UNRAVELLING THE MOLECULAR MECHANISMS OF PLANT-PATHOGEN INTERACTIONS IN GRAPEVINE AND POTATO THROUGH FUNCTIONAL ANALYSIS

Špela Tomaž

Doctoral Dissertation
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia

Supervisor: Dr. Anna Coll, National Institute of Biology, Ljubljana, Slovenia

Evaluation Board:
Prof. Dr. Jože Pungerčar, Chair, Jožef Stefan Institute, Ljubljana, Slovenia and Jožef Stefan International Postgraduate School, Ljubljana, Slovenia
Prof. Dr. Jose Manuel Franco-Zorilla, Member, Centro Nacional de Biotechnologia, Consejo Superior de Investigaciones Científicas, Madrid, Spain
Dr. Špela Baebler, Member, National Institute of Biology, Ljubljana, Slovenia

MEDNARODNA PODIPLOMSKA ŠOLA JOŽEFA STEFANA
JOŽEF STEFAN INTERNATIONAL POSTGRADUATE SCHOOL

Špela Tomaž

UNRAVELLING THE MOLECULAR MECHANISMS OF PLANT-PATHOGEN INTERACTIONS IN GRAPEVINE AND POTATO THROUGH FUNCTIONAL ANALYSIS

Doctoral Dissertation

RAZKRIVANJE MOLEKULARNIH MEHANIZMOV INTERAKCIJ MED RASTLINAMI IN PATOGENI PRI VINSKI TRTI IN KROMPIRJU S FUNKCIONALNO ANALIZO

Doktorska disertacija

Supervisor: Dr. Anna Coll

Ljubljana, Slovenia, October 2022

Acknowledgments

This dissertation is the result of an excellent collaboration with many colleagues from different institutions. It was financially supported by the Slovenian Research Agency through the research core funding program no. P4-0165, projects no. J7-7248 and J1-7151, and the young researchers' program no. 1000-22-0105 in accordance with the agreement on (co)financing of research activities in 2022. The work was mainly carried out at the National Institute of Biology, Department of Biotechnology and Systems Biology.

First, I would like to thank my supervisor, Anna Coll, for her excellent guidance, advice, and many fruitful discussions throughout my studies, as well as for taking the time to answer all my questions. I would also like to express my sincere gratitude to Kristina Gruden, who offered me the position in her group and contributed greatly to this work with her immense knowledge and wisdom. A big "thank you" also goes to Marina Dermastia and Maruša Pompe Novak for their expertise and collaboration on the grapevine studies. I would like to thank Tjaša Lukan for her help with protein localization and confocal microscopy. Additionally, I would like to thank Erica Prates from Oak Ridge National Laboratory, for providing her expertise on structural modeling and discussions related to protein structure, as well as Matej Butala and Gregor Bajc from Biotechnical Faculty, University of Ljubljana, for their help with the protein-DNA interaction studies. I would like to thank the members of the dissertation committee, Jože Pungerčar, Jose Manuel Franco-Zorilla, and Špela Baebler, for their thorough review of the thesis.

Thanks to all of the FITO members for providing a great working environment, cooking delicious Friday lunches, and the many moments we shared, whether during or outside of work. You have all contributed in some way to the completion of this work. In particular, I would like to thank Maja Križnik, Marko Petek, Karmen Pogačar, Katja Stare, and Tjaša Stare for their office and laboratory company. Finally, yet importantly, I would like to thank my family and friends for their encouragement and support.

Abstract

Plants are constantly exposed to various environmental stressors. Pathogen infections causing disease can affect economically important crops and impact global food production. To counter disease development and spread, it is imperative that we acquire a thorough knowledge of its pathogenesis, as well as of the mechanisms underlying plant immune response. Plant-pathogen interactions mostly play out at the molecular level, encompassing a range of molecular processes and interactions between biological molecules that are unique to each pathosystem. Exploring these complex mechanisms entails using various functional analysis approaches, from high-throughput sequencing analyses to *in vitro* protein interaction assays. In the scope of this thesis, we adapt and apply different methodologies to study the molecular processes underlying grapevine (*Vitis vinifera*) infection with phytoplasma and the mechanisms of potato (*Solanum tuberosum*) immune response against potato virus Y (PVY).

In studying the grapevine-phytoplasma pathosystem, we first optimized an RNA isolation protocol for the efficient isolation of total RNA, including small RNAs, from samples of '*Candidatus* Phytoplasma solani' phytoplasma-infected grapevines. Our optimized procedure enabled the isolation of pure RNA of high integrity and concentration, which we used for library preparation and high-throughput sequencing. We further used a protein pull-down assay coupled with mass spectrometry to screen for plant protein targets of a putative '*Candidatus* Phytoplasma solani' effector and subsequently confirmed the interactions between the effector and identified targets with *in planta* co-immunoprecipitation. The results of this study provide important functional information about the poorly known mechanisms of phytoplasma effector-mediated pathogenicity.

To improve our understanding of the potato immune response to PVY infection, we investigated the role of potato TGA transcription factors, regulatory proteins known for their importance in salicylic acid-mediated plant immunity in the model plant Arabidopsis (*Arabidopsis thaliana*). Our *in silico* analyses of potato TGA (StTGA) protein sequences revealed the presence of truncated proteins, which we named mini-TGAs. As mini-TGAs are not known in Arabidopsis, we performed different functional analyses to characterize the potato mini-TGA StTGA2.1. The results of this study revealed that StTGA2.1 compensates for salicylic acid deficiency in potato immune response to PVY and associates with other full-length StTGAs to modulate transcription. We showed that StTGA2.1 is involved in the regulation of class III peroxidase gene expression and used advanced computational modeling to gain insight into the structural aspects of mini-TGA function.

Altogether, the results presented in this thesis demonstrate the application of functional analysis methodologies in plants and provide new insights into the mechanisms of plant-pathogen interactions in two economically important crop species: grapevine and potato. The newly acquired knowledge improves our understanding of the molecular biology of plant diseases, which is critical for the development of effective crop protection strategies.

Povzetek

Rastline so nenehno izpostavljene različnim okoljskim stresorjem. Okužbe s povzročitelji bolezni lahko prizadenejo gospodarsko pomembne pridelke in vplivajo na svetovno proizvodnjo hrane. Za preprečevanje razvoja in širjenja bolezni moramo pridobiti podrobno znanje o njeni patogenezi ter o mehanizmih, na katerih temelji imunski odziv rastlin. Interakcije med rastlinami in patogeni se večinoma odvijajo na molekularnem nivoju in vključujejo vrsto molekularnih procesov ter interakcij med biološkimi molekulami, ki so edinstvene vsakemu patosistemu. Raziskovanje teh zapletenih mehanizmov vključuje uporabo različnih pristopov funkcionalne analize, od visoko zmogljivih analiz sekvenciranja do testov interakcije proteinov *in vitro*. V okviru te doktorske naloge smo prilagodili in uporabili različne metodologije za preučevanje molekularnih procesov okužbe vinske trte s fitoplazmo ter mehanizmov imunskega odziva krompirja na okužbo z virusom krompirja Y (PVY).

Pri preučevanju patosistema vinska trta-fitoplazme smo najprej optimizirali protokol izolacije RNA za učinkovito izolacijo skupne RNA, vključujoč male RNA, iz vzorcev vinske trte, okužene s fitoplazmo '*Candidatus* Phytoplasma solani'. Naš optimizirani postopek je omogočil izolacijo čiste RNA visoke integritete in koncentracije, ki smo jo uporabili za pripravo knjižnic in sekvenciranje z visoko zmogljivostjo. V nadaljevanju smo za iskanje rastlinskih proteinskih tarč domnevnega efektorja fitoplazme '*Candidatus* Phytoplasma solani' uporabili test »pull down« v kombinaciji z masno spektrometrijo in nato interakcije med efektorjem in identificiranimi tarčami potrdili s ko-imunoprecipitacijo *in planta*. Rezultati te študije podajajo pomembne funkcionalne informacije o slabo poznanih mehanizmih patogenosti fitoplazem, posredovanih z efektorji.

Da bi izboljšali naše razumevanje imunskega odziva krompirja na okužbo s PVY, smo raziskali vlogo krompirjevih transkripcijskih faktorjev TGA, regulatornih proteinov, ki so v modelni rastlini navadni repnjakovec (*Arabidopsis thaliana*) znani po svoji vlogi v rastlinski imunosti, posredovani s salicilno kislino. Naše analize proteinskih zaporedij krompirjevih TGA (StTGA) *in silico* so pokazale prisotnost skrajšanih proteinov, ki smo jih poimenovali mini-TGA. Ker proteini mini-TGA v navadnem repnjakovcu niso poznani, smo izvedli različne funkcionalne analize, da bi okarakterizirali krompirjev mini-TGA StTGA2.1. Rezultati te študije so pokazali, da lahko StTGA2.1 kompenzira pomanjkanje salicilne kisline pri imunskem odzivu krompirja na PVY in se povezuje z drugimi proteini StTGA, polne dolžine, za modulacijo transkripcije. Pokazali smo, da je StTGA2.1 vključen v uravnavanje izražanja genov peroksidaz razreda III, in uporabili napredno računalniško modeliranje za vpogled v strukturne vidike delovanja proteinov mini-TGA.

