sequence analysis of the 5′ends of tomato spotted wilt virus N mRNAs. *Arch. Virol.* 141, 177–184. doi: 10.1007/BF01718599.

Pozzer, L., Bezerra, I. C., Kormelink, R., Prins, M., Peters, D., Resende, R. de O., et al. (1999). Characterization of a Tospovirus Isolate of Iris Yellow Spot Virus Associated with a Disease in Onion Fields in Brazil. *Plant Dis.* 83, 345–350. doi: 10.1094/PDIS.1999.83.4.345.

Prasanth, K. R., Chuang, C., and Nagy, P. D. (2017). Co-opting ATP-generating glycolytic enzyme PGK1 phosphoglycerate kinase facilitates the assembly of viral replicase complexes. *PLOS Pathog.* 13, e1006689. doi: 10.1371/journal.ppat.1006689.

Prigigallo, M. I., Križnik, M., De Paola, D., Catalano, D., Gruden, K., Finetti-Sialer, M. M., et al. (2019). Potato Virus Y Infection Alters Small RNA Metabolism and Immune Response in Tomato. *Viruses* 11, 1100. doi: 10.3390/v11121100.

Prins, M. (1995). Broad Resistance to Tospoviruses in Transgenic Tobacco Plants Expressing Three Tospoviral Nucleoprotein Gene Sequences. *Mol. Plant-Microbe Interact.* 8, 85. doi: 10.1094/MPMI-8-0085.

Prins, M., Kikkert, M., Ismayadi, C., de Graauw, W., de Haan, P., and Goldbach, R. (1997). Characterization of RNA-mediated resistance to tomato spotted wilt virus in transgenic tobacco plants expressing NS(M) gene sequences. *Plant Mol. Biol.* 33, 235–243. doi: 10.1023/A:1005729808191.

Pruitt, R. N., Locci, F., Wanke, F., Zhang, L., Saile, S. C., Joe, A., et al. (2021). The EDS1–PAD4–ADR1 node mediates Arabidopsis pattern-triggered immunity. *Nature* 598, 495–499. doi: 10.1038/s41586-021-03829-0.

Q

Qi, D., de Young, B. J., Innes, R. W., DeYoung, B. J., and Innes, R. W. (2012). Structure-Function Analysis of the Coiled-Coil and Leucine-Rich Repeat Domains of the RPS5 Disease Resistance Protein. *Plant Physiol.* 158, 1819–1832. doi: 10.1104/pp.112.194035.

Qi, S., Shen, Y., Wang, X., Zhang, S., Li, Y., Islam, M. M., et al. (2022). A new NLR gene for resistance to Tomato spotted wilt virus in tomato (Solanum lycopersicum). *Theor. Appl. Genet.* 135, 1493–1509. doi: 10.1007/s00122-022-04049-4.

Qian, L., Zhao, J., Du, Y., Zhao, X., Han, M., and Liu, Y. (2018). Hsp90 Interacts With Tm-2² and Is Essential for Tm-2²-Mediated Resistance to Tobacco mosaic virus. *Front. Plant Sci.* 9, 1–8. doi: 10.3389/fpls.2018.00411.

R

Ramesh, S. V., Williams, S., Kappagantu, M., Mitter, N., and Pappu, H. R. (2017). Transcriptome-wide identification of host genes targeted by tomato spotted wilt virus-derived small interfering RNAs. *Virus Res.* 238, 13–23. doi: 10.1016/j.virusres.2017.05.014.

Reguera, J., Weber, F., and Cusack, S. (2010). Bunyaviridae RNA Polymerases (L-Protein) Have an N-Terminal, Influenza-Like Endonuclease Domain, Essential for Viral Cap-Dependent Transcription. *PLoS Pathog.* 6, e1001101. doi: 10.1371/journal.ppat.1001101.

Reitz, S. R., Gao, Y., Kirk, W. D. J., Hoddle, M. S., Leiss, K. A., and Funderburk, J. E. (2020). Invasion Biology, Ecology, and Management of Western Flower Thrips. *Annu. Rev. Entomol.* 65, 17–37. doi: 10.1146/annurev-ento-011019-024947.

Ren, T., Qu, F., and Morris, T. J. (2005). The nuclear localization of the Arabidopsis transcription factor TIP is blocked by its interaction with the coat protein of Turnip crinkle virus. *Virology* 331, 316–324. doi: 10.1016/j.virol.2004.10.039.

Ribeiro, D., Borst, J. W., Goldbach, R., and Kormelink, R. (2009a). Tomato spotted wilt virus nucleocapsid protein interacts with both viral glycoproteins Gn and Gc in planta. *Virology*. doi: 10.1016/j.virol.2008.09.028.

Ribeiro, D., Foresti, O., Denecke, J., Wellink, J., Goldbach, R., and Kormelink, R. J. M. (2008). Tomato spotted wilt virus glycoproteins induce the formation of endoplasmic reticulum- and Golgi-derived pleomorphic membrane structures in plant cells. *J. Gen. Virol.* doi: 10.1099/vir.0.2008/001164-0.

Ribeiro, D., Goldbach, R., and Kormelink, R. (2009b). Requirements for ER-Arrest and Sequential Exit to the Golgi of Tomato Spotted Wilt Virus Glycoproteins. *Traffic* 10, 664–672. doi: 10.1111/j.1600-0854.2009.00900.x.

Ribeiro, D., Jung, M., Moling, S., Borst, J. W., Goldbach, R., and Kormelink, R. (2013). The Cytosolic Nucleoprotein of the Plant-Infecting Bunyavirus Tomato Spotted Wilt Recruits Endoplasmic Reticulum–Resident Proteins to Endoplasmic Reticulum Export Sites. *Plant Cell* 25, 3602–3614. doi: 10.1105/tpc.113.114298.

Richard, M. M. S., Knip, M., Schachtschabel, J., Beijaert, M. S., and Takken, F. L. W. (2021). Perturbation of nuclear–cytosolic shuttling of Rx1 compromises extreme resistance and translational arrest of potato virus X transcripts. *Plant J.* 106, 468–479. doi: 10.1111/tpj.15179.

de Ronde, D. (Dryas) (2013). Analysis of Tomato spotted wilt virus effector-triggered immunity.

de Ronde, D., Butterbach, P., Lohuis, D., Hedil, M., van Lent, J. W. M. M., and Kormelink, R. (2013). Tsw gene-based resistance is triggered by a functional RNA silencing suppressor protein of the Tomato spotted wilt virus. *Mol. Plant Pathol.* 14, 405–415. doi: 10.1111/mpp.12016.

de Ronde, D., Butterbach, P., and Kormelink, R. (2014a). Dominant resistance against plant viruses. *Front. Plant Sci.* 5, 1–17. doi: 10.3389/fpls.2014.00307.

de Ronde, D., Pasquier, A., Ying, S., Butterbach, P., Lohuis, D., and Kormelink, R. (2014b). Analysis of Tomato spotted wilt virus NSs protein indicates the importance of the N-terminal domain for avirulence and RNA silencing suppression. *Mol. Plant Pathol.* 15, 185–195. doi: 10.1111/mpp.12082.

de Ronde, D., Lohuis, D., and Kormelink, R. (2019). Identification and characterization of a new class of Tomato spotted wilt virus isolates that break Tsw -based resistance in a temperature-dependent manner. *Plant Pathol.* 68, 60–71. doi: 10.1111/ppa.12952.

Roshan, P., Kulshreshtha, A., Kumar, S., Purohit, R., and Hallan, V. (2018). AV2 protein of tomato leaf curl Palampur virus promotes systemic necrosis in Nicotiana benthamiana and interacts with host Catalase2. *Sci. Rep.* 8, 1273. doi: 10.1038/s41598-018-19292-3.

Rotenberg, D., Jacobson, A. L., Schneweis, D. J., and Whitfield, A. E. (2015). Thrips transmission of tospoviruses. *Curr. Opin. Virol.* 15, 80–89. doi: 10.1016/j.coviro.2015.08.003.

Rouster, J., Leah, R., Mundy, J., and Cameron-Mills, V. (1997). Identification of a methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. *Plant J.* 11, 513–523. doi: 10.1046/j.1365-313X.1997.11030513.x.

S

Sanfaçon, H. (2015). Plant Translation Factors and Virus Resistance. *Viruses* 7, 3392–3419. doi: 10.3390/v7072778.

Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., et al. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641. doi: 10.1038/nature11119.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019.

Schnettler, E., Hemmes, H., Huismann, R., Goldbach, R., Prins, M., and Kormelink, R. (2010). Diverging Affinity of Tospovirus RNA Silencing Suppressor Proteins, NSs, for Various RNA Duplex Molecules. *J. Virol.* 84, 11542–11554. doi: 10.1128/JVI.00595-10.

Scholthof, K.-B. G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., et al. (2011). Top 10 plant viruses in molecular plant pathology. *Mol. Plant Pathol.* 12, 938–54. doi: 10.1111/j.1364-3703.2011.00752.x.

Schultink, A., Qi, T., Bally, J., and Staskawicz, B. (2019). Using forward genetics in Nicotiana benthamiana to uncover the immune signaling pathway mediating recognition of the Xanthomonas perforans effector XopJ4. *New Phytol.* 221, 1001–1009. doi: 10.1111/nph.15411.

Sekine, K.-T., Tomita, R., Takeuchi, S., Atsumi, G., Saitoh, H., Mizumoto, H., et al. (2012). Functional Differentiation in the Leucine-Rich Repeat Domains of Closely Related Plant Virus-Resistance Proteins That Recognize Common Avr Proteins. *Mol. Plant-Microbe Interact.* 25, 1219–1229. doi: 10.1094/MPMI-11-11-0289.

Senthil-Kumar, M., and Mysore, K. S. (2014). Tobacco rattle virus–based virus-induced gene silencing in Nicotiana benthamiana. *Nat. Protoc.* 9, 1549–1562. doi: 10.1038/nprot.2014.092.

Seo, E., Kim, T., Park, J. H., Yeom, S.-I. I., Kim, S., Seo, M.-K. K., et al. (2018). Genome-wide comparative analysis in Solanaceous species reveals evolution of microRNAs targeting defense genes in Capsicum spp. *DNA Res.* 25, 561–575. doi: 10.1093/dnares/dsy025.

Seo, Y.-S., Rojas, M. R., Lee, J.-Y., Lee, S.-W., Jeon, J.-S., Ronald, P., et al. (2006). A viral resistance gene from common bean functions across plant families and is up-regulated in a non-virus-specific manner. *Proc. Natl. Acad. Sci.* 103, 11856–11861. doi: 10.1073/pnas.0604815103.

Sha, A., Zhao, J., Yin, K., Tang, Y., Wang, Y., Wei, X., et al. (2014). Virus-Based MicroRNA Silencing in Plants. *Plant Physiol.* 164, 36–47. doi: 10.1104/pp.113.231100.

Sharma, M., Fuertes, D., Perez-Gil, J., and Lois, L. M. (2021a). SUMOylation in Phytopathogen Interactions: Balancing Invasion and Resistance. *Front. Cell Dev. Biol.* 9. doi: 10.3389/fcell.2021.703795.