Rezultati, predstavljeni v tej disertaciji, prikazujejo uporabo metodologij funkcionalne analize v rastlinah in prinašajo nova spoznanja o mehanizmih interakcij med rastlinami in patogeni pri dveh gospodarsko pomembnih rastlinskih vrstah, vinski trti in krompirju. Novo pridobljeno znanje izboljšuje naše razumevanje molekularne biologije rastlinskih bolezni, kar je ključnega pomena za razvoj učinkovitih strategij za zaščito poljščin.

Contents

List of Figures

List of Tables

Abbreviations

3D	... three-dimensional
AMP	... antigenic membrane protein
bZIP	... basic region leucine zipper
CTAB	... cetyltrimethylammonium bromide
DNA	... deoxyribonucleic acid
DOG	... Delay of Germination
ETI	... effector-triggered immunity
FBA	... fructose-bisphosphate aldolase
GFP	... enhanced green fluorescent protein
GDR	... geranylgeranyl diphosphate reductase
H2B-YFP	... red Histone 2B-eYFP nuclear marker
mini-TGA	... a truncated TGA protein found in several plant species
mRNA	... messenger RNA
MS	... mass spectrometry
NLR	... nucleotide-binding leucine-rich repeat receptor
NPR	... Non-expressor of Pathogenesis Related
PGM	... phosphoglucomutase
PRR	... pattern recognition receptor
PRX	... class III peroxidases
PTI	... pattern-triggered immunity
PVY	... potato virus Y
RFP	... monomeric red fluorescent protein 1
RNA	... ribonucleic acid
rRNA	... ribosomal RNA
sRNA	... small RNA
SA	... salicylic acid
SAP	... secreted AY-WB protein
SDS-PAGE	... sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TGA	... TGACG-binding
WRKY	... WRKY transcription factors
YFP	... enhanced yellow fluorescent protein

Chapter 1

Introduction

1.1 Plant Diseases

Plant diseases impact the product quality and production yield of economically important plant species worldwide. About 20–30% of yield losses in the five major crops (i.e., wheat, rice, maize, potato, and soybean) are caused by various pathogens or pests (Oerke, 2006; Savary et al., 2019). The main causal agents of plant diseases include viruses, bacteria, oomycetes, fungi, nematodes, and parasitic plants (Strange & Scott, 2005). Disease outbreaks are usually difficult to predict or control, as different pathogens often cause similar symptoms and frequently overcome plant resistance due to continuous and swift evolution. The interactions between pathogens and plants are highly specific and can lead to different outcomes depending on the interaction compatibility, which can differ between pathogen strains or the host cultivars of the same species (Sharma et al., 2022; Strange & Scott, 2005). It is thus not feasible to generalize research in phytopathology; moreover, while we have learned a lot by studying the few model plant species, most importantly Arabidopsis (*Arabidopsis thaliana*), each pathosystem needs to be addressed individually, and our knowledge of crop diseases is still lacking.

In the following subsections, we briefly introduce two types of plant diseases, phytoplasma and viral infections, and describe one representative of each type in more detail, namely the '*Candidatus* Phytoplasma solani' ('*Ca.* P. solani'), a causative agent of the Grapevine Yellows disease in grapevine (*Vitis vinifera*), and the potato virus Y (PVY), considered the most dangerous virus infecting potato (*Solanum tuberosum*) (Kreuze et al., 2019).

1.1.1 Phytoplasma Infections

Phytoplasmas are small bacterial organisms that belong to the class of *Mollicutes* (Bertaccini & Lee, 2018). Taxonomically, they are distributed into subgroups based on their 16S ribosomal DNA and were assigned to the '*Candidatus*' genus (IRPCM, 2004). Unlike most bacteria, they lack a cell wall and are thus obligate parasites. They reside in the phloem tissue and are mainly transmitted by insects feeding on the phloem sap, including leafhoppers, plant hoppers, and psyllids (Weintraub & Beanland, 2006). Because phytoplasma are difficult to cultivate outside their hosts, their study is limited to infected plant tissue. Phytoplasma are the causal agents of Grapevine Yellows diseases, which affect global grapevine production. Infected plants exhibit several symptoms, such as leaf yellowing or reddening, leaf curling, tissue necrosis, and shriveling of berries (Dermastia, Bertaccini, Constable, & Mehle, 2017). On a subcellular level, phytoplasma impede photosynthetic processes and manipulate the host's metabolism, particularly carbohydrate

synthesis, to their advantage (Albertazzi et al., 2009; Bertamini, Nedunchezhian, Tomasi, & Grando, 2002; Hren et al., 2009; Santi, Grisan, Pierasco, De Marco, & Musetti, 2013). Two of the most damaging Grapevine Yellows in Europe are the "flavescence dorée", associated with subgroups 16SrV-C and 16SrV-D (Martini, Murari, Mori, & Bertaccini, 1999), and "bois noir", associated with '*Ca.* P. solani' from the stolbur subgroup 16SrXII-A (Quaglino et al., 2013). The pathogenesis of phytoplasma infection has been mainly studied in grapevine infected with "bois noir" due to the "flavescence dorée" quarantine status (Dermastia et al., 2017).

1.1.1.1 '*Candidatus* Phytoplasma solani'

The phytoplasma species '*Ca.* P. solani' was first described in 2013 by Quaglino *et al.* and is represented by different strains that form a unique genetic cluster determined on the basis of phylogenetic analysis of the 16S rRNA, *tuf*, *secY,* and *rplV-rpsC* genes and the assessment of biological properties. The '*Ca.* P. solani' strains cause "bois noir" in grapevine, the stolbur disease in tomato (Pracros, Renaudin, Eveillard, Mouras, & Hernould, 2006), potato (Mitrović et al., 2016), and other wild or cultivated plants, maize redness (Jović et al., 2007), and other diseases. Its primary hosts among uncultivated plants are the stinging nettle (*Urtica dioica*) and bindweed (*Convolvulus arvensis* and *Calystegia sepium*). Although endemic to southern Europe and the Mediterranean area, the presence of '*Ca.* P. solani' has been reported in vineyards of many European and Asian countries, as well as North America, South America, and Oceania, according to the European and Mediterranean Plant Protection Organization Global Database ("EPPO Global Database," 2022). In the European and Mediterranean region, '*Ca.* P. solani' is mainly transmitted by the planthopper vector *Hyalesthes obsoletus* Signoret, which frequently feeds on stinging nettle or bindweed, yet only occasionally on grapevines, making them an incidental or a dead-end host for the phytoplasma (Dermastia et al., 2017).

The grapevine "bois noir" causes symptoms similar to other Grapevine Yellows diseases. None of the grapevine cultivars studied thus far have proved resistant to phytoplasma infection (Albertazzi et al., 2009; Laimer et al., 2009); however, some plants are able to recover spontaneously from "bois noir", with disease symptoms suddenly declining or disappearing (Murolo, Garbarino, Mancini, & Romanazzi, 2020). Although the processes underlying grapevine recovery are not well understood, it is accompanied by biochemical changes to the phloem, callose deposition, changes in sugar transport and metabolism, and has been connected to reactive oxygen species production (Gambino, Boccacci, Margaria, Palmano, & Gribaudo, 2013; Santi, De Marco, Polizzotto, Grisan, & Musetti, 2013).

1.1.2 Viral Infections

Viruses are obligate parasites, unable to replicate without the use of their host's intracellular machinery. According to the 2022 update of the International Committee on Taxonomy of Viruses (https://ictv.global/taxonomy/), about 2000 species of plant-infecting viruses are distributed among 38 families (Walker et al., 2021). The genomes of plant viruses are usually small yet often complex, harboring overlapping open reading frames and encoding multifunctional proteins, such as the potyviral Helper Component Proteinase (Valli, Gallo, Rodamilans, López-Moya, & García, 2018). Although they can have different genome organizations, about 80% of all plant viruses are RNA viruses (Mandahar, 2006). RNA viruses have high genetic variability, and a single viral species is usually comprised of many variants (Rubio, Galipienso, & Ferriol, 2020). Once the viral particle enters the host cell, its previously encapsulated genome is released and the virus starts replicating at the site of infection. The virus can then spread throughout the host

using cell-to-cell and long-distance movement, ultimately causing a systemic infection (Calil & Fontes, 2017; Križnik, Gruden, & Baebler, 2020).

While they can be transmitted in several ways, transmission by insect vectors is the most efficient and also the most widespread with insect species from the order Hemiptera transmitting over 70% of known plant viruses (Lefeuvre et al., 2019). Insects can transmit viruses in a non-persistent/semi-persistent or persistent manner. In non-persistent or semi-persistent transmission, the virus reversibly interacts with the components of the insect's mouthparts and can be shed by salivary secretion. Persistently transmitted viruses are ingested by the insect and enter the salivary glands through the hemolymph. They can remain there until the end of the insect's lifetime, repeatedly infecting plants on which it feeds. Sometimes the viruses can even influence the behavior of their insect vectors or plant hosts and manipulate them to their advantage (Lefeuvre et al., 2019).

1.1.2.1 Potato Virus Y

PVY is a filamentous plant virus from the *Potyviridae* family (genus *Potyvirus*) that infects a range of economically important crops. It is the most dangerous virus infecting potato (Kreuze et al., 2019) and the causal agent of potato tuber necrotic ringspot disease, which can severely affect tuber production. PVY is transmitted by aphids, which feed on the leaves of infected plants (Karasev & Gray, 2013). Its single-stranded positive-sense RNA genome encodes for eleven viral proteins, which are largely multifunctional and are involved in various stages of viral infection, including replication, movement, and/or other processes. The viral particle is about 730 nm long and flexuous, with over 2000 viral coat protein copies assembled around the RNA genome in a left-handed helical arrangement (Kežar et al., 2019).

PVY infection can elicit different responses in potato plants, depending on the viral strain and potato cultivar. The responses of susceptible cultivars range from no visible symptoms in tolerant plants to severe disease symptoms and tuber necrosis in sensitive cultivars (Baebler, Coll, & Gruden, 2020). Resistant potato cultivars respond to infection with extreme resistance or hypersensitive response, limiting viral replication and movement through mechanisms not well understood (Baebler et al., 2020). The hypersensitive response is prompted by the *Ny-1* resistance gene in the potato cultivar Rywal and manifests in the form of necrotic lesions at the infection site and inhibition of systemic viral spread (Szajko et al., 2008). At the molecular level, the hypersensitive response involves several processes, such as reactive oxygen species production, callose deposition, cell wall strengthening, and induction defense gene expression (Baebler et al., 2020). Studies with transgenic potato plants, impaired in accumulation of salicylic acid (SA), showed that the SA hormonal pathway plays a crucial role in potato immune response to PVY infection (Baebler et al., 2011, 2014; Lukan, Baebler, et al., 2018; Lukan et al., 2020); however, the mechanisms of SA perception in potato have not been investigated.

1.2 Plant Immune Response

Due to their sessile nature, plants' response to pathogen infection largely depends on intracellular processes. In the innate immune response, pattern recognition receptors (PRRs) in the plasma membrane recognize the conserved molecular patterns of invading pathogens known as pathogen-associated molecular patterns, activating the so-called pattern-triggered immunity (PTI). Some pathogens evade the PTI response by releasing effectors, virulence proteins or secondary metabolites that inhibit its development. In response, plants can recognize effectors via intracellular receptors, mostly belonging to the

nucleotide-binding leucine-rich repeat receptor (NLR) family of proteins. This is called the effector-triggered immunity (ETI) (Jones & Dangl, 2006).

A gene-for-gene concept, pioneered by Harold Flor in 1942, has long been valid and is based on the hypothesis that individual genes are responsible for the outcome of plant-pathogen interactions: one effector molecule is recognized by one plant receptor (Flor, 1942). However, it is becoming apparent that effector recognition involves complex networks, the result of hundreds of millions of years of competitive evolution. Many pathogens have a diverse set of effectors recognized directly or indirectly by plant receptors, which sometimes require other molecules to act before they can transmit the signal. There are many differences between receptors of different plant species due to the fast evolution of effector proteins; this redundancy makes receptor networks much more robust and resilient to changes in the environment (C.-H. Wu, Derevnina, & Kamoun, 2018). Plant receptors often contain so-called integrated domains that are targeted by the effector. These can include domains of different plant proteins, often transcription factors. Increasing numbers of these domains have now been discovered in plant receptors and are thought to serve as decoys that effectors bind instead of the actual targets, leading to recognition by the receptor (Kroj, Chanclud, Michel-Romiti, Grand, & Morel, 2016).

Pathogen recognition through PTI or ETI is followed by a series of downstream signaling cascades that lead to the activation of various defense mechanisms. Although PRRs and NLRs involve different activation mechanisms and early signaling components, PTI and ETI eventually result in similar outcomes, including the elevation of intracellular calcium levels and reactive oxygen species, activation of mitogen-activated protein kinases, and ultimately transcriptional reprogramming. ETI is also often accompanied by programmed cell death and hypersensitive response. In PTI, signal transduction is mediated by the protein kinase activity of PRRs and intracellular protein kinases, as well as by various co-receptor proteins, whereas the signaling pathways of ETI are less well known (Yuan, Ngou, Ding, & Xin, 2021). Upon activation, NLRs have been shown to form oligomeric complexes called resistosomes (Martin et al., 2020; J. Wang et al., 2019), which are thought to form pores in the cell membrane, facilitating calcium influx or even leading to cell death. Recent research shows that PRRs and NLRs function synergistically to ensure a fully functional immune response during ETI (Yuan, Jiang, et al., 2021). Additionally, Ngou, Ahn, Ding, & Jones (2021) showed that NLRs potentiate protein activation in PTI, while the hypersensitive response of ETI is strongly enhanced by the activation of PPRs. The interconnectedness of PTI and ETI thus seems to be a key aspect of plant immunity.

Although most phytopathology studies focus on understanding the interplay between host receptors and pathogen effector proteins, small RNAs (sRNAs) and RNA interference became known as vital regulators of immunity-related gene expression and transcriptional reprogramming in plant-pathogen interactions (Huang, Wang, Hu, Hamby, & Jin, 2019). Plant sRNAs can be classified as micro RNAs or small-interfering RNAs and are usually generated by Dicer or Dicer-like enzymes. They are then incorporated into Argonaute proteins to form an RNA-induced silencing complex and induce gene silencing in a sequence-specific manner, through the degradation of target messenger RNA (mRNA), the inhibition of translation, or transcriptional gene silencing (Huang et al., 2019; Križnik et al., 2020). For example, RNA interference is an important defense mechanism against viral infections, with the viral genome as its target (Križnik et al., 2020). It also provides an additional level of regulation of host immune signaling (Križnik et al., 2017). Some sRNAs can even cross the boundaries between hosts and pathogens, silencing genes in interacting organisms, a mechanism referred to as ‘‘cross-kingdom RNA interference’’ (Huang et al., 2019).

1.2.1 Hormonal Signaling

The complex intracellular signaling cascades required for the establishment of an efficient defense response are coordinated by plant hormones. Major plant hormones involved in immune response include salicylic acid (SA), jasmonic acid, abscisic acid, ethylene, and auxins. SA (2-hydroxy benzoic acid), a phenolic compound and a secondary plant metabolite, is the main hormone regulating plant defenses and promotes immunity against biotrophic and semibiotrophic pathogens (Peng, Yang, Li, & Zhang, 2021). In the absence of stress, most plant species maintain relatively low basal levels of SA, which can rapidly increase after infection. SA is synthesized from chorismate through either the isochorismate or phenylalanine ammonia lyase pathway and its production in response to stress is tightly regulated through the modulation of enzymes included in its biosynthesis. SA homeostasis is also modulated by chemical modifications, including hydroxylation, glycosylation, methylation, and amino acid conjugation (Peng et al., 2021). The most studied SA receptors are the Non-Expressor of Pathogenesis Related (NPR) transcription cofactors (W. Wang et al., 2020; Y. Wu et al., 2012), key regulatory proteins involved in immunity-related transcriptional reprogramming. While the Arabidopsis AtNPR1 functions as the main positive regulator of plant immunity, its paralogs AtNPR3 and AtNPR4 act as redundant negative regulators (Ding et al., 2018; Zhang et al., 2006). Additionally, numerous SA-binding proteins with different affinities for SA were biochemically identified (Klessig, Tian, & Choi, 2016); however, their function remains unknown.

1.2.2 Transcriptional Regulation

Dynamic adaptations of plant cell function are largely dependent on the transcriptional regulation of gene expression. Plants are equipped with a wide arsenal of transcription factors, activating or repressing gene expression while heeding the signals from hormonal pathways. Transcription factors regulate transcription by binding to short DNA sequence patterns called motifs in gene promoter regions, influencing the recruitment of RNA-polymerase II to the transcription initiation site. Eukaryotic transcription initiation is highly complex as it involves multiple transcription factors as well as interacting cofactors, collectively modulating the transcription of a single gene (Reiter, Wienerroither, & Stark, 2017; Roeder, 2019). Depending on the conditions and cofactors present, transcription factors can act as activators, positively affecting gene expression, or repressors, which prevent transcription. While transcription factors interact with target motifs primarily through their DNA-binding domain, they usually contain additional domains enabling interactions with other regulatory proteins (Gonzalez, 2016).

Several plant transcription factor families have been identified as important regulators of plant immunity, including the APETALA2/ETHYLENE-RESPONSE ELEMENT BINDING FACTOR family, the basic-helix-loop-helix family, the WRKY family, and others (Tsuda & Somssich, 2015). SA signaling is directly connected to SA-receptors, NPR cofactors (W. Wang et al., 2020; Y. Wu et al., 2012), and TGACG-binding (TGA) transcription factors, which cooperatively modulate the expression of key defense-related genes and genes involved in SA synthesis (Ding et al., 2018; Zhang, Fan, Kinkema, Li, & Dong, 1999).