Sharma, N., Sahu, P. P. P., Prasad, A., Muthamilarasan, M., Waseem, M., Khan, Y., et al. (2021b). The Sw5a gene confers resistance to ToLCNDV and triggers an HR response after direct AC4 effector recognition. *Proc. Natl. Acad. Sci. U. S. A.* 118. doi: 10.1073/pnas.2101833118.

Shen, Y., Zhao, X., Yao, M., Li, C., Miriam, K., Zhang, X., et al. (2014). A versatile complementation assay for cell-to-cell and long distance movements by cucumber mosaic virus based agro-infiltration. *Virus Res.* 190, 25–33. doi: 10.1016/j.virusres.2014.06.013.

Shi, J. L., Zai, W. S., Xiong, Z. L., Wan, H. J., and Wu, W. R. (2021). NB-LRR genes: characteristics in three Solanum species and transcriptional response to Ralstonia solanacearum in tomato. *Planta* 254, 96. doi: 10.1007/s00425-021-03745-7.

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., et al. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35, 441–443. doi: 10.1038/nbt.3833.

Shivaprasad, P. V, Chen, H.-M., Patel, K., Bond, D. M., Santos, B. A. C. M., and Baulcombe, D. C. (2012). A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* 24, 859–874. doi: 10.1105/tpc.111.095380.

Sillitoe, I., Bordin, N., Dawson, N., Waman, V. P., Ashford, P., Scholes, H. M., et al. (2021). CATH: increased structural coverage of functional space. *Nucleic Acids Res* 49, D266–D273. doi: 10.1093/nar/gkaa1079.

da Silva, L. L. P., Snapp, E. L., Denecke, J., Lippincott-Schwartz, J., Hawes, C., and Brandizzi, F. (2004). Endoplasmic Reticulum Export Sites and Golgi Bodies Behave as Single Mobile Secretory Units in Plant Cells. Plant Cell 16, 1753–1771. doi: 10.1105/tpc.022673.

A

Simons, J.F.; Persson, R.; Pettersson, R.F.; Branch, S. (1992) Association of the Nonstructural Protein NSS of Uukuniemi Virus with the 40S Ribosomal Subunit; *J Virol.* 66, 4233-4241. doi: 10.1128/JVI.66.7.4233-4241.1992.

Slootweg, E. J., Spiridon, L. N., Martin, E. C., Tameling, W. I. L., Townsend, P. D., Pomp, R., et al. (2018). Distinct Roles of Non-Overlapping Surface Regions of the Coiled-Coil Domain in the Potato Immune Receptor Rx1. *Plant Physiol.* 178, 1310–1331. doi: 10.1104/pp.18.00603.

Slootweg, E. J., Spiridon, L. N., Roosien, J., Butterbach, P., Pomp, R., Westerhof, L., et al. (2013). Structural determinants at the interface of the ARC2 and leucine-rich repeat domains control the activation of the plant immune receptors Rx1 and Gpa2. *Plant Physiol.* 162, 1510–1528. doi: 10.1104/pp.113.218842.

Slootweg, E., Koropacka, K., Roosien, J., Dees, R., Overmars, H., Lankhorst, R. K., et al. (2017). Sequence Exchange between Homologous NB-LRR Genes Converts Virus Resistance into Nematode Resistance, and Vice Versa. *Plant Physiol.* 175, 498–510. doi: 10.1104/pp.17.00485.

Soellick, T.-R., Uhrig, J. F., Bucher, G. L., Kellmann, J.-W., and Schreier, P. H. (2000). The movement protein NSm of tomato spotted wilt tospovirus (TSWV): RNA binding, interaction with the TSWV N protein, and identification of interacting plant proteins. *Proc. Natl. Acad. Sci.* 97, 2373–2378. doi: 10.1073/pnas.030548397.

Sonoda, S., and Tsumuki, H. (2004). Analysis of RNA-mediated virus resistance by NSs and NSm gene sequences from Tomato spotted wilt virus. *Plant Sci.* 166, 771–778. doi: 10.1016/j.plantsci.2003.11.018.

Spassova, M. I., Prins, T. W., Folkertsma, R. T., Klein-Lankhorst, R. M., Hille, J., Goldbach, R. W., et al. (2001). The tomato gene Sw5 is a member of the coiled coil, nucleotide binding, leucine-rich repeat class of plant resistance genes and confers resistance to TSWV in tobacco. *Mol. Breed.* doi: 10.1023/A:1011363119763.

Srinivasan, R., Abney, M. R., Culbreath, A. K., Kemerait, R. C., Tubbs, R. S., Monfort, W. S., et al. (2017). Three decades of managing Tomato spotted wilt virus in peanut in southeastern United States. *Virus Res.* 241, 203–212. doi: 10.1016/j.virusres.2017.05.016.

Stavolone, L., and Lionetti, V. (2017). Extracellular Matrix in Plants and Animals: Hooks and Locks for Viruses. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.01760.

Stefano, G., Hawes, C., and Brandizzi, F. (2014). ER – the key to the highway. *Curr. Opin. Plant Biol.* 22, 30–38. doi: 10.1016/j.pbi.2014.09.001.

Stefano, G., Renna, L., Chatre, L., Hanton, S. L., Moreau, P., Hawes, C., et al. (2006). In tobacco leaf epidermal cells, the integrity of protein export from the endoplasmic reticulum and of ER export sites depends on active COPI machinery. *Plant J.* doi: 10.1111/j.1365-313X.2006.02675.x.

Storms, M. M. H., Kormelink, R., Peters, Dickvan Lent, J. W. M., and Goldbach, R. W. (1995). The Nonstructural NSm Protein of Tomato Spotted Wilt Virus Induces Tubular Structures in Plant and Insect Cells. *Virology* 214, 485–493. doi: 10.1006/viro.1995.0059.

Storms, M. M. H., Va, C., De, N., Schoot, R., Prins, M., Kormelink, R., et al. (2002). A comparison of two methods of microinjection for assessing altered plasmodesmal gating in tissues expressing viral movement proteins. *Plant J.* 13, 131–140. doi: 10.1046/j.1365-313X.1998.00007.x.

Sukarta, O. C. A., Slootweg, E. J., and Goverse, A. (2016). Structure-informed insights for NLR functioning in plant immunity. *Semin. Cell Dev. Biol.* 56, 134–149. doi: 10.1016/j.semcdb.2016.05.012.

Sun, Y., Li, L., Macho, A. P., Han, Z., Hu, Z., Zipfel, C., et al. (2013). Structural basis for flg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex. *Science* 342, 624–8. doi: 10.1126/science.1243825.

Swanepoel, R., and Blackburn, N. K. (1977). Demonstration of Nuclear Immunofluorescence in Rift Valley Fever Infected Cells. *J. Gen. Virol.* 34, 557–561. doi: 10.1099/0022-1317-34-3-557.

Szcześniak, M. W., and Makałowska, I. (2014). miRNEST 2.0: a database of plant and animal microRNAs. *Nucleic Acids Res.* 42, D74–D77. doi: 10.1093/nar/gkt1156.

T

Takahashi, F., Mizoguchi, T., Yoshida, R., Ichimura, K., and Shinozaki, K. (2011). Calmodulin-Dependent Activation of MAP Kinase for ROS Homeostasis in Arabidopsis. *Mol. Cell* 41, 649–660. doi: 10.1016/j.molcel.2011.02.029.

Takahashi, H., Miller, J., Nozaki, Y., Takeda, M., Shah, J., Hase, S., et al. (2002). RCY1 , an Arabidopsis thaliana RPP8/HRT family resistance gene, conferring resistance to cucumber mosaic virus requires salicylic acid, ethylene and a novel signal transduction mechanism. *Plant J.* 32, 655–667. doi: 10.1046/j.1365-313X.2002.01453.x.

Takahashi, H., Tabara, M., Miyashita, S., Ando, S., Kawano, S., Kanayama, Y., et al. (2022). Cucumber Mosaic Virus Infection in Arabidopsis: A Conditional Mutualistic Symbiont? *Front. Microbiol.* 12, 1–25. doi: 10.3389/fmicb.2021.770925.

Takeda, A., Sugiyama, K., Nagano, H., Mori, M., Kaido, M., Mise, K., et al. (2002). Identification of a novel RNA silencing suppressor, NSs protein of Tomato spotted wilt virus. *FEBS Lett.* 532, 75–79. doi: 10.1016/S0014-5793(02)03632-3.

Takken, F. L. W., and Tameling, W. I. L. (2009). To Nibble at Plant Resistance Proteins. *Science (80-.).* 324, 744–745. doi: 10.1126/science.1171666.

Taller, D., Bálint, J., Gyula, P., Nagy, T., Barta, E., Baksa, I., et al. (2018). Expansion of Capsicum annum fruit is linked to dynamic tissue-specific differential expression of miRNA and siRNA profiles. *PLoS One* 13, e0200207. doi: 10.1371/journal.pone.0200207.

Tameling, W. I. L., and Baulcombe, D. C. (2007). Physical Association of the NB-LRR Resistance Protein Rx with a Ran GTPase–Activating Protein Is Required for Extreme Resistance to Potato virus X. *Plant Cell* 19, 1682–1694. doi: 10.1105/tpc.107.050880.

Tao, H., Jia, Z., Gao, X., Gui, M., Li, Y., and Liu, Y. (2022). Analysis of the miRNA expression profile involved in the tomato spotted wilt orthotospovirus–pepper interaction. *Virus Res.* 312, 198710. doi: 10.1016/j.virusres.2022.198710.

Tareen, A., and Kinney, J. B. (2020). Logomaker: beautiful sequence logos in Python. *Bioinformatics* 36, 2272–2274. doi: 10.1093/bioinformatics/btz921.

The PyMOL Molecular Graphics System, Version 2.2.3 Schrödinger, LLC.

Thomas, D., Blakqori, G., Wagner, V., Banholzer, M., Kessler, N., Elliott, R. M., et al. (2004). Inhibition of RNA polymerase II phosphorylation by a viral interferon antagonist. *J. Biol. Chem.* 279, 31471–31477. doi: 10.1074/jbc.M400938200.

Thompson, G. J., and van Zijl, J. J. B. (1996). Control of Tomato Spotted Wilt Virus in Tomatoes in South Africa. *Acta Hortic.*, 379–384. doi: 10.17660/ActaHortic.1996.431.32.

Tian, F., Yang, D.-C. C., Meng, Y.-Q. Q., Jin, J., and Gao, G. (2019). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.* 48, D1104–D1113. doi: 10.1093/nar/gkz1020.

Tian, H., Wu, Z., Chen, S., Ao, K., Huang, W., Yaghmaiean, H., et al. (2021). Activation of TIR signalling boosts pattern-triggered immunity. *Nature* 598, 500–503. doi: 10.1038/s41586-021-03987-1.

Tomita, R., Sekine, K.-T., Mizumoto, H., Sakamoto, M., Murai, J., Kiba, A., et al. (2011). Genetic basis for the hierarchical interaction between Tobamovirus spp. and L resistance gene alleles from different pepper species. *Mol. Plant. Microbe. Interact.* 24, 108–117. doi: 10.1094/MPMI-06-10-0127.