1.2.2.1 TGA Transcription Factors

TGA transcription factors are members of the basic region leucine zipper (bZIP) protein family and are indispensable regulators of gene expression in various cellular processes, from biotic and abiotic stress (Zhang, Tessaro, Lassner, & Li, 2003; Zhong et al., 2015) to plant growth and development (Murmu et al., 2010). TGAs bind to their target motif, the

TGACG sequence, and its variations, through dimerization, while they can also form higher order complexes (Boyle et al., 2009; Niggeweg, Thurow, Weigel, Pfitzner, & Gatz, 2000; Schiermeyer, Thurow, & Gatz, 2003) and interact with other regulatory proteins (Chen et al., 2019; Hussain, Sheikh, Haider, Quareshy, & Linthorst, 2018; Li et al., 2019). DNA recognition and binding of TGA factors occur through a conserved bZIP DNA-binding domain, while their N-terminal and C-terminal regions are involved in establishing interactions with other proteins and have a transcription activation function (Boyle et al., 2009; Chai et al., 2020; Fan & Dong, 2002; Kumar et al., 2022).

Most notably, TGAs act as NPR-interacting proteins and serve as DNA-binding anchors in the NPR-TGA-DNA regulatory complex at target promoters (Boyle et al., 2009; Kumar et al., 2022; Zhou et al., 2000). Their mechanism of action has mainly been studied in relation to transcriptional regulation of the *Pathogenesis related-1* promoter, where the AtNPR1 interacts with AtTGA2 to activate gene expression (Boyle et al., 2009; Rochon, Boyle, Wignes, Fobert, & Després, 2006; Zhang et al., 2003). The structural aspects of NPR-TGA interaction have been revealed only recently by Kumar *et al.* (2022), showing that an AtNPR1 dimer bridges two DNA-bound AtTGA3 homodimers through their C-terminal regions.

1.3 Protein Functional Analysis Approaches

Thorough understanding of cellular processes and the function of specific molecular components is a crucial part of plant immune response research. Protein functional analysis is complex and requires the use of different approaches, most often beginning with target identification through "omics" techniques, including genomics, transcriptomics, proteomics, and metabolomics. The function of a specific protein is, to an extent, defined by its structural characteristics and modulated by various external parameters, such as interactions with other biological molecules or post-translational modifications. Sequence analyses *in silico*, studying interactions with different biological molecules, tissue and subcellular localization, enzyme activity assays, and other experimental techniques provide essential information about protein behavior. It is also important to consider how the protein of interest is integrated into the overall system and by which mechanisms it may be regulated. Structural data, as well as determining the contribution of individual protein parts, allow for a comprehensive analysis of protein activity, yet they are often harder to obtain (Alberts et al., 2002). While the determination of plant protein three-dimensional (3D) structures is steadily increasing, their number still lags behind the number of structures obtained from other organisms.

1.4 Aims of the Research

1.4.1 Adaptation of Experimental Approaches for Crop Studies

Experimental techniques used for plant molecular analyses are either based on established methodologies used in other experimental systems or have been developed specifically for studying plants. However, most of these methods have been applied to model plant species, whereas the analysis of less studied crops often requires further adaptation of basic and/or advanced techniques to address the specific traits of particular species or even cultivars. In the scope of this thesis, our **first aim** is to adapt and apply experimental techniques for studying the grapevine-phytoplasma pathosystem on a molecular level.

1.4.2 Identification of Protein Targets of Phytoplasma Infections

Phytoplasma-mediated disease development is closely linked to the secretion and action of effectors. Resolving the genomic sequences of several phytoplasma allowed the identification of secretory proteins that could act as effectors, but few of them have been described thus far (Bertaccini, Oshima, Maejima, & Namba, 2019). For example, '*Ca.* P. asteris' secreted AY-WB protein 11 (SAP11) effector was found to localize in cell nuclei, where it can destabilize Arabidopsis CINCINNATA-TEOSINTE BRANCHED1, CYCLOIDEA, PROLIFERATING CELL FACTORS through protein-protein interactions, leading to changes in leaf shape and stem proliferation (Bai et al., 2009; Sugio, Kingdom, MacLean, Grieve, & Hogenhout, 2011; Sugio, Maclean, & Hogenhout, 2014). A SAP11 homologue and other putative effector genes have been identified in the '*Ca.* P. solani' strain SA-1 genome (Music et al., 2019), but to our knowledge, their function remains unexplored. Our **second aim** is to explore the pathogenesis of '*Ca.* P. solani' infection, through identification of effector targets in grapevine.

1.4.3 Analysis of TGA Transcription Factors in Potato Immunity

TGA transcription factors are among the most important components of SA-mediated transcriptional regulation following biotic stress, and their involvement in pathogen infection has been reported in several crops, such as rice (Moon et al., 2018), tobacco (Thurow et al., 2005) and tomato (Ekengren, Liu, Schiff, Dinesh-Kumar, & Martin, 2003). The SA signaling pathway is also crucial in potato immune response to PVY infection (Baebler et al., 2014), yet little is known about the role of TGAs in this crop. In order to improve our understanding of the molecular mechanisms underlying potato immunity, our **third aim** is to identify the Arabidopsis TGA orthologues in potato, involved in the immune response to PVY. Finally, the **fourth aim** of this doctoral thesis is to study the molecular interactions of potato TGAs and their mechanisms of transcriptional regulation.

1.5 Research Hypotheses

1. Optimization of established methodology for studying plant processes will enable efficient analysis of key molecular components in phytoplasma infection of grapevine.

2. The putative '*Candidatus* Phytoplasma solani' effectors interact with grapevine proteins.

3. The mechanisms of potato TGA function differ from those of their orthologues in the model plant *Arabidopsis thaliana*.

4. Potato TGA transcription factors are important in the salicylic acid-mediated potato immune response against potato virus Y.

5. Heterodimerization of potato TGAs modulates transcriptional regulation of genes involved in potato immune response.

1.6 Publications Included and Candidate's Contributions

Our first publication (*Differential Response of Grapevine to Infection with* 'Candidatus *Phytoplasma solani' in Early and Late Growing Season Through Complex Regulation of*

mRNA and Small RNA Transcriptomes) focuses on unraveling the transcriptional changes in gene expression and sRNA levels in grapevine during '*Ca.* P. solani' infection before and after symptom development. The PhD candidate is a co-author of this publication. She was involved in the optimization of the RNA isolation procedure from grapevine leaf vein-enriched samples and isolated total RNA from healthy and phytoplasma-infected samples, collected in two seasons of the first year of analysis, using the optimized protocol. She also wrote the "RNA Extraction and Sequencing" methods section.

Our second publication (*Candidate Effector Proteins of* 'Candidatus *Phytoplasma solani*' *are Associated with Modulation of Plant Carbohydrate Metabolism Toward Effective Glycolysis, with Accelerated Ascorbate-Glutathione Cycle, and with Induction of Autophagosomes,* manuscript draft), connected to studying grapevine-phytoplasma interactions, focuses on the identification and characterization of putative '*Ca.* P. solani' effectors and their target proteins in grapevine. The PhD candidate is a co-author of this publication. She has adapted and performed a protein pull-down assay for the detection of effector targets. She prepared the samples for mass spectrometry analysis, searched for orthologues of identified tobacco targets in grapevine, and confirmed the identified interactions with co-immunoprecipitation. The PhD candidate also wrote the results and methods sections connected to these experiments. Additional co-immunoprecipitation assay results, not included in the manuscript, are included in Appendix A.

The third publication (*TGA Transcription Factors – Structural Characteristics as Basis for Functional Variability*) provides a concise summary of plant TGA transcription factors from a structural point of view and relates their structural characteristics to their functional roles in transcription regulation, focusing on the characterized TGA factors from Arabidopsis and tobacco. The PhD candidate is the first author of this publication. She studied the available reports on TGA factor structural and functional characteristics, conducted *in silico* protein sequence analyses, and wrote the first manuscript draft.

Our fourth publication (*A Mini-TGA Protein Modulates Gene Expression through Heterogeneous Association with Transcription Factors,* currently accepted under minor revisions) focuses on the identification of potato TGAs and characterizes the StTGA2.1 protein, which has unusually compact molecular architecture and lacks a complete DNA-binding domain. We referred to this particular type of TGAs as mini-TGAs. The PhD candidate is the first author of this publication. She planned and performed *in silico* analyses, protein-protein interaction experiments, and plant immunity studies. She was actively involved in localization experiments, protein-DNA interaction studies, targeted genomic sequencing, and gene expression analyses and wrote the first draft of the manuscript.

Although the results have not yet been included in a publication, we also analyzed the subcellular localization of two potato NPR proteins and performed preliminary co-localization studies of StNPR1 with StTGA2.1 proteins (Appendix B) to complement our results on StTGA2.1 function published in the fourth publication. The PhD candidate drafted the experimental design and performed the analyses. The results obtained are included in the Discussion.

Chapter 2

Scientific Publications

2.1 Differential Response of Grapevine to Infection with '*Candidatus* Phytoplasma solani' in Early and Late Growing Season through Complex Regulation of mRNA and Small RNA Transcriptomes

Marina Dermastia, Blaž Škrlj, Rebeka Strah, Barbara Anžič, Špela Tomaž, Maja Križnik, Christina Schönhuber, Monika Riedle-Bauer, Živa Ramšak, Marko Petek, Aleš Kladnik, Nada Lavrač, Kristina Gruden, Thomas Roitsch, Günter Brader and Maruša Pompe-Novak

International Journal of Molecular Sciences, 2021, 22:3531. DOI: 10.3390/ijms22073531

In our first publication, we investigated the changes in gene expression levels and sRNA profiles in grapevine cv. Zweigelt due to '*Ca.* P. solani' infection through high-throughput transcriptomic analyses. Comparing results from grapevine leaf-vein-enriched samples collected in early and late growing seasons, we found high transcriptional activity in response to infection already in the early season, while the regulation of sRNAs was more pronounced in the late season. Our analysis showed most of the regulated genes and sRNAs were associated with biotic stress and revealed the perturbations to hormonal pathways due to infection. We further analyzed transcriptomic data using the network enrichment methodology to extract distinct gene groups formed at different time points, in different plant groups, and those that showed disintegration between growing seasons and between infected and healthy plants. Our results uncovered new genes that have previously not been associated with phytoplasma infection.