Turina, M., Kormelink, R., and Resende, R. O. (2016). Resistance to Tospoviruses in Vegetable Crops: Epidemiological and Molecular Aspects. *Annu. Rev. Phytopathol.* 54, 347–371. doi: 10.1146/annurev-phyto-080615-095843.

U

Underwood, W. (2012). The Plant Cell Wall: A Dynamic Barrier Against Pathogen Invasion. *Front. Plant Sci.* 3. doi: 10.3389/fpls.2012.00085.

V

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, research0034.1. doi: 10.1186/gb-2002-3-7-research0034.

Venkatesh, J., and Kang, B.-C. C. (2019). Current views on temperature-modulated R gene-mediated plant defense responses and tradeoffs between plant growth and immunity. *Curr. Opin. Plant Biol.* 50, 9–17. doi: 10.1016/j.pbi.2019.02.002.

Vidal, S., Eriksson, A. R. B. B., Montesano, M., Denecke, J., Palva, E. T., and Tapio Palva, E. (1998). Cell wall-degrading enzymes from Erwinia carotovora cooperate in the salicylic acid-independent induction of a plant defense response. *Mol. Plant-Microbe Interact.* 11, 23–32. doi: 10.1094/MPMI.1998.11.1.23.

Voeltz, G. K., Prinz, W. A., Shibata, Y., Rist, J. M., and Rapoport, T. A. (2006). A Class of Membrane Proteins Shaping the Tubular Endoplasmic Reticulum. *Cell* 124, 573–586. doi: 10.1016/j.cell.2005.11.047.

Voinnet, O., and Baulcombe, D. (1997). Systemic signalling in gene silencing. *Nature* 389, 553.

Vos, P., Simons, G., Jesse, T., Wijbrandi, J., Heinen, L., Hogers, R., et al. (1998). The tomato Mi-1 gene confers resistance to both root-knot nematodes and potato aphids. *Nat. Biotechnol.* 16, 1365–1369. doi: 10.1038/4350.

van der Vossen, E. A. G., van der Voort, J. N. A. M. R., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D. C., et al. (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant J.* 23, 567–576. doi: 10.1046/j.1365-313x.2000.00814.x.

van der Vossen, E. A. G., Gros, J., Sikkema, A., Muskens, M., Wouters, D., Wolters, P., et al. (2005). The Rpi-blb2 gene from Solanum bulbocastanum is an Mi-1 gene homolog conferring broad-spectrum late blight resistance in potato. *Plant J.* 44, 208–222. doi: 10.1111/j.1365-313X.2005.02527.x.

Vossen, J. H., van Arkel, G., Bergervoet, M., Jo, K.-R. R., Jacobsen, E., and Visser, R. G. F. F. (2016). The Solanum demissum R8 late blight resistance gene is an Sw-5 homologue that has been deployed worldwide in late blight resistant varieties. *Theor. Appl. Genet.* 129, 1785–1796. doi: 10.1007/s00122-016-2740-0.

de Vries, S., Kukuk, A., von Dahlen, J. K., Schnake, A., Kloesges, T., and Rose, L. E. (2018). Expression profiling across wild and cultivated tomatoes supports the relevance of early miR482/2118 suppression for Phytophthora resistance. *Proc. R. Soc. B Biol. Sci.* 285, 20172560. doi: 10.1098/rspb.2017.2560.

W

Wang, G., Roux, B., Feng, F., Guy, E., Li, L., Li, N., et al. (2015a). The Decoy Substrate of a Pathogen Effector and a Pseudokinase Specify Pathogen-Induced Modified-Self Recognition and Immunity in Plants. *Cell Host Microbe* 18, 285–295. doi: 10.1016/j.chom.2015.08.004.

Wang, J. J., Hu, M., Wang, J. J., Qi, J., Han, Z., Wang, G., et al. (2019a). Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science (80-.).* 364, 1–11. doi: 10.1126/science.aav5870.

Wang, J., Li, H., Han, Z., Zhang, H., Wang, T., Lin, G., et al. (2015b). Allosteric receptor activation by the plant peptide hormone phytosulfokine. *Nature* 525, 265–268. doi: 10.1038/nature14858.

Wang, J., Wang, J., Hu, M., Wu, S., Qi, J., Wang, G., et al. (2019b). Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science (80-.).* 364. doi: 10.1126/science.aav5868.

Wang, W., Chen, L., Fengler, K., Bolar, J., Llaca, V., Wang, X., et al. (2021). A giant NLR gene confers broad-spectrum resistance to Phytophthora sojae in soybean. *Nat. Commun.* 12, 6263. doi: 10.1038/s41467-021-26554-8.

Wang, Z., Hardcastle, T. J., Canto Pastor, A., Yip, W. H., Tang, S., and Baulcombe, D. C. (2018). A novel DCL2-dependent miRNA pathway in tomato affects susceptibility to RNA viruses. *Genes Dev.* 32, 1155–1160. doi: 10.1101/gad.313601.118.

Webb, B., and Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* 54, 5.6.1-5.6.37. doi: 10.1002/cpbi.3.

Weber, E., Gruetzner, R., Werner, S., Engler, C., and Marillonnet, S. (2011). Assembly of designer tal effectors by golden gate cloning. *PLoS One* 6. doi: 10.1371/journal.pone.0019722.

Weber, F., Bridgen, A., Fazakerley, J. K., Streitenfeld, H., Kessler, N., Randall, R. E., et al. (2002). Bunyamwera Bunyavirus Nonstructural Protein NSs Counteracts the Induction of Alpha/Beta Interferon. *J. Virol.* 76, 7949–7955. doi: 10.1128/JVI.76.16.7949-7955.2002.

Wen, R.-H., Khatabi, B., Ashfield, T., Maroof, M. A. S., and Hajimorad, M. R. (2013). The HC-Pro and P3 Cistrons of an Avirulent Soybean mosaic virus Are Recognized by Different Resistance Genes at the Complex Rsv1 Locus. *Mol. Plant-Microbe Interact.* 26, 203–215. doi: 10.1094/MPMI-06-12-0156-R.

van Wersch, S., and Li, X. (2019). Stronger When Together: Clustering of Plant NLR Disease resistance Genes. *Trends Plant Sci.* 24, 688–699. doi: 10.1016/j.tplants.2019.05.005.

Whitfield, A. E., Ullman, D. E., and German, T. L. (2005). Tospovirus-Thrips Interactions. *Annu. Rev. Phytopathol.* 43, 459–489. doi: 10.1146/annurev.phyto.43.040204.140017.

Widana Gamage, S. M. K., and Dietzgen, R. G. (2017). Intracellular Localization, Interactions and Functions of Capsicum Chlorosis Virus Proteins. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00612.

Wijkamp, I., van Lent, J., Kormelink, R., Goldbach, R., and Peters, D. (1993). Multiplication of tomato spotted wilt virus in its insect vector, Frankliniella occidentalis. *J. Gen. Virol.* 74, 341–349. doi: 10.1099/0022-1317-74-3-341.

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* 27, 293–315. doi: 10.1002/pro.3330.

Williams, S. J., Sohn, K. H., Wan, L., Bernoux, M., Sarris, P. F., Segonzac, C., et al. (2014). Structural Basis for Assembly and Function of a Heterodimeric Plant Immune Receptor. *Science (80-.).* 344, 299–303. doi: 10.1126/science.1247357.

Williams, S. J., Sornaraj, P., DeCourcy-Ireland, E., Menz, R. I., Kobe, B., Ellis, J. G., et al. (2011). An Autoactive Mutant of the M Flax Rust Resistance Protein Has a Preference for Binding ATP, Whereas Wild-Type M Protein Binds ADP. *Mol. Plant-Microbe Interact.* 24, 897–906. doi: 10.1094/MPMI-03-11-0052.

Williams, L. V., Lambertini, P. M. L., Shohara, K., and Biderbost, E. B. (2001). Occurrence and Geographical Distribution of Tospovirus Species Infecting Tomato Crops in Argentina. *Plant Dis.* 85, 1227–1229. doi: 10.1094/PDIS.2001.85.12.1227.

Wróblewski, T., Spiridon, L., Martin, E. C., Petrescu, A.-J. J., Cavanaugh, K., Truco, M. J., et al. (2018). Genome-wide functional analyses of plant coiled–coil NLR-type pathogen receptors reveal essential roles of their N-terminal domain in oligomerization, networking, and immunity. *PLOS Biol.* 16, e2005821. doi: 10.1371/journal.pbio.2005821.

Wu, C.-H., Abd-El-Haliem, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., et al. (2017). NLR network mediates immunity to diverse plant pathogens. *Proc. Natl. Acad. Sci.* 114, 8113–8118. doi: 10.1073/pnas.1702041114.

Wu, C., Belhaj, K., Bozkurt, T. O., Birk, M. S., and Kamoun, S. (2016). Helper NLR proteins NRC2a/b and NRC3 but not NRC1 are required for Pto-mediated cell death and resistance in Nicotiana benthamiana. *New Phytol.* 209, 1344–1352. doi: 10.1111/nph.13764.

Wu, X., Xu, S., Zhao, P., Zhang, X., Yao, X., Sun, Y., et al. (2019). The Orthotospovirus nonstructural protein NSs suppresses plant MYC-regulated jasmonate signaling leading to enhanced vector attraction and performance. *PLOS Pathog.* 15, e1007897. doi: 10.1371/journal.ppat.1007897.

X

Xia, S., Cheng, Y. T., Huang, S., Win, J., Soards, A., Jinn, T.-L., et al. (2013). Regulation of Transcription of Nucleotide-Binding Leucine-Rich Repeat-Encoding Genes SNC1 and RPP4 via H3K4 Trimethylation. *PLANT Physiol.* 162, 1694–1705. doi: 10.1104/pp.113.214551.

Xu, D., and Nussinov, R. (1998). Favorable domain size in proteins. *Fold. Des.* 3, 11–17. doi: 10.1016/S1359-0278(98)00004-2.

Y

Yang, C.-F., Chen, K.-C., Cheng, Y.-H., Raja, J. A. J., Huang, Y.-L., Chien, W.-C., et al. (2015). Generation of Marker-free Transgenic Plants Concurrently Resistant to a DNA Geminivirus and a RNA Tospovirus. *Sci. Rep.* 4, 5717. doi: 10.1038/srep05717.

Yang, T., Qiu, L., Huang, W., Xu, Q., Zou, J., Peng, Q., et al. (2020). Chilli veinal mottle virus HCPro interacts with catalase to facilitate virus infection in Nicotiana tabacum. *J. Exp. Bot.* 71, 5656–5668. doi: 10.1093/jxb/eraa304.

Yang, X., Wen, Z., Zhang, D., Li, Z., Li, D., Nagalakshmi, U., et al. (2021). Proximity labeling: an emerging tool for probing in planta molecular interactions. *Plant Commun.* 2, 100137. doi: 10.1016/j.xplc.2020.100137.