The PhD candidate was involved in the optimization of the RNA isolation procedure from grapevine leaf-vein-enriched samples to obtain total RNA, including sRNAs, of high quality, purity, and integrity. With the optimized procedure, she isolated total RNA from healthy and phytoplasma-infected samples, collected in two seasons of the first year of analysis, and wrote the “RNA Extraction and Sequencing” methods section.

International Journal of
Molecular Sciences

MDPI

Article

Differential Response of Grapevine to Infection with '*Candidatus* Phytoplasma solani' in Early and Late Growing Season through Complex Regulation of mRNA and Small RNA Transcriptomes

Marina Dermastia [1,*], Blaž Škrlj [2,3], Rebeka Strah [1,3], Barbara Anžič [1], Špela Tomaž [1,3], Maja Križnik [1], Christina Schönhuber [4], Monika Riedle-Bauer [5], Živa Ramšak [1], Marko Petek [1], Aleš Kladnik [6], Nada Lavrač [2], Kristina Gruden [1], Thomas Roitsch [7], Günter Brader [4] and Maruša Pompe-Novak [1,8]

1 National Institute of Biology, 1000 Ljubljana, Slovenia; rebeka.strah@gmail.com (R.S.); barbara.anzic@gmail.com (B.A.); spela.tomaz@nib.si (Š.T.); maja.kriznik@nib.si (M.K.); ziva.ramsak@nib.si (Ž.R.); marko.petek@nib.si (M.P.); kristina.gruden@nib.si (K.G.); marusa.pompe.novak@nib.si (M.P.-N.)
2 Jožef Stefan Institute, 1000 Ljubljana, Slovenia; blaz.skrlj@ijs.si (B.Š.); nada.lavrac@ijs.si (N.L.)
3 Jožef Stefan International Postgraduate School, 1000 Ljubljana, Slovenia
4 Bioresources Unit, Austrian Institute of Technology, 3430 Tulln, Austria; christina.schoenhuber@gmail.com (C.S.); guenter.brader@ait.ac.at (G.B.)
5 Federal College and Research Institute for Viticulture and Pomology, 3400 Klosterneuburg, Austria; monika.riedle-bauer@weinobst.at
6 Department of Biology, Biotechnical Faculty, University of Ljubljana, 1000 Ljubljana, Slovenia; ales.kladnik@bf.uni-lj.si
7 Department of Plant and Environmental Sciences, University of Copenhagen, 2630 Taastrup, Denmark; roitsch@plen.ku.dk
8 School of Viticulture and Enology, University of Nova Gorica, 5271 Vipava, Slovenia
* Correspondence: marina.dermastia@nib.si

check for updates

Citation: Dermastia, M.; Škrlj, B.; Strah, R.; Anžič, B.; Tomaž, Š.; Križnik, M.; Schönhuber, C.; Riedle-Bauer, M.; Ramšak, Ž.; Petek, M.; et al. Differential Response of Grapevine to Infection with '*Candidatus* Phytoplasma solani' in Early and Late Growing Season through Complex Regulation of mRNA and Small RNA Transcriptomes. *Int. J. Mol. Sci.* **2021**, *22*, 3531. https://doi.org/10.3390/ijms22073531

Academic Editors: Giorgio Gambino and Chiara Pagliarani

Received: 23 February 2021
Accepted: 25 March 2021
Published: 29 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Bois noir is the most widespread phytoplasma grapevine disease in Europe. It is associated with '*Candidatus* Phytoplasma solani', but molecular interactions between the causal pathogen and its host plant are not well understood. In this work, we combined the analysis of high-throughput RNA-Seq and sRNA-Seq data with interaction network analysis for finding new cross-talks among pathways involved in infection of grapevine cv. Zweigelt with '*Ca.* P. solani' in early and late growing seasons. While the early growing season was very dynamic at the transcriptional level in asymptomatic grapevines, the regulation at the level of small RNAs was more pronounced later in the season when symptoms developed in infected grapevines. Most differentially expressed small RNAs were associated with biotic stress. Our study also exposes the less-studied role of hormones in disease development and shows that hormonal balance was already perturbed before symptoms development in infected grapevines. Analysis at the level of communities of genes and mRNA-microRNA interaction networks revealed several new genes (e.g., expansins and cryptdin) that have not been associated with phytoplasma pathogenicity previously. These novel actors may present a new reference framework for research and diagnostics of phytoplasma diseases of grapevine.

Keywords: '*Candidatus* Phytoplasma solani'; grapevine; bois noir; RNA-Seq; sRNA-Seq; miRNA; phasiRNA; hormones; interaction network

1. Introduction

Bois noir is the most widespread phytoplasma disease of grapevines (*Vitis vinifera* L.) in Europe and can lead regionally to losses of up to 50% [1]. Its causal agent is '*Candidatus* Phytoplasma solani' from the stolbur group 16SrXII-A [2]. The interactions between grapevine and '*Ca.* P. solani' have been extensively studied and are reviewed elsewhere [3].

Several developmental processes and metabolic pathways in the host plant are affected by infection as revealed at the transcriptional, protein, and metabolic levels [4–8]. In infected grapevines, several steps in photosynthesis are repressed during infection. There is growing evidence that feedback inhibition of photosynthesis results in leaf yellowing (i.e., chlorosis) because of carbohydrate accumulation in the source leaves [9–14]. As was proposed early in phytoplasma research, this accumulation is a consequence of the manipulation of the host metabolism by the phytoplasmas, which can turn infected plant tissues into a carbohydrate sink that provides phytoplasmas with hexoses [8,15,16]. Several studies have analyzed the expression of genes involved in carbohydrate metabolism and their enzyme products and sugar metabolites upon infection of grapevines with '*Ca.* P. solani' [8,16–20]. As a sign of stress conditions, grapevines infected with '*Ca.* P. solani' accumulate several amino acids, including serine, glycine, valine, leucine, alanine, β-alanine, threonine, aspartate, pyroglutamate, and proline [17]. In addition, infection with '*Ca.* P. solani' changes the flavonoid pathways in grapevine [21–26]. Although evidence points to important roles of the plant hormones in the signaling networks involved in grapevine responses to '*Ca.* P. solani', the underlying molecular mechanisms of these interactions remain poorly understood [27].

While the main plant processes in which phytoplasmas are involved are now recognized, information is scarce in terms of fine-tuning the expression of the regulators involved, such as transcription factors and small RNAs (sRNAs). The combination of these factors defines the genetic regulatory circuits of transcriptional control [28,29]. sRNAs are involved in multiple cellular processes, and they modulate the expression of other regulators, including transcription factors [28,30]. In plants, two types of sRNAs regulate posttranscriptional gene expression, microRNAs (miRNAs) and phased small-interfering RNAs (phasiRNAs), although their targets are not well defined. However, several conserved and species-specific miRNAs have been identified in grapevines [31]. It has also been shown that miRNAs have an important role in host plant responses to phytoplasma infections. They have been identified in phytoplasma-infected Mexican lime [32], *Ziziphus jujuba* [33], mulberry [34], and paulownia [35], together with their putative targets, which are genes involved in plant morphology, signaling, nutrient homeostasis, environmental stress responses, and hormonal metabolism and regulation. In addition, several miRNAs that respond to phytoplasma infection have been identified in grapevines infected with Flavescence dorée phytoplasmas [36] and with '*Ca.* P. asteris' [37].

New high-throughput technologies such as whole transcriptome sequencing allow simultaneous analysis of multiple gene expression snapshots of the same organism or of its tissues. These can be used for more accurate analysis of time-dependent phenomena, such as phytoplasma pathogenesis. In the present study, we addressed the temporal dynamics of the responses of grapevine cv. Zweigelt (Rotburger) to infection with '*Ca.* P. solani' under natural vineyard conditions in leaf-vein-enriched samples and analyzed gene expression by RNA-Seq and sRNA profiles by sRNA-Seq.

As molecules in the cell seldom work completely independently, several interaction network-based approaches have been adopted as tools of choice to study their interconnectivity [38]. In a study parallel to the present one, we developed methods that can operate with such information-rich structures and applied these to modeling the bois noir disease [39]. The network inference we described offers an exploration of temporal network dynamics at the level of communities that can be revealed from RNA-Seq data. This method was developed on grapevine cv. Zweigelt infected with '*Ca.* P. solani' late in the growing season. In the present study, we aimed to apply this method to grapevines sampled early in the growing season prior to symptom development, which has not been investigated in detail to date, and to perform a comparative analysis with the results of the late growing season. In addition, to reveal hierarchical interactions and co-regulation among different RNA classes involved in these grapevine infections, data-driven bioinformatics approaches were used to analyze the genome-wide data obtained over two consecutive growing seasons. We have identified complex regulatory networks that pro-

vide us with a reference framework for detailed studies of the regulation of the main processes in grapevines infected with '*Ca.* P. solani'.