Yang, Z., and Li, Y. (2018). Dissection of RNAi-based antiviral immunity in plants. *Curr. Opin. Virol.* 32, 88–99. doi: 10.1016/j.coviro.2018.08.003.

Yao, M., Zhang, T., Zhou, T., Zhou, Y., Zhou, X., and Tao, X. (2012). Repetitive prime-and-realign mechanism converts short capped RNA leaders into longer ones that may be more suitable for elongation during rice stripe virus transcription initiation. *J. Gen. Virol.* 93, 194–202. doi: 10.1099/vir.0.033902-0.

Yazhisai, U., Rajagopalan, P. A., Raja, J. A. J., Chen, T.-C., and Yeh, S.-D. (2015). Untranslatable tospoviral NSs fragment coupled with L conserved region enhances transgenic resistance against the homologous virus and a serologically unrelated tospovirus. *Transgenic Res.* 24, 635–649. doi: 10.1007/s11248-015-9865-9.

Yin, K., Gao, C., and Qiu, J.-L. (2017). Progress and prospects in plant genome editing. *Nat. Plants* 3, 17107. doi: 10.1038/nplants.2017.107.

Yuan, M., Jiang, Z., Bi, G., Nomura, K., Liu, M., Wang, Y., et al. (2021). Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature* 592, 105–109. doi: 10.1038/s41586-021-03316-6.

Yuan, P., Bartlam, M., Lou, Z., Chen, S., Zhou, J., He, X., et al. (2009). Crystal structure of an avian influenza polymerase PAN reveals an endonuclease active site. *Nature* 458, 909–913. doi: 10.1038/nature07720.

Z

Zaccardelli, M., Perrone, D., Del Galdo, A., Campanile, F., Parrella, G., and Giordano, I. (2008). Tomato Genotypes Resistant to Tomato Spotted Wilt Virus Evaulated in Open Field Crops in Southern Italy. *Acta Hortic.*, 147–150. doi: 10.17660/ActaHortic.2008.789.20.

Zhai, Y., Gnanasekaran, P., and Pappu, H. R. (2021). Identification and Characterization of Plant-Interacting Targets of Tomato Spotted Wilt Virus Silencing Suppressor. *Pathogens* 10, 27. doi: 10.3390/pathogens10010027.

Zhan, J., Shi, H., Li, W., Zhang, C., and Zhang, Y. (2021). NbTMP14 Is Involved in Tomato Spotted Wilt Virus Infection and Symptom Development by Interaction with the Viral NSm Protein. *Viruses* 13, 427. doi: 10.3390/v13030427.

Zhang, H., Zhang, J., Lang, Z., Botella, J. R., and Zhu, J.-K. (2017). Genome Editing—Principles and Applications for Functional Genomics Research and Crop Improvement. *CRC. Crit. Rev. Plant Sci.* 36, 291–309. doi: 10.1080/07352689.2017.1402989.

Zhang, L., Chen, S., Ruan, J., Wu, J., Tong, A. B., Yin, Q., et al. (2015). Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization. *Science (80-.).* 350, 404–409. doi: 10.1126/science.aac5789.

Zhang, X., Yuan, Y. R., Pei, Y., Lin, S. S., Tuschl, T., Patel, D. J., et al. (2006). Cucumber mosaic virus-encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. *Genes Dev.* 20, 3255–3268. doi: 10.1101/gad.1495506.

Zhang, Y., Cheng, T. C., Huang, G., Lu, Q., Surleac, M. D., Mandell, J. D., et al. (2019). Transposon molecular domestication and the evolution of the RAG recombinase. *Nature* 569, 79–84. doi: 10.1038/s41586-019-1093-7.

Zhang, Z., Yu, J., Li, D., Zhang, Z., Liu, F., Zhou, X., et al. (2010). PMRD: plant microRNA database. *Nucleic Acids Res.* 38, D806–D813. doi: 10.1093/nar/gkp818.

Zhao, J.-H., Liu, X.-L., Fang, Y.-Y., Fang, R.-X., and Guo, H.-S. (2018). CMV2b-Dependent Regulation of Host Defense Pathways in the Context of Viral Infection. *Viruses* 10, 618. doi: 10.3390/v10110618.

Zhao, W., Jiang, L., Feng, Z., Chen, X., Huang, Y., Xue, F., et al. (2016). Plasmodesmata targeting and intercellular trafficking of Tomato spotted wilt tospovirus movement protein NSm is independent of its function in HR induction. *J. Gen. Virol.* 97, 1990–1997. doi: 10.1099/jgv.0.000496.

Zhao, X., Chen, Z., Wu, Q., Cai, Y., Zhang, Y., Zhao, R., et al. (2021). The Sw-5b NLR nucleotide-binding domain plays a role in oligomerization, and its self-association is important for activation of cell death signaling. *J. Exp. Bot.* 72, 6581–6595. doi: 10.1093/jxb/erab279.

Zhou, M., Li, Y., Hu, Q., Bai, X., Huang, W., Yan, C., et al. (2015). Atomic structure of the apoptosome: mechanism of cytochrome c - and dATP-mediated activation of Apaf-1. *Genes Dev.* 29, 2349–2361. doi: 10.1101/gad.272278.115.

Zhu, M., Jiang, L., Bai, B., Zhao, W., Chen, X., Li, J., et al. (2017). The Intracellular Immune Receptor Sw-5b Confers Broad-Spectrum Resistance to Tospoviruses through Recognition of a Conserved 21-Amino Acid Viral Effector Epitope. *Plant Cell* 29, 2214–2232. doi: 10.1105/tpc.17.00180.

Zhu, M., van Grinsven, I. L., Kormelink, R., and Tao, X. (2019). Paving the Way to Tospovirus Infection: Multilined Interplays with Plant Innate Immunity. *Annu. Rev. Phytopathol.* 57, 41–62. doi: 10.1146/annurev-phyto-082718-100309.

Zhu, T., Zhou, X., Zhang, J., Zhang, W., Zhang, L., You, C., et al. (2022). Ethylene-induced NbMYB4L is involved in resistance against tobacco mosaic virus in Nicotiana benthamiana. *Mol. Plant Pathol.* 23, 16–31. doi: 10.1111/mpp.13139.

Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends Immunol.* 35, 345–351. doi: 10.1016/j.it.2014.05.004.

Zou, Y., Wang, S., and Lu, D. (2020). MiR172b-TOE1/2 module regulates plant innate immunity in an age-dependent manner. *Biochem. Biophys. Res. Commun.* 531, 503–507. doi: 10.1016/j.bbrc.2020.07.061.

A

List of abbreviations

aa	amino acid
ADP	adenosine diphosphate
AGO	Argonaute protein
Amp R	ampicillin resistance
AMV	Alfalfa mosaic virus
ANOVA	analysis of variance
ARF	auxin response factor
ATP	adenosine triphosphate
ATTA	*Agrobacterium tumefaciencs* transcient expression assay
Avr	avirulence factor
BAC library	bacterial artificial chromosome library
BCA	bicinchoninic acid
BiFC	bimolecular fluorescence complementation
bp	base pair
C-terminus	carboxyl-terminus
CaCV	Capsicum chlorosis virus
CARE	cis-acting regulatory elements
Cat1	Catalase 1
CC	coiled-coil
cDNA	complementary DNA
CDS	coding sequence
CFP	cyan fluorescent protein
CHEF	clamped homogeneous electric field
CMV	Cucumber mosaic virus
CNL	CC-NLR
Co-IP	co-Immunoprecipitation
CP	coat protein
CRISPR	clustered regularly interspaced short palindromic repeat
cryo-EM	cryogenic electronic microscopy
DAB	3,3'-diaminobenzidine
DCL	Dicer-like
dpa	days post agro-infiltration
DPE	downstream promoter element
dpi	days post infection
dsRNA	double-stranded RNA
EF-Tu	elongation factor Tu
eF1α	eukaryotic elongation factor 1A
ER	endoplasmic reticulum
ERES	ER export sites

ERF	ethylene responsive factor
ETI	effector-triggered immunity
FoTF	*Frankiniella occidentalis* transcription factor
Fw	forward
GAPC1	cytosolic glyceraldehyde-3-phosphate dehydrogenase
Gc	glycoprotein C-terminal
gDNA	genomic DNA
GFP	green fluorescent protein
Gn	glycoprotein N-terminal
GO	gene ontology
GP	glycoprotein
GRSV	Groundnut ringspot virus
GUS	Beta-glucuronidase
HC-Pro	helper component proteinase
HR	hypersensitive response
HSP	heat-shock protein
IGR	intergenic region
Inr	Initiator
INSV	Impatiens necrotic spot virus
JA	jasmonate
Kana R	Kanamycin resistance
kDa	kilodalton
L	left border
LB	Luria Broth medium
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LRR	leucine-rich repeat
MAMP	microbe-associated molecular patterns
MAPK	mitogen-activated protein kinase
MeJA	methyl jasmonate
miRNA	micro RNA
MMAi	Murashige-Skoog induction medium
MP	movement protein
mRNA	messenger RNA
MYMIV	Mungbean yellow virus
N protein	nucleocapsid protein
N-terminus	amino terminus
NB-ARC	nucleotide binding Apaf-1, R-protein, and CED-4 domain
NBS	nucleotide binding domain
NHR	non-host resistance
NL	NBS-LRR receptors

A

NLR	nucleotide-binding leucine-rich repeat receptor/Nod-like receptor
NRC	NLR-required for cell death
NSm	nonstructural protein encoded by the medium RNA segment
NSs	nonstructural protein encoded by the small RNA segment
nt	nucleotide
OD600	optical density at 600 nm
P bodies	processing bodies
PAMP	pathogen-associated molecular pattern
PCD	programmed cell death
PCR	polymerase chain reaction
PDS	phytoene desaturase
phasiRNA	phasi secondary small interfering RNA
PRR	pattern-recognition receptor
PTI	PAMP-triggered immunity
PVX	Potato virus X
R gene	resistance gene
RB	resistance breaking
RDR, RdRp	RNA-dependent RNA polymerase
RFP	red fluorescent protein
RI	resistance-inducing
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
RNL	RPW8-NLR
RNP	ribonucleo(capsid)protein
ROS	reactive oxygen species
RPW8	Resistance to Powdery Mildew 8-like domain
RSS	RNA silencing suppressor
RT-qPCR	quantitative reverse transcription PCR
Rv	reverse
RVFV	Rift Valley fever virus
S gene	susceptibility gene
SA	salicylic acid
SAS	SUMO attachment site
SD	Solanaceae domain
SIM	SUMO interaction motif
siRNA	small interfering RNAs
SMV	Soybean mosaic virus
SN	systemic necrosis

SNV	Sin nombre Hantavirus
STTM	short tandem target mimic
TCSV	Tomato chlorotic spot virus
TCV	Turnip crinkle virus
TF	transcription factor
TIR	Toll/interleukin-1
TLR	Toll-like receptor
TMD	transmembrane domain
TMV	Tobacco mosaic virus
TNL	TIR-NLR
TRV	Tobacco rattle virus
TSS	transcription start site
TSWV	Tomato spotted wilt virus
TuMV	Turnip mosaic virus
TYRV	Tomato yellow ring virus
UTR	untranslated region
VIGS	virus induced gene silencing
VRF	viral replication factory
vsiRNA	viral siRNA

A

Summary

Plants are a fundamental part of (daily) life. They give us oxygen, nutrition, vitamins, and are used as feed for our cattle. To feed the growing world population, crops are often grown in large monocultures. This makes them more vulnerable and sensitive to insect pests and plant pathogens (fungi, bacteria, viruses, etc.) of which some are spread by (pest) insects; in principle, all plants are appetizing to the pest, and all are equally sensitive to the pathogen they might bring along. If left unchecked, pests and pathogens can wipe out large parts of a cultivated crop, resulting in harvest losses, reduced income for the farmer, and less food on the global market. A way to block the spread of disease is by using chemicals that kill pest and disease spreading insects. However, pesticides also exterminate insects that are beneficial. The overuse of such pesticides has also resulted in increased occurrence of pests that are insensitive to these chemicals. Furthermore, several types of pesticides are no longer allowed due to health and safety issues.