2. Results and Discussion

2.1. mRNAs and sRNAs Show Differential Temporal Involvement in Grapevine Infected with 'Ca. P. solani'

High-throughput RNA-Seq (Table S1) and sRNA-Seq (Table S2) were performed on leaf-vein-enriched samples collected in the early and late growing seasons in 2017; for mRNAs, this was repeated again early in the season of 2020. For the 2017 experiment, the grapevines were chosen based on checking for phytoplasma infection the previous year, as presumably uninfected and infected with '*Ca.* P. solani'. The presumed infected grapevines were asymptomatic at the time of the first sampling in the early growing season. However, they developed symptoms over the growing season, and the phytoplasma infection with a nettle type (CPsM4_At1 [40]) was confirmed at the time of the second sampling in the late growing season. The same grapevines were sampled again in the early growing season of 2020. In 2020, the grapevines did not develop symptoms over the summer, and phytoplasmas were not detected in samples from these grapevines later in the growing season, which suggested that these grapevines had already recovered from their previous infection or were in the recovery process [41]. Moreover, the RNA-Seq analysis of samples from these grapevines revealed that there was not even a single differentially expressed gene defined between the recovered and uninfected grapevines (Table S1), which indicated that the recovery appeared to have occurred previously, in 2018 or 2019 [20]. Of note, these samples were collected in a production vineyard where the vines are pruned extensively each year, which is a practice that can have positive effects on grapevine recovery from bois noir disease [42].

Analysis of the RNA-Seq data produced an average of 82,335,298 reads per sample that were aligned to the grapevine reference genome. Out of the 42,413 annotated genes, 24,279 were removed by filtering out those with raw counts over 50 in less than four samples (i.e., genes that were not expressed under any of the conditions). From the remaining 18,134 genes, 15,319 genes (correspond to 84% of the 18,134 expressed genes, and 36% of the 42,413 annotated genes) were significantly differentially expressed in at least one of the comparisons (Table S1).

sRNA-Seq revealed 178 miRNAs (Table S2) and 261 phasiRNAs (Table S3) with raw counts over 50 in at least four samples. Of these, 238 (54%) were significantly differentially expressed. Eighteen of the miRNAs were novel (18 novel MIR loci) and might be involved in novel miRNA-mRNA interactions. Using our pipelines, the novel miRNAs were grouped into 17 novel miRNA families, with two of the loci grouped into the same family, miR14 (Table S4). The prevailing number of differentially expressed sRNAs in the infected grapevines (i.e., 125 sRNAs) in comparison with the number of differentially expressed sRNAs in uninfected plants (i.e., 54 sRNAs) suggests an important role for sRNAs in grapevine responses to infection with '*Ca.* P. solani'.

Venn diagrams reveal some interesting comparisons (Figure 1). Gene expression obtained by RNA-Seq was initially compared between the infected and uninfected grapevine samples (Figure 1a). A total of 6942 and 6288 genes were significantly differentially expressed in infected grapevines compared to uninfected grapevines in the early and late growing seasons, respectively. Together with the same average absolute $\log_2$ value of fold-change in both cases (i.e., 0.70), these data indicate that the grapevine transcriptional response to infection with '*Ca.* P. solani' in the early season is similar to that of the late growing season. For the differentially expressed genes, 35% (3438 genes) were expressed differentially in both growing seasons, while the remaining 6354 genes were expressed differentially, as either in the early (3504 genes) or the late (2850 genes) growing season.

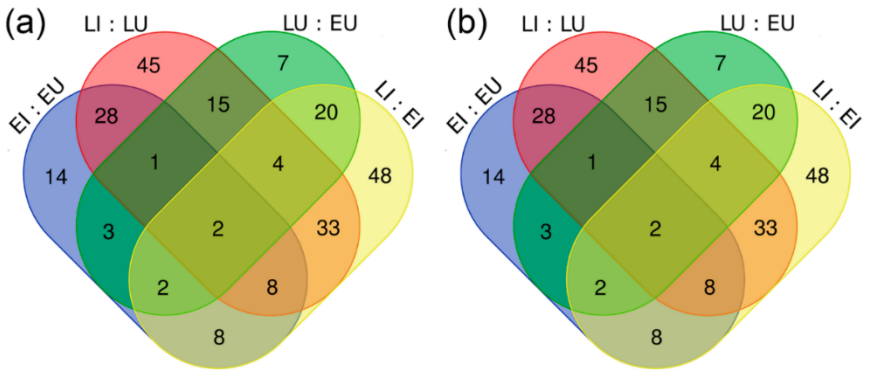


Figure 1. Venn diagrams showing the significantly differentially expressed genes (**a**) and sRNA (**b**) from the 2017 experiment. E, early growing season; L, late growing season; U, uninfected; I, infected with '*Ca.* P. solani'.

In the next step, gene expression was compared between grapevine samples from the late and early growing seasons (Figure 1a). In total, 10,941 and 10,998 genes were significantly differentially expressed in the late growing season compared to the early growing season in the uninfected and infected grapevine samples, respectively. Together with the same average absolute $\log_2$ value of fold-change in both cases (i.e., 1.03), these data suggest that the infection itself did not affect the number of differentially expressed genes during the growing seasons. Moreover, 60% (8187) of the genes were expressed differentially in the uninfected and infected grapevines, while the other 40% (5565) of the genes were expressed differentially either in uninfected (2754 genes) or infected (2811 genes) grapevines.

In contrast with the gene expression revealed, the Venn diagram of sRNAs shows a different picture (Figure 1b). In infected grapevines compared to uninfected grapevines, 66 and 136 sRNAs were differentially expressed in the early or late growing season, respectively. Thirty-nine sRNAs (19%) were expressed differentially in both growing seasons, while the other 163 sRNAs (81%) were expressed differentially either in the early (27 sRNAs) or the late (97 sRNAs) growing season.

Most phytoplasma studies have been focused on the symptomatic phases of pathogenesis, and the consequent conclusions for the more metabolically dynamic late growing season phase are based on the associated results [3,43]. In this regard, the results of the present study are striking, as they show that the asymptomatic early phase of the annual development of bois noir disease is very active at the grapevine transcriptional level, whereas its regulation at the level of sRNAs is more pronounced later on.

Analyzed miRNA and phasiRNA targets have confirmed the growing evidence that any single miRNA might regulate many genes, as well as that any single gene might be targeted by a number of miRNAs (Figure 2; Table S2). An interesting observation was that 57% and 53% of the miRNAs detected were down-regulated in grapevines infected with '*Ca.* P. solani' in the early and late growing seasons, respectively (Figure 2). The significance of this finding is not known at the moment. However, in animals, a global down-regulation of miRNA expression is often the rule in cancers [44]. While 80% of the same miRNAs were similarly differentially expressed either early or late in the growing season, 20% of all miRNAs changed their mode of up-regulation or down-regulation through the year (Figure 2). Among these, there are the most prominent isomiRs with sequences that varied with respect to the reference sequences of *vvi-miR166* and *vvi-miR3623* that regulate genes that encode proteins involved in disease resistance [45].

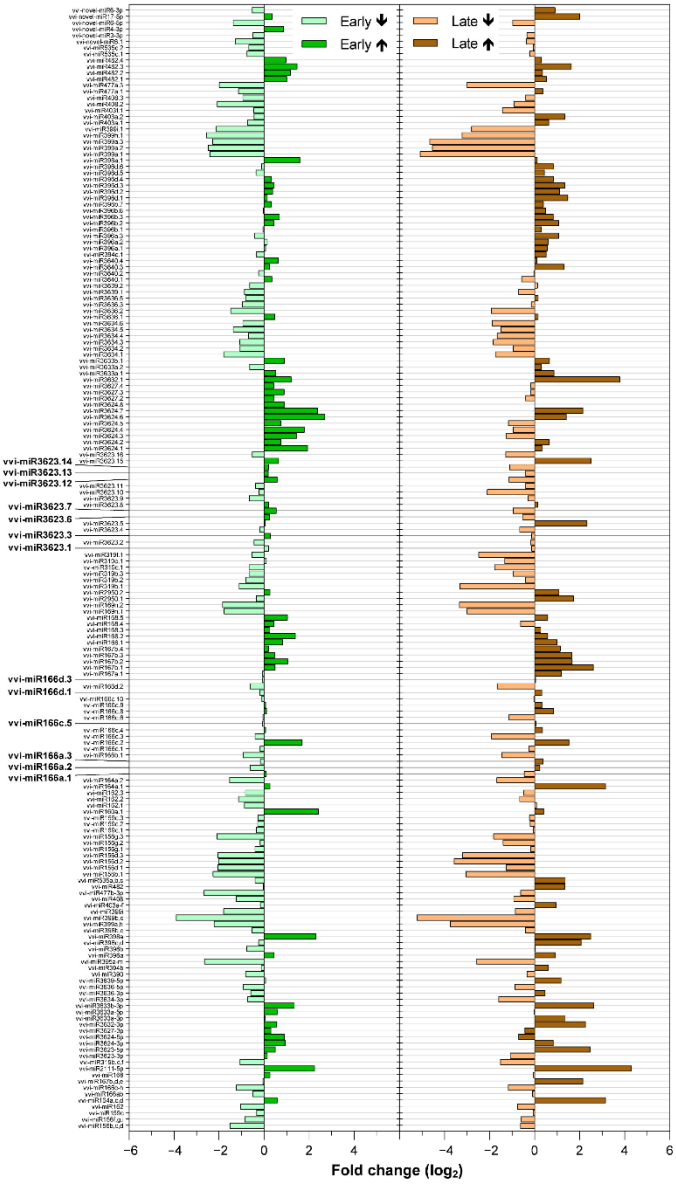


Figure 2. All of the differentially expressed miRNAs in grapevines infected with '*Ca*. P. solani' in comparison with uninfected grapevine during the growing seasons. Exposed are isomiRs of miRNAs, which regulate genes coding for proteins involved in disease resistance with changing mode of up- or down-regulation during the year. ↑, up-regulated; ↓, down-regulated.