Figure 1 – Pepper running away from virus particles.

A different way to block the spread of pests and pathogens, in addition to for instance releasing the pest's natural enemy in the field, is the use of plants that are more resistant to pathogens and pests. As plants are not able to run away from so-called biological stressors (Figure 1), they need a good immune system to protect themselves. Some plants literally arm themselves, in the form of the bark of a tree that acts as a shield, or the spikey thorns on a rose making it difficult to get close. Not all plants have such armor, but they do all have ways to recognize invaders once their first defenses have been breached. In some cases, plants recognize a pathogen protein that is also present and conserved in other, similar, pathogens. In others, it recognizes the presence of DNA or RNA molecules (or unwanted derivatives) at locations in the cell where they normally should not occur. Their presence is one indication that intruders are present and multiplying. Targeting shared molecules and proteins is handy, as you can recognize different pathogens with one system. In other cases, the plant recognizes a very specific part that is only found in one or a few pathogens. These proteins are called resistance proteins or R proteins. In cases where these R proteins do not directly recognize the pathogen by one of its proteins, a so-called guard or a decoy protein of the plant, being targeted by the pathogen, is recognized by the R protein. The interaction between the guard/decoy and the pathogen leads to structural changes of the guardee/decoy, which are sensed by these R proteins. Regardless of the way in which the pathogen is recognized, directly or indirectly, the identification leads to a defense response to eradicate the pathogen. Considering that many of these resistance genes inherit in a single dominant manner, it is easy to see why R proteins are interesting traits for breeders to use and improve their crops with increased resilience towards insect pests and pathogens.

In this thesis, I looked at the interaction between a specific R protein and a particular pathogen; the R protein Tsw, in a pepper plant, that exclusively senses the presence of Tomato spotted wilt virus, or TSWV, in the cell. This virus can infect over 1000 different types of plants, many of which are important and high value crops such as tomato and bell pepper (Figure 2), as well as ornamental plants such as Chrysanthemum and Iris, important products from the Dutch

Figure 2 – A bell pepper infected with tomato spotted wilt virus (TSWV) and an uninfected bell pepper.

flower industries. TSWV is the representative of the group of tospoviruses, that share similar genetical, morphological and biological properties, but differ in their host range and thrips, the insect vectors, that spread these viruses. TSWV is one of the most important and devastating viruses from this group. The tospovirus RNA genome only has five genes that are needed for multiplication and the formation of new progeny virus particles. Aside from these five genes, the virus needs and highjacks many plant host proteins to fulfill its entire life cycle. One of its own genes encodes the protein NSs. This protein blocks the pathway in the plant defense that recognizes and removes and/or inhibits replicating viral DNA and RNA elements. By blocking this so-called antiviral defense response, TSWV viral genome copies are not destroyed, which allows the virus to establish an infection and cause disease in the plant. In turn, the pepper plant has developed additional methods to sense the presence of specifically TSWV, which is where the R protein Tsw comes into play. Tsw is able to recognize the NSs protein of TSWV, and upon detection start a rapid defense response that ends in the death of the cell within a few days. As a result, the virus remains contained within these dying cells and to escape to other tissues is thereby prevented. This blocks the virus infection, as the virus depends on living plants cells to survive. Chapter 1 and 2 together provide an in-depth overview of plant immune systems, the TSWV/tospovirus replication cycle, and the interplay between the host and the virus.

To investigate the interaction between Tsw and TSWV, the sequence of Tsw was needed. The successful search for the genomic nucleotide sequence of the plant gene that codes for Tsw is described in Chapter 3. At first glance, the gene seems to produce a normal R protein. However, one of its segments, or domains, called LRR, appeared unusually long. Generally, protein domains don't go over a certain size, as they are more prone to break apart the larger they get. The LRR domain in Tsw is almost three times the normal size of a protein domain, and is also much longer than the LRR domain of other R proteins (Figure 3). To take a visual look at how that might work, a 3D model of Tsw was made (Chapter 4), as well as of a non-functional variant tsw* that is found in non-resistant, or susceptible, peppers. This 3D model confirmed that Tsw is quite unconventional in its shape, while

A

| CC | Linker | NBS | LRR |

Figure 3 – A 3D model of A) an average sizes R protein ZAR1, *Arabidopsis thaliana*) and B) the exceptionally long R protein Tsw (*Capsiscum chinense*). CC – coiled coil (red); NBS – nucleotide binding domain (green); LRR – Leucine rich repeat (blue).
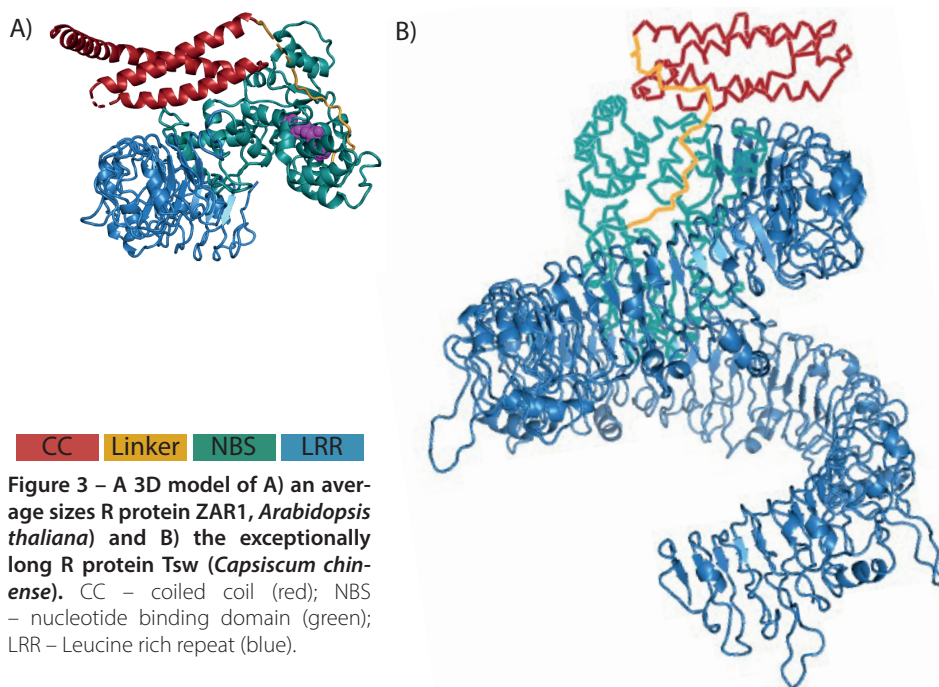
tsw looks like a normal R protein. This raises many questions. For instance, whether the long domain is needed for the recognition of TSWV, or whether Tsw even works in the same way as other R proteins do. A method was devised to construct the entire gene, which additionally allows for domains swaps between Tsw and tsw, which will help in studying the role of the long domain in the R protein function, and the purposes of the other domains (Chapter 5).

Computer predictions were used to search for targets within the gene encoding for Tsw that might indicate a way in which the production of Tsw is regulated (Chapter 5). Several potential targets were found that may influence Tsw production. To further analyze these, the amount of micro RNAs (miRNAs), master regulators of gene expression, that bind to these targets were changed using special techniques in pepper plants carrying Tsw or the non-active tsw. Plants with Tsw in which nothing was changed were resistant to the virus. However, Tsw plants with altered miRNA levels became infected with TSWV. Even more so than the susceptible tsw plants, with or without changed miRNA levels, suggesting that Tsw regulation appears being controlled by these specific miRNAs.

The quest for answers regarding the interaction between TSWV and Tsw was tackled from both sides. By using the NSs protein from TSWV as bait, I fished for plant proteins that interacted with NSs (Figure 4, Chapter 6) and found an interaction between NSs and Catalase. Catalase is a plant protein that converts hydrogen peroxide (H_2O_2) into water and oxygen. Harmful molecules like H_2O_2 are normally released during cellular stress and
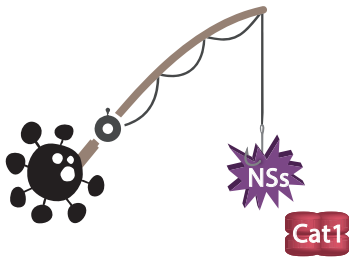
Figure 4 – By using the NSs protein of TSWV as bait, the plant protein Catalase was "caught" to interact with NSs.

defense responses caused by e.g. a virus infection. The interaction between NSs and Catalase lowers the levels of stress molecules, which seems to benefit the virus. Whether the interaction between NSs and Catalase is required for and/or part of the signaling cascade of the Tsw-mediated resistance response remains to be further investigated.

Several objectives set at the start of my research have been successfully addressed, but a lot of questions still remain or arise as a result of my thesis research. Continuing on these will help to create a better understanding of the Tsw-mediated defense response against the virus, and in the future may help towards developing new or alternative defense strategies in the plant to better withstand an infection with TSWV. By that time, this thesis will have been a small steppingstone to lower yield losses of pepper and slightly strengthen the food chain.

* If you are lost by now, don't worry:
  - TSWV   the virus
  - Tsw   the R protein that recognizes the virus
  - tsw   the variant of Tsw that does not work and therefore is unable to confer resistance

Samenvatting

Planten zijn een fundamenteel onderdeel van het (dagelijkse) leven. Ze zorgen voor zuurstof, voeding, vitamines en worden gebruikt als veevoer. Om de alsmaar groeiende wereldbevolking te voeden worden gewassen vaak verbouwd in grote monoculturen. Deze velden zijn hierdoor kwetsbaarder en gevoeliger voor insectenplagen en plantpathogenen (schimmels, bacteriën, virussen, etc.) waarvan sommige worden verspreid door (plaag)insecten; in principe zijn alle planten op zo'n veld appetijtelijk voor de plaag en zijn alle planten even gevoelig voor de ziekteverwekker die ze kunnen meebrengen. Als er niets

Figuur 1 – Paprika die weg rent van virus deeltjes.

aan wordt gedaan, kunnen plagen en ziekteverwekkers grote delen van een gewas vernietigen, wat resulteert in oogstverliezen, een lager inkomen voor de boer en minder voedsel op de wereldmarkt. Een manier om de verspreiding van ziekten te stoppen, is door chemicaliën te gebruiken die ziekteverspreidende insecten doden. Pesticiden verdelgen echter ook nuttige insecten. Het overmatig gebruik van dergelijke pesticiden heeft ook geleid tot een toename van plagen die ongevoelig zijn voor deze chemicaliën. Bovendien zijn verschillende soorten bestrijdingsmiddelen niet langer toegestaan vanwege gezondheids- en veiligheidsproblemen.