2.2. Gene Set Enrichment Analysis Confirms High Transcriptional Activity in the Early Growing Season

Gene set enrichment analysis is a computational method that determines whether an a priori defined set of genes shows statistically significant concordant differences between two biological states. Here it was used to identify any relationships between the expression and function of the differentially expressed genes in these grapevines infected with '*Ca.* P. solani' (Figure 3). This gene set enrichment analysis showed a very strong response of grapevine to infection with '*Ca.* P. solani' early in the growing season, before symptoms develop. Eight functional bins [46] were differentially up-regulated only in the early growing season, specifically as: major CHO metabolism.degradation; glycolysis; fermentation; TCA/organic transformation; mitochondrial electron transport/ATP synthesis; DNA synthesis/chromatin structure; signaling receptor kinases, and cell vesicle transport. Twelve bins were enriched early and late in the growing season, and six bins only late in the growing season (Figure 3).

The largest enriched bins (in terms of percentages of genes) were associated with protein amino acid activation, synthesis of ribosomal proteins, synthesis initiation, elongation, targeting, degradation, folding, glycosylation, and assembly and cofactor ligation. Early in the growing season, the majority of the genes in these bins were up-regulated. Late in the growing season, the majority of the genes in these bins were down-regulated.

The bins related to RNA and transcription factors (i.e., ARF, bHLH, C2C2(Zn) constans-like zinc finger families) were down-regulated for both growing seasons. On the other hand, several genes in the bins associated with the cell wall and secondary metabolism were up-regulated throughout growing seasons.

2.3. Genes and miRNAs Not Associated with Phytoplasma Diseases Contribute to Resolving the Sanitary Status of Grapevines

Based on their mRNA profiles, uninfected and infected grapevine samples were clearly separated according to the sanitary status, as well as by the time of sampling (Figure 4a). In addition to the genes that encode a thaumatin protein from pathogenesis-related protein from class 5 and pathogenesis-related protein from class 10 (which have been associated with phytoplasma pathogenesis [3,27]), the other main genes with seasonal variations in these grapevine samples were ones that still have unclear roles during grapevine infection with '*Ca.* P. solani'; namely, genes that encode metallothionein, cysteine proteinase1, cellulose synthase-like G3, chloroplast β-amylase, and a gene related to cold, circadian rhythm, and RNA binding2 (Table 1).

The genes that best separated the grapevine samples according to their sanitary status encode metallothionein, which is different from the contributor to the seasonal origin, and the same genes that contribute to separation by season and encode chloroplast β-amylase (i.e., *Vitvi02g00605*) and a protein from the ubiquitin-protein family. Additional contributors that separated the infected and uninfected grapevines are involved in photosynthesis, as the main known process to be down-regulated in phytoplasma-infected grapevines; i.e., the gene that encodes ribulose bisphosphate carboxylase (small chain) family protein, and two genes that encode rubisco activase.

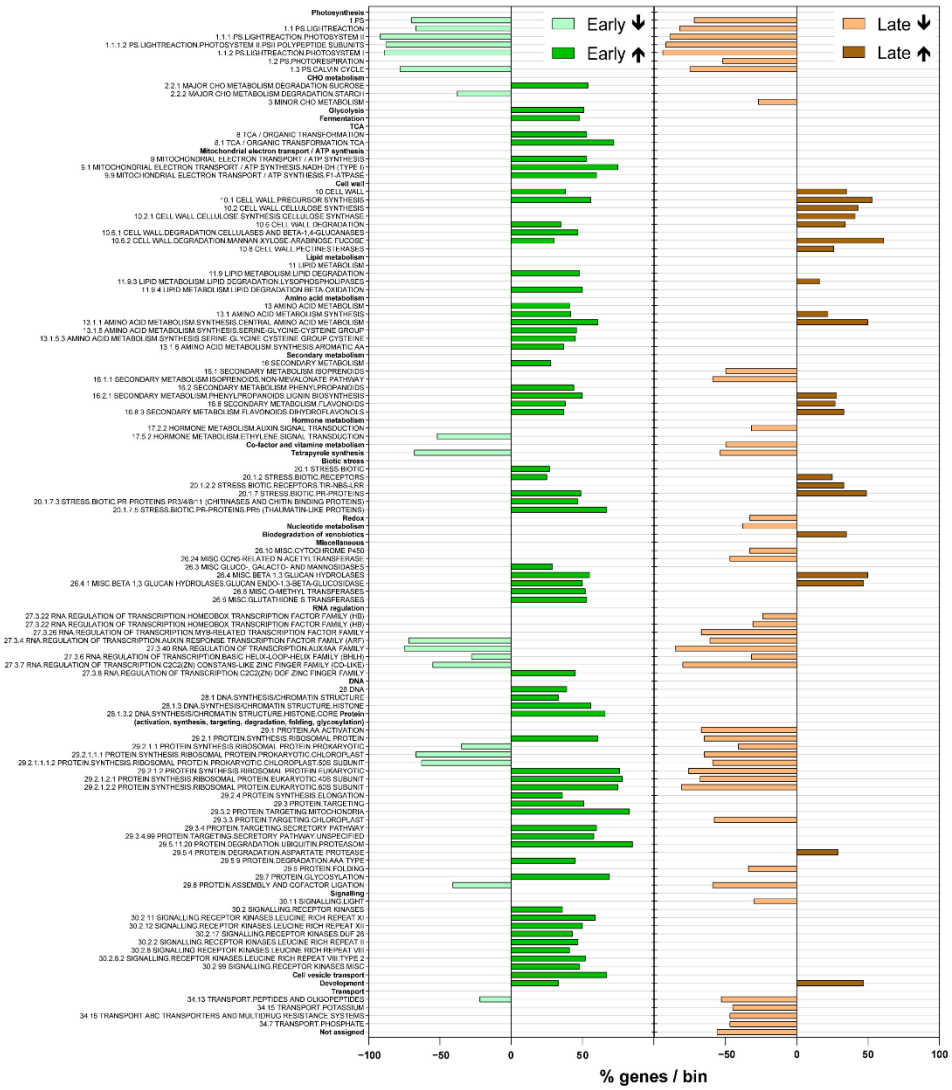


Figure 3. Enriched bins [46] according to gene set enrichment analysis showing proportions (%) of genes up-regulated (↑) or down-regulated (↓) in the particular bins in the grapevines infected with *'Ca.* P. solani' compared to the uninfected grapevines.

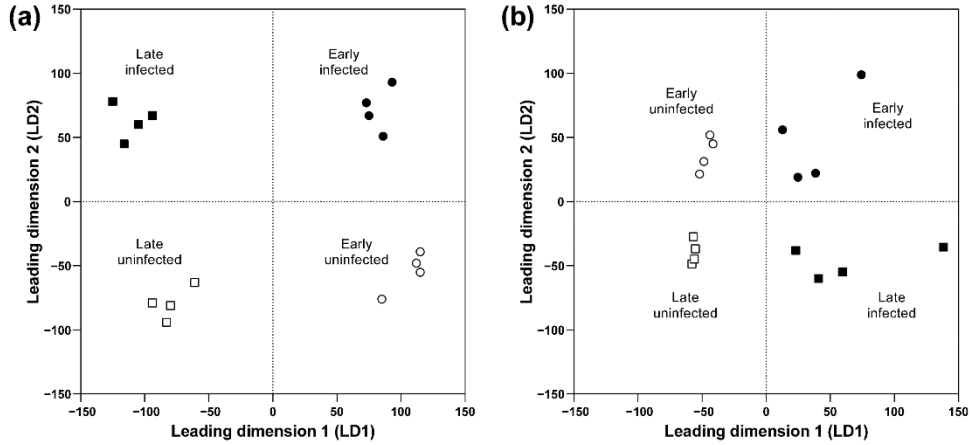


Figure 4. Multidimensional scaling for the normalized gene expression (**a**) and sRNA expression (**b**) in grapevine during infection with '*Ca.* P. solani'.

Table 1. The 10 genes with the most important contributions to separation of the grapevine samples according to the early and late growing seasons (Early/Late), and the same for their sanitary status (Uninfected/Infected), for the grapevines during infection with '*Ca.* P. solani'. E, early; L, late; U, uninfected; I, infected.