Een andere manier om de verspreiding van plagen en ziekteverwekkers tegen te gaan, naast bijvoorbeeld het uitzetten van de natuurlijke vijand van de plaag in het veld, is het gebruik van planten die beter bestand zijn tegen ziekteverwekkers en plagen. Omdat planten niet in staat zijn om weg te komen van zogenaamde biologische stressoren (Figuur 1), hebben ze een goed immuunsysteem nodig om zichzelf te beschermen. Sommige planten bewapenen zich letterlijk, in de vorm van de bast van een boom die als schild fungeert, of de stekelige doornen van een roos, die het moeilijk maken om dichtbij te komen. Niet alle planten hebben zo'n duidelijk pantser, maar alle planten hebben manieren om indringers te herkennen zodra hun eerste verdediging is doorbroken. In sommige gevallen herkennen planten een geconserveerd eiwit van een pathogeen; een soort eiwit dat ook aanwezig en geconserveerd is in andere, soortgelijke pathogenen. In andere gevallen herkent de plant de aanwezigheid van DNA- of RNA-moleculen (of ongewenste afgeleiden) op plaatsen in de cel waar ze normaliter niet voorkomen. De aanwezigheid van zulke moleculen is daarin een indicatie dat indringers aanwezig zijn en zich vermenigvuldigen. Voor het afweersysteem is het praktisch om moleculen en eiwitten te herkennen die karakteristiek zijn voor een bepaalde groep ziekteverwekkers, zodat een hele groep met een systeem kan worden aangepakt. In andere gevallen herkent de plant een heel specifiek onderdeel dat maar in één of enkele ziekteverwekkers voorkomt. Deze eiwitten worden resistentie-eiwitten of R-eiwitten genoemd. In gevallen waarin deze R-eiwitten de ziekteverwekker indirect herkennen, letten ze op een

zogenaamd bewakings- of een lokeiwit van de plant dat een interactie kan aangaan met de ziekteverwekker De interactie tussen de bewaker/lokaas en de ziekteverwekker leidt tot veranderingen in de structuur van de bewaker/lokaas, en het is deze verandering die door het R-eiwit wordt geregistreerd. Ongeacht de manier waarop de ziekteverwekker direct of indirect wordt herkend, leidt de identificatie tot een afweerreactie om de ziekteverwekker uit te roeien. Omdat veel van deze resistentiegenen op een enkele dominante manier overerven zijn veredelaars heel geïnteresseerd om R-eiwitten in hun gewassen te gebruiken en zo de weerstand van planten tegen insectenplagen en ziekteverwekkers te kunnen verhogen.

In dit proefschrift heb ik gekeken naar de interactie tussen een specifiek R-eiwit en een bepaalde ziekteverwekker; het R-eiwit Tsw, dat uit een paprikaplant komt, en de enige ziekteverwekker waarvan dit R-eiwit de aanwezigheid kan registreren, namelijk Tomato spotted wilt virus, of TSWV. Dit virus is in staat om meer dan 1000 verschillende soorten planten te infecteren, waarvan vele economisch belangrijke gewassen zoals tomaat en paprika (Figuur 2), maar

Figuur 2 – Een paprika geïnfecteerd met TSWV en een ongeïnfecteerde paprika.

ook sierplanten zoals chrysant en Iris, beiden belangrijke producten uit de Nederlandse bloemenindustrie. TSWV is de vertegenwoordiger van de groep tospovirussen, die vergelijkbare genetische, morfologische en biologische eigenschappen delen, maar verschillen in hun gastheerbereik en tripssoort, de insectenvectoren, die deze virussen verspreiden. TSWV is een van de belangrijkste en meest verwoestende virussen uit deze groep. Het tospovirus RNA-genoom heeft maar vijf genen, die nodig zijn voor vermenigvuldiging en de vorming van nieuwe virusdeeltjes. Naast van deze vijf genen heeft het virus hiervoor ook planteneiwitten nodig van zijn gastheer en die plundert het om zijn levenscyclus te vervullen. Een van de virusgenen codeert voor het eiwit NSs. Dit eiwit blokkeert de route in de plantenafweer die replicerende virale DNA- en RNA-elementen herkent en verwijdert en/of remt. Door deze antivirale afweerreactie te blokkeren worden genoomkopietjes van TSWV niet meer vernietigd, waardoor het virus een cel kan infecteren en zich kan gaan verspreiden door de plant. De paprikaplant heeft op diens beurt aanvullende methoden ontwikkeld om de aanwezigheid van specifiek TSWV te detecteren, en dat is waar het R-eiwit Tsw een rol speelt. Tsw is in staat om het NSs-eiwit van TSWV te herkennen en bij detectie van NSs start Tsw een snelle verdedigingsreactie die binnen enkele dagen eindigt in de dood van de cel. Als gevolg hiervan blijft het virus in deze stervende cellen en wordt voorkomen dat het naar andere weefsels ontsnapt. Dit blokkeert de virusinfectie, omdat het virus afhankelijk is van levende plantencellen om te overleven. Hoofdstuk 1 en 2 geven een goed overzicht van

A

het immuunsysteem van planten, de TSWV/tospovirus-replicatiecyclus en het samenspel tussen de gastheer en het virus.

Om de interactie tussen Tsw en TSWV te onderzoeken, was het nodig om de sequentie van Tsw te hebben. De succesvolle zoektocht naar de genomische nucleotidesequentie van het plantengen dat codeert voor Tsw wordt beschreven in Hoofdstuk 3. Op het eerste gezicht lijkt het gen een normaal R-eiwit te produceren. Een van zijn segmenten, of domeinen, genaamd LRR, blijkt echter ongewoon lang. Over het algemeen geldt hoe groter het eiwitdomein, hoe groter de kans is dat ze uit elkaar vallen. Daarom zijn eiwitdomeinen vaak niet groter dan een bepaalde maat, maar het LRR-domein in Tsw is bijna drie keer de normale grootte van een eiwitdomein. Daarnaast is het ook veel langer dan LRR-domeinen in andere R-eiwitten (Figuur 3). Om een beeld te krijgen van hoe dat zou kunnen werken is een 3D-model gemaakt van Tsw (Hoofdstuk 4), evenals van een niet-functionele variant tsw* die aanwezig is in niet-resistente, oftewel gevoelige, paprika's. Dit 3D-model bevestigde dat Tsw vrij onconventioneel van vorm is, terwijl tsw eruitziet als een normaal R-eiwit. Dit roept veel vragen op. Bijvoorbeeld of het lange domein nodig is voor de herkenning van TSWV NSs, en of dat Tsw op dezelfde manier of juist een andere manier werkt dan andere R-eiwitten. Er is een methode ontworpen om het hele gen te construeren (Hoofdstuk 5), wat bovendien de mogelijkheid bood om domeinen te wisselen tussen Tsw en tsw. Dit zal helpen bij het bestuderen van de rol van het lange domein in de R-eiwitfunctie en de doeleinden van de andere domeinen.

Met behulp van computervoorspellingen zijn mogelijke regulatiedoelwitten gevonden in het gen dat codeert voor Tsw (Hoofdstuk 5). Dit soort potentiële doelwitten zouden een aanwijzing kunnen zijn voor manieren waarop de productie van Tsw wordt gereguleerd. Om deze verder te analyseren is de hoeveelheid micro-RNA's (miRNA's), belangrijke
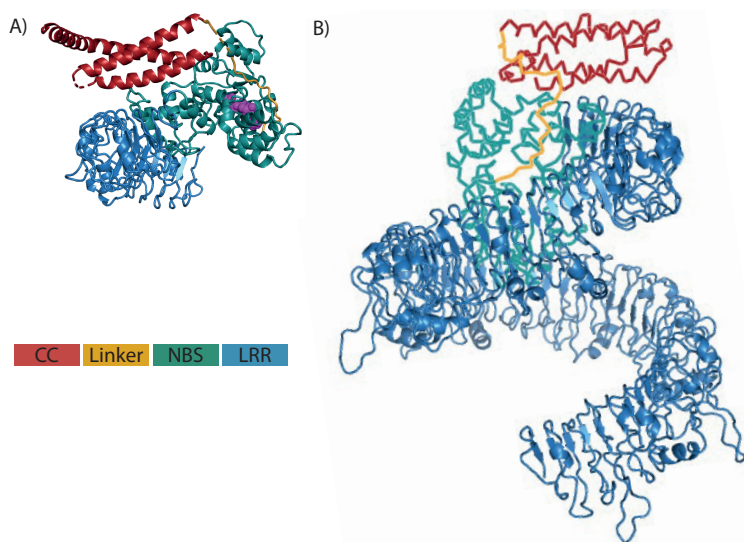
Figuur 3 – Een 3D model van A) an gemiddeld R eiwit (ZAR1, *Arabidopsis thaliana*) en B) het zeer grote R eiwit Tsw (*Capsiscum chinense*). CC – coiled coil (rood); NBS – nucleotide binding domain (groen); LRR – Leucine rich repeat (blauw).

regulatoren van genexpressie, die aan deze doelen binden, aangepast met behulp van speciale technieken. Dit is gedaan in peperplanten die Tsw of de niet-actieve tsw dragen. Planten met Tsw waarin niets veranderd was, waren resistent tegen het virus. Tsw-planten met aangepaste miRNA-niveaus raakten echter wel geïnfecteerd met TSWV. Meer zelfs dan de vatbare tsw-planten, met of zonder gewijzigde miRNA-niveaus, wat suggereert dat Tsw-regulatie lijkt te worden gecontroleerd door deze specifieke miRNA's.

Figuur 4 – Catalase 1 is gevangen als interacterend eiwit, door met het NSs eiwit van TSWV te vissen.

De zoektocht naar antwoorden over de interactie tussen TSWV en Tsw werd van beide kanten aangepakt. Door het NSs-eiwit van TSWV als aas te gebruiken, hebben we naar de planteneiwitten gevist die een interactie aangaan met NSs (Figuur 4, Hoofdstuk 6) en vonden we een interactie tussen NSs en Catalase. Catalase is een plantaardig eiwit dat waterstofperoxide (H_2O_2) omzet in water en zuurstof. Schadelijke moleculen, zoals H_2O_2, komen normaal gesproken vrij tijdens cellulaire stress- en afweerreacties veroorzaakt door b.v. een virusinfectie. De interactie tussen NSs en Catalase verlaagt de niveaus van stressmoleculen, wat het virus ten goede lijkt te komen. Of de interactie tussen NSs en Catalase vereist is voor en/of een deel van de signaalcascade van de Tsw-gemedieerde resistentierespons moet nog verder worden onderzocht.

Verschillende doelen die aan het begin van mijn onderzoek zijn gesteld zijn behaald, maar er blijven nog altijd veel vragen over en nieuwe vragen komen naar boven naar aanleiding van dit onderzoek. Antwoord op deze vragen zullen uiteindelijk voor zorgen dat we beter begrijpen hoe de afweerreactie van Tsw tegen het virus werkt, en dit kan in de toekomst helpen bij het ontwikkelen van nieuwe of alternatieve verdedigingsstrategieën in de plant om een infectie met TSWV beter te weerstaan. Tegen die tijd zal dit proefschrift een kleine opstap zijn geweest om de opbrengstverliezen van paprika te verlagen en de voedselketen enigszins te versterken.

A

* Als u nu helemaal de weg kwijt bent, geen zorgen:
  - TSWV   het virus
  - Tsw   het R-eiwit dat een onderdeel van het virus herkent
  - tsw   de variant van Tsw die niet werkt, en dus niet voor resistentie kan zorgen

Words of thanks

I have a strong suspicion that the only thing everyone, whether colleague or family, always reads in a dissertation are the acknowledgements. For me, it has nothing to do with my interest in the research, and everything to do with the celebration of other people that happens in this section. While doing a PhD can at times feel lonely, it is impossible to do it by yourself. As a PhD candidate are supported, critiqued, celebrated by numerous people. These interactions shape you and give direction to your research, but all those lovely people doing the shaping are usually on the sidelines. The acknowledgment section is the place to put these people in the spotlight they deserve (and the place to snoop around to find your own name), so here we go!

The first people to thank are Richard and Monique, my PhD team. **Richard**, as my daily supervisor we have had countless of discussions ranging from constructive talks about (my) work, to debates over coffee about, just as a purely hypothetical example, the proper etiquette of disagreeing with someone's choice for a hedge. I consider myself very lucky with a supervisor who will always make time for my pressing questions and discussions about work. Thank you very much for all the input, critical remarks, and ideas. And of course, the nice meals during the plant group dinners at your place! **Monique**, as a professor you are usually more in the background until the last part of the PhD. However, in my case you came in a bit earlier, and as such have you to thank for the synthetic approach towards getting Tsw. Thank you for creating the opportunity in which I could take this route, and for the fast response whenever I sent something for feedback. I also want to thank you for helping me get back on my feet after an insignificant cold wiped me and my vestibular system out completely. Our talks went from research-related to person-related, and I appreciated you sharing this side of you with me.

From promoters we move on to the rest of Virology, starting off with the plant people. When I started, Magda and André had just joined Corien in the lab. I don't think it was due to our (read: their) loudness in and outside of the lab, but within half a year we moved to our own little plant lab. I remember a lot of fun times and games, good discussions during coffee, plenty of good and bad music coming out of the speakers in the lab, and *so* many students roaming around. **Magda**, you were the post-doc on our shared project, although we didn't work together much on the lab. Thank you for your sharp eye and input, and especially for giving me a piece of advice I still thank you for, namely, to immediately learn how to use Illustrator. **André**, thank you for the good music in the lab, and the even better talks. You are the kind of scientist I aspire to be; someone with knowledge and insight, one who is willing to share this knowledge, and has the smarts to pretend you have the knowledge when necessary. And, who has a kind heart to boot. **Corien**, while it won't come as a surprise that I won't thank you for your taste in music, I am thankful for many other things you bring to the table. You seem to know someone everywhere you go, and I thank you for letting me tag along at conferences and meetings. As an introvert,

having such a sociable colleague and friend like you is almost a prerequisite to get some networking done. Thanks for your laughter and the lightness you bring besides that smart brain of yours, and for helping out a friend in need, several times.

Other people joined the plant lab. Like **Min**, my next-door-neighbor, who brought cute gifts from China, and is always interested in to hear what you have been up! Or like **Sharella**, who started out as a thesis student with André, and now is a PhD candidate. Thanks to the both of you for your kindness and the nice chats during coffee breaks. Several technicians also came, and went, in the plant lab. When I started **Dick** was the plant technician, who went with well-deserved retirement, and was followed by the lovely Maria, after which **Cristina** came. **Dennis** came along as Emilyn's new technician. And now, just before I finished, **Mark** started. All plant technicians have added their own flair to the job, helped keep the lab organized and enabled us to do our research. PhD life is so much better with you in it! Thanks! During my time at Virology I also supervised nine thesis students, who each taught me something. It was great helping you on your way to become independent scientists, thank you, **Francesco**, **Martijn**, **Yoga**, **Rik**, **Yixuan**, **Lonne**, **Niek**, **Amanda**, and **Iris**.

Two plant people who deserve the spotlight even more are my two paranymphs. Having you two by my side at my defense feels like the proper finish to the last years at Virology. **Cristina**, first, thank you for your help with experiments and thank your luck for getting our "golden fish" (the BAC clone with *Tsw*) with just one last colony PCR. You are kind, sweet, and I think it is great that you grew into the teacher you never knew you wanted to be! And then we have **Mandy**, my desk buddy, coffee companion, and fellow feminist. I won't thank you for your taste in music either, but that dislike goes both ways :D. I will however thank you profusely for being a such a good friend and colleague. It feels like we held the plant lab afloat at times, and each other as well. You looked out for me, listened to my stories, and sent me home whenever my brain didn't register it was shutting down. I got to return the favor, and am just very glad we both are on the up and up again! Good luck with the last stretch of your PhD and if you need an ear, I am here for you.

Apart from the lovely plant people I have mentioned, there are many just as lovely non-plant people who help make Virology a good place to be. **Jan**, **Vera**, **Gorben**, **Jelke**, **Just**, **Jeroen**, **Marcel**, **Jelmer**, **René**, thank you for the nice (non) scientific talks during coffee breaks. The same goes for **Emilyn**, a very welcome addition to the plant virus family. Thank you for your interest in me and my project, and I am looking forward seeing where your research takes you. **Karen**, I always appreciated it when you joined our plant meetings. Thank you for your insight and interest.

I owe plenty of thanks to the other technicians on the lab. Of course, for the good organization of the labs, but even more so for all the other stuff. These are just some small things that come to mind, but there is plenty more. **Hanke**, seeing your name on things

A

in the plant lab still makes me smile. **Marleen**, thanks for taking me ice skating early in the morning at the semi-secret spot in the Uiterwaarden. **Els** and **Dorothy**, thank you for allowing me to help you in the insect rearing as occupational therapy. Els, best of luck and take care! **Gwen**, I appreciated our talks and your interest in plants. **Corinne**, to me you are instrumental in making Virology feel like more than just a department. And, thanks for taking me to Sprookjeswonderland. If Pepijn or Rutger ever want to go again, let me know! Many thanks as well to all the previous and current PhDers and post-docs at Virology: **Ahmed**, **Annemaria**, **Astrid**, **Aydin**, **Bob**, **Caroline**, **Fengqiao**, **Gabriela**, **Giel**, **Han**, **Hannah**, **Haidong**, **Irene Meki**, **Irene Ontiveros Espinel**, **Janna**, **Jerôme**, **Jirka**, **Jitte**, **John**, **Joyce**, **Linda van Oosten**, **Linda de Jong**, **Lisa**, **Luzhao**, **Mark**, **Melissa**, **Miao**, **Sara**, **Sandra**, **Simone**, and **Tessy**.

I also have plenty of non-Virology people to thank. For instance, the people from PRI, especially **Jan**, **Annette**, and **Petra**, thank you helping me out at times with your (plant virus) expertise and equipment. Or those researchers from Phytopathology and Nematology also working on *R* genes, with whom we had semi-regular meetings to talk about our research and problems. Thank you **Aska**, **Matthieu**, **Aranka**, **Octavina**, **Erik**, **Yinbin**, **Qi**, **Sergio**, and **Wen** for these interesting talks. I also want to thank the EPS PhD council members. It was great to organize activities for PhD students, and to feel like together we could make life for PhDers just slightly better. Thanks **Susan**, **Ingrid**, **Daan**, **Daniel**, **Davy**, **Hao**, **Ivo**, **Jasper**, **Lena**, **Michelle**, **Mandy**, **Octavina**, **Tieme**, **Ties**, **Valérie**, **Vera**, **Ivo**, **Sietske**, and **Stuart**.

I have used, and thus killed, plenty of plants for this research and owe a huge thanks to the people from Unifarm, **Bert**, **Bertus**, and most of all **Henk**, for taking such good care of my plants. Henk thank you so much for accommodating my last-minute request of more plants, and for the nice chats in the greenhouse.

While I was a competitive race rower during my PhD (You may ask, 'Why, isn't a PhD already enough hard work?'. Yes. But it was terrible fun. Terrible. And fun.), I started as a regular rower at Argo, and it is with that group of lovely people that I had most fun with. From rowing on the Rijn, the canals of Venice, or by the Eifel tower, to sharing grown-up stories involving actual grown-up jobs, kids (or dogs), and emigrations. **Carina**, **Esther**, **Merel**, **Rosanne**, and **Rosita**, a.k.a. gRoeifactor, thanks!

Ava, **Koen**, and **Vivian**, thank you for the nice walks and equally good talks, whether through the forests in Lunteren, traipsing through a city in search of a botanical garden, or just through the mud in Wageningen.

Besides my very nice collogues and friends, I have a great family that I can fall back on, and without whom I would be nowhere near where I am today. **Maarten**, **Annemarie**, and **Atje** (sorry Alexander, this is now your official name), thank you for your interest in me

Education Statement of the Graduate School
Experimental Plant Sciences

Issued to:	Irene Louise van Grinsven
Date:	10 October 2022
Group:	Laboratory of Virology
University:	Wageningen University and Reseach
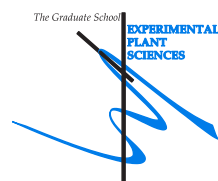
1) Start-Up Phase	**date**
▸ First presentation of your project	
Tsw-mediated resistance and the innate immunity modulator NSs from TWSV	16 Jan 2017
▸ Writing or rewriting a project proposal	
▸ MSc courses	
Subtotal Start-Up Phase	*1.5*

2) Scientific Exposure	**date**
▸ EPS PhD student days	
Get2Gether for EPS PhD students, Soest, NL	9-10 Feb 2017
Get2Gether for EPS PhD students, Soest, NL	11-12 Feb 2019
Get2Gether for EPS PhD students, Soest, NL	10-11 Feb 2020
▸ EPS theme symposia	
EPS Theme 2 Symposium & Willie Commelin Scholten Day 'Interactions between plants and biotic agents', Wageningen, NL	23 Jan 2017
EPS Theme 2 Symposium & Willie Commelin Scholten Day 'Interactions between plants and biotic agents', Amsterdam, NL	24 Jan 2018
EPS Theme 2 Symposium & Willie Commelin Scholten Day 'Interactions between plants and biotic agents', Wageningen, NL	1 Feb 2019
▸ Lunteren Days and other national platforms	
Dutch Annual Virology Symposium (DAVS), Amsterdam, NL	3 Mar 2017
Dutch Annual Virology Symposium (DAVS), Amsterdam, NL	9 Mar 2018
Dutch Annual Virology Symposium (DAVS), Amsterdam, NL	8 Mar 2019
Dutch Annual Virology Symposium (DAVS), online	5 Mar 2021
Annual meeting Experimental Plant Sciences, Lunteren, NL	10-11 Apr 2017
Annual meeting Experimental Plant Sciences, Lunteren, NL	9-10 Apr 2018
Annual meeting Experimental Plant Sciences, Lunteren, NL	8-9 Apr 2019
Annual meeting Experimental Plant Sciences, online	12-13 Apr 2021
▸ Seminars (series), workshops and symposia	
Seminar: Michael Diamond, 'New Insights into Zika Virus Pathogenesis'	17 Nov 2016
Seminar: Mart Krupovic, 'Natural history of viral capsids'	23 Feb 2017
Seminar: Martin Cann, 'The immune receptor Rx1 remodels chromatin and chromatin interactors in immunity'	11 Jul 2017
Seminar: Richard Lenski, 'Dynamics of adaptation and genome evolution in a long-term experiment'	31 Aug 2017
Seminar: Andrea Gröne, 'Why bother? Disease is normal in wildlife'	28 Sep 2017
Seminar: Peter Simmonds, 'RNA virus genomes: more than just a coding sequence?'	20 Nov 2017
Seminar: Anne-Nathalie Volkoff, 'Endogenous viruses used by parasitic wasps to deliver virulence molecules to their hosts'	14 Dec 2017
Seminar: Sophien Kamoun, 'Beyond the single gene: NLR network mediates immunity to diverse plant pathogens'	6 Mar 2018
Seminar: Isreal Pagán, 'Extreme tactics of plant virus transmission: from seed dispersion to parasitic castration'	5 Apr 2018
Seminar: Esther Schnettler, 'The antiviral RNAi response in arthropods'	8 Oct 2018
Seminar: Adly Abd-Alla, 'Virus problems in insect mass rearing and a new protocol for colony management'	8 Oct 2018
Seminar: Olga Haenen, 'Novel proteins: insects and fish, healthy sustainable and safe'	9 Oct 2018
Seminar: Jørgen Eilenberg, 'The framework for a coordinated European effort to manage insect diseases in production system'	9 Oct 2018

Seminar: Michael Strand, 'Evolution, Function and Endogenization: Lessons learned from virus associations with parasitic insects'	8 Nov 2018
Seminar: Jijie Chai, 'Structure, mechanism and biochemical insight of plant NLR protein'	5 Jun 2019
Seminar: Arvind Varsani, 'Notes on viral discovery, evolution and more: From the tropics to the frozen continent'	25 Jun 2020
Seminar: Stefan Giesen, 'Challenges as a PhD and postdoc in Science and tips to overcome those'	20 Jan 2021
Seminar: Ab Osterhaus, 'Newly emerging viruses, from the animal world'	26 Apr 2021
Seminar: Björn Krenz, 'Virus-host plant interaction studies: from stress granules to nuclear architecture manipulation'	6 Oct 2021
Symposium: 'WURomics: Technology-Driven Innovation for Plant Breeding", Wageningen, NL	15 Dec 2016
Symposium: '60 years Virology: Ceremony and lectures on 60 years of research in the Virology department', Wageningen, NL	26 Oct 2017
▸ **Seminar plus**	
Master Class with Michael Diamond	18 Nov 2016
Discussion with Martin Cann	11 Jul 2017
Master Class with Micheal Strand	9 Nov 2018
▸ **International symposia and congresses**	
Sixth Joint Meeting of the DPG working Group "Virus Diseases of Plants" and the "Nederlandse Kring voor Plantenvirologie", Bonn, DE	27-28 Mar 2017
European Plant Science Retreat, Utrecht, NL	3-6 Jul 2018
XVIII International Congress on Molecular Plant Microbe Interactions Congress, Glasgow, GB	14-18 Jul 2019
▸ **Presentations**	
Poster: Unravelling Tsw-mediated resistance and the interplay with NSs of Tomato Spotted Wilt Virus at EPS course Transcription Factors and Transcriptional Regulation.	12 Dec 2016
Presentation: Tsw-mediated resistance and the innate immunity modulator NSs from TSWV (First consortium presentation)	21 Dec 2016
Poster: Cell death triggering and effector recognition by dominant resistance genes *Sw5* and *Tsw* at Sixth Joint Meeting DPG Working Group.	27 Mar 2017
Presentation: Tsw-mediated resistance and the RNA silencing suppressor NSs from TSWV (First *R* gene meeting)	16 Jun 2017
Presentation: Unravelling Tsw-mediated resistance and the interplay with the innate immunity modulator NSs from TSWV (Pitch at workshop Michael Diamond)	18 Nov 2016
Presentation: Tsw-mediated resistance and the RNA silencing suppressor NSs from TSWV (Pitch at workshop Michael Strand)	9 Nov 2018
Presentation: Introduction into Tomato Spotted Wilt Virus, the RNA silencing suppressor NSs and resistance gene *Tsw* (Final *R* gene meeting)	25 Jun 2020
Presentation: Tomato Spotted Wilt Virus, the RNA silencing suppressor NSs and resistance gene *Tsw* (Final consortium meeting)	19 Oct 2020
▸ **3rd year interview**	
▸ **Excursions**	
EPS Company Visit to TomatoWorld	14 Oct 2016
EPS Company Visit to Koppert Biological Systems	27 Oct 2018
EPS Online Company visit to RijkZwaan	16 Jun 2021
Subtotal Scientific Exposure	*20.7*

3) In-Depth Studies	**date**
▸ **Advanced scientific courses & workshops**	
EPS course 'Transcription Factors and Transcriptional Regulation', Wageningen, NL	12-14 Dec 2016
EMBO workshop 'Pathogen immunity and signalling', Oxford, GB	1-5 Apr 2019
PE&RC & WIMEK course 'Introduction to R and Rscript', online	12 Jan - 2 Feb 2021
▸ **Journal club**	
▸ **Individual research training**	
Confocal training, Norbert de Ruijter, Wageningen, NL	9 Oct 2017
Subtotal In-Depth Studies	*3.1*

A

Appendices

4) Personal Development	date
▸ General skill training courses	
WGS PhD Competence Assessment, Wageningen, NL	8 Nov 2016
EPS Introduction Course, Wageningen, NL	16 Feb 2017
WGS course 'Supervising BSc and MSc Thesis Students', Wageningen, NL	23-24 Oct 2017
WGS course 'Scientific Writing', Wageningen, NL	13 Nov 2017 - 23 Jan 2018
WGS workshop 'How to fail', online	17 May 2021
WGS workshop 'Managing your supervisor in an online world', online	19 May 2021
WGS workshop 'Drawing essentials for impactful communication, peace of mind and lots of fun', online	25 May 2021
EPS workshop 'Becoming a mindful scientist', online	10 Sept 2021
▸ Organisation of meetings, PhD courses or outreach activities	
Get2Gether EPS PhD Student Days 2019	11-12 Feb 2019
Get2Gether EPS PhD Student Days 2020	10-11 Feb 2020
Member of advisory committee for the appointment of two new chairs for the chairgroup of Plant Development Biology	2019
Participation in peerreview meeting of Biotic Interactions and Plant Health unit	23 Nov 2021
Membership of EPS PhD Council	
Member of EPS PhD council	2018-2021
Subtotal Personal Development	*7.8*

5) Teaching & Supervision Duties	date
▸ Courses	
PHP-30806 - Molecular Aspects of Bio-Interactions	2016, '17, '19, '21
CBI-30806 - Immunotechnology	2017-2019
VIR-30806 - Fundamental and Applied Virology	2017
ENT-30306 - Ecological Aspects of Bio-Interactions	2017
VIR-30306 - Molecular Virology	2018
PHP-21303 - Fundamentals of Plant Pathology and Entomology	2020
▸ Supervision of BSc/MSc students	
New insight for Tsw gene: Physical Mapping of the Resistance Gene Tsw and Optimization of *A. tumefaciens* transient assay in Capsicum spp., Francesco Desiderio	2017
Identification of interaction partners of the TSWV NSs protein active in RNAi and *Tsw* gene resistance, Martijn Vogelaar	2018
Elucidation of the interaction between NSs and catalse, I Gede Yoga Anantawijaya	2018
Elucidation of the Tomato spotted wilt virus NSs gene, unraveling its interaction with catalase, Rik Muilwijk	2019
Screening the Capsicum chinense BAC library for *Tsw*-related genes, Yixuan Yang	2019
Investigation of the potential interactions of TSWV NSs with the proteins HPH2B, eIF4A-14, and PABP, Lonne Siegers	2019
Mapping the binding domain of TSWV NSs for Catalase 1 using co-immunoprecipitation, Niek de Jonge	2019
In kaart brengen van het resistentie-gen Tsw in *Capsicum chinense* en de detectie van *Tsw*-F in aanverwante plantensoorten, Amanda Franken	2020
Antiviral RNAi-mediated regulation of the resistance gene *Tsw* in pepper, Iris Goet	2020
Subtotal Teaching & Supervision Duties	*6.0*

TOTAL NUMBER OF CREDITS*   **39.1**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits. *One credit represents a normative study load of 28 study hours.

About the author

Irene Louise van Grinsven was born on the 4th of December 1990, in Hoorn. At age 7, she lived in Santa Cruz (USA) for a year, were she obtained her American accent while attending 1st grade. She completed her primary and secondary school in Enkhuizen, the Netherlands. After graduating in 2009, she worked and travelled for a year in France, the USA, Mexico, and the Philippines. In 2010, she started with her study Molecular Life Sciences at Wageningen University. This combined all her favorite subjects: biology, physics, and chemistry. During her study she discovered that plants are as interesting as her dad had been telling her, and she followed a minor in Plant Biotechnology. For her MSc thesis, she worked with Dr. Erin Bakker and Dr. Aska Goverse at the Laboratory of Nematology on homologues of the resistance gene *Rx* in wild and cultivated tomato. As a final part of her study, she had the opportunity to work on novel applications of CRISPR Cas9 in tomato protoplasts at KeyGene, Wageningen. In the last year of her MSc, Irene was a student assistant in the course Immunotechnology, for which she developed course material. Her rowing ambitions also intensified during her MSc, and Irene became an active competitive race rower at W.S.R. Argo, with which she continued until 2018.

Since 2016, Irene has been a PhD candidate at the Laboratory of Virology. Under the supervision of Dr. Richard Kormelink, she worked on the interaction between the resistance gene *Tsw* and the plant virus Tomato Spotted Wilt Virus against which it confers resistance. The results of the conducted research are described in this thesis. During her time at Virology she won the prize for Best Dessert at Virology's International Dinner, was a member of the Experimental Plant Sciences PhD council for two years, supervised the thesis research of one HBO and eight MSc students, and was involved in teaching six different courses.

From July 2022 onwards, Irene works as a postdoc in the group of prof. Dr. Rob Schuurink at the Swammerdam Institute for Life Sciences in Amsterdam to research the effect of viruses and microsporidia on the effectorome of thrips and the plant proteins with which these thrips effectors interact.

A