Gene ID	Description	Log_2 FC			
		E-I : E-U	L-I : L-U	L-U : E-U	L-I : E-I
Early/Late					
Vitvi02g01406	Thaumatin family	6.62	3.27	5.52	2.18
Vitvi06g01696	Metallothionein	1.86	2.88	1.14	2.16
Vitvi18g00740	Granulin repeat cysteine protease family protein	0.93	1.19	1.72	1.98
Vitvi19g00434	Ubiquitin family protein	0.93	0.31	2.10	1.48
Vitvi05g01756	Pathogenesis-related protein 10	1.50	−0.27	1.91	0.13
Vitvi02g00605	Chloroplast β-amylase	0.56	−0.40	5.71	4.75
Vitvi06g01697	Metallothionein	0.55	1.02	−0.43	0.05
Vitvi02g01341	Cellulose synthase-like G3	0.73	1.13	4.34	4.75
Vitvi03g00327	Cold circadian rhythm and RNA binding 2	0.01	0.28	1.81	2.07
Vitvi07g01690	Cysteine proteinase1	1.04	0.96	2.05	1.98
Uninfected/Infected					
Vitvi19g01871	Metallothionein 3	−1.68	−0.47	0.26	1.48
Vitvi02g00605	Chloroplast β-amylase	0.56	−0.40	5.71	4.75
Vitvi08g01245	Rubisco activase	−1.14	−1.80	2.25	1.59
Vitvi01g00714	Galactinol synthase 4	−0.55	0.07	0.87	1.49
Vitvi19g00549	GDP-L-galactose phosphorylase vitamin C defective 5	−1.13	−1.51	0.54	0.16
Vitvi17g00038	CLPC homolog 1	−0.24	−1.05	2.01	1.19
Vitvi05g00563	Early light-induced protein 1, chloroplastic-related	0.24	−0.73	2.00	1.03
Vitvi17g00320	Ribulose bisphosphate carboxylase (small chain) family protein	−1.14	−1.67	−0.28	−0.81
Vitvi19g00434	Ubiquitin family protein	0.93	0.31	2.10	1.48
Vitvi06g00513	Rubisco activase	−1.48	−1.84	−0.35	−0.71

It is worth noting that the Arabidopsis ortholog of chloroplast β-amylase gene *Vitvi02g00605*, *BAM3* (*At4g17090*) encodes one of the six plastidic β-amylases in Arabidopsis, which is transcriptionally induced by cold stress and is dominantly active in mesophyll cells during the night. This might explain the contribution of a gene related to

the circadian rhythm to the separation by seasonal origin. The importance of β-amylase to multidimensional scaling might at least partially explain the synthesis and accumulation of starch during phytoplasma infections [9,14,19]. Although one of the prominent symptoms of phytoplasma infections is an accumulation of starch in leaves [9,11,12], the source of this starch is not clear. ADP-glucose-pyrophosphorylase (AGPase) is a rate-limiting enzyme in starch biosynthesis [47]. Hitherto, transcript analysis of the gene that encodes its large regulatory subunit in grapevine cv. Chardonnay when infected with '*Ca*. P. solani' revealed its transcriptional up-regulation [20], but no significant difference on AGPase enzyme activity [19]. Our RNA-Seq of grapevine cv. Zweigelt here showed several genes that encode AGPase, with differential expression in the uninfected and infected grapevines for both growing seasons (Table S1). However, non-significant differences were found in total AGPase enzyme activity (data not shown) in agreement with our previous study [19]. On the other hand, a transcript that encodes chloroplastic β-amylase was up-regulated in infected grapevines in the early growing season and less so in the late growing season (Table 1). This is in agreement with a study on phytoplasma-infected symptomatic mulberry leaves in which the transcript levels of the β-amylase gene were lower compared to uninfected leaves together with a significant reduction in the corresponding β-amylase enzyme activity [12]. These results suggested that the accumulation of starch in the infected leaves results from reduced starch degradation and not from its de-novo synthesis [12]. However, data for β-amylase in mulberry leaves are in agreement with the reduced gene expression and enzyme activity of α-amylase [12]. This was not the case in grapevines of cv. Chardonnay infected with '*Ca*. P. solani', where the infection resulted in increased expression of an α-amylase gene [8]. In grapevine cv. Zweigelt (this study), infection with this pathogen was associated with down-regulation of *Vitvi03g00500* and *Vitvi01g00932*, which are genes that encode α-amylase-like and α-amylase protein, respectively. An additional two genes that encode α-amylase-like proteins, *Vitvi03g01571* and *Vitvi18g00144*, were up-regulated. Whether these data and the previous data on amylases in infected plants are species-related or related to ecological factors is currently not known.

miRNA expression data were also plotted with multidimensional scaling, which clearly revealed the sanitary status of the grapevine samples as well as their early or late growing season (Figure 4b). Interestingly, among sRNAs that differentiated the grapevine samples by sanitary status, there were miRNAs (*vvi-miR482.4*, *vvi-miR166d.2*, *vvi-miR482*, and *vvi-miR156g.1*) that regulated the expression of genes involved in biotic stress, several disease resistance proteins, a heat shock protein, and β-galactosidase (Table 2; Tables S2 and S5). Four sRNAs (*vvi-miR166c-h*, *vvi-miR3623.5*, *vvi-miR3623-5p*, and *vvi-miR3624-3p*) contributed significantly to both sanitary status and season differentiation. These mainly regulate the expression of genes involved in RNA regulation of transcription, protein degradation, and metal transport (Table 2; Tables S2 and S5). The contributors to sanitary status separation also included *vvi-miR162*, *vvi-miR162.3*, *vvi-miR159c*, *vvi-miR159c.1*, and *vvi-miR3623.4*, which regulate the expression of genes involved in nucleotide metabolism, RNA processing and degradation, RNA regulation of transcription, posttranslational protein modification, and protein degradation (Table 2; Tables S2 and S5).

Table 2. The 10 miRNAs with the most important contributions to separation of the grapevine samples according to the early and late growing seasons (Early/Late), and the same for their health status (Uninfected/Infected), for the grapevines during infection with '*Ca.* P. solani'. E, early; L, late; U, uninfected; I, infected.

miRNA ID	Log_2 FC			
	E-I : E-U	L-I : L-U	L-U : E-U	L-I : E-I
Early/Late				
vvi-miR166c-h	−1.24	−1.18	0.37	0.43
vvi-miR162	−1.05	−0.77	−0.05	0.24
vvi-miR3623.5	0.05	2.33	0.65	2.92
vvi-miR3624-3p	0.94	0.83	1.08	0.96
vvi-miR3623.4	−0.20	−0.67	0.83	0.35
vvi-miR159c	−0.34	−0.06	−0.26	0.02
vvi-miR162.3	−0.84	−0.50	−0.01	0.33
vvi-miR3623-5p	0.50	2.47	0.50	2.47
vvi-miR159c.1	−0.36	−0.07	−0.16	0.12
vvi-miR3634.3	−1.09	−1.84	−0.14	−0.89
Uninfected/Infected				
vvi-miR3624-3p	0.94	0.83	1.08	0.96
vvi-miR3623.5	0.05	2.33	0.65	2.92
vvi-miR3623-5p	0.50	2.47	0.50	2.47
vvi-miR166c-h	−1.24	−1.18	0.37	0.43
vvi-miR156g.1	−0.42	−0.18	1.74	1.97
vvi-miR482	−0.05	1.33	−0.05	1.34
vvi-miR398b,c	−0.54	−0.43	1.68	1.79
vvi-miR166d.2	−0.63	−1.67	2.08	1.04
vvi-miR482.4	0.98	0.31	1.00	0.32
vvi-miR168.5	1.05	0.58	1.01	0.55

2.4. The Most Differentially Expressed Genes and sRNAs Are Associated with Different Aspects of Biotic Stress Signaling

Analysis of the mRNA-Seq and sRNA-Seq data from grapevine cv. Zweigelt during infection with '*Ca.* P. solani' using the MapMan tool [46] revealed several novel patterns of gene expression related to the bins that were putatively related to biotic stress. These included: cell wall; hormones; proteolysis; NBS-LRR receptors; pathogenesis-related proteins; signaling, including sugar and nutrient physiology, receptor kinases, calcium signaling, G-proteins, MAP kinases, and light; transcription factors; and secondary metabolites.

In the bins that included the pathogenesis-related proteins and secondary metabolites, the analysis of the corresponding gene expression in pre-symptomatic but infected plants revealed several genes that were highly up-regulated compared to uninfected grapevines, although their expression decreased late in the growing season (Table S1). Among these, there were genes that encode for pathogenesis-related proteins from class 1 and 5, chitinases, and *Vitvi16g01336*, which encodes 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein downy mildew resistance 6 (DMR6) ([43,48]. DMR6 has an essential role in the mediation of salicylic acid homeostasis during plant development, leaf senescence, and pathogen responses and acts as a susceptibility S gene in a class of suppressors of plant immunity [27].

In agreement with the reported roles of sRNAs in stresses, the greatest number of predicted targets of differentially expressed miRNAs and phasiRNAs corresponded to the bins associated with the biotic stress signaling (Figures 5 and 6). In this group of sRNAs, 65% of the miRNAs were down-regulated in comparison with the up-regulated miRNAs in both of the growing seasons.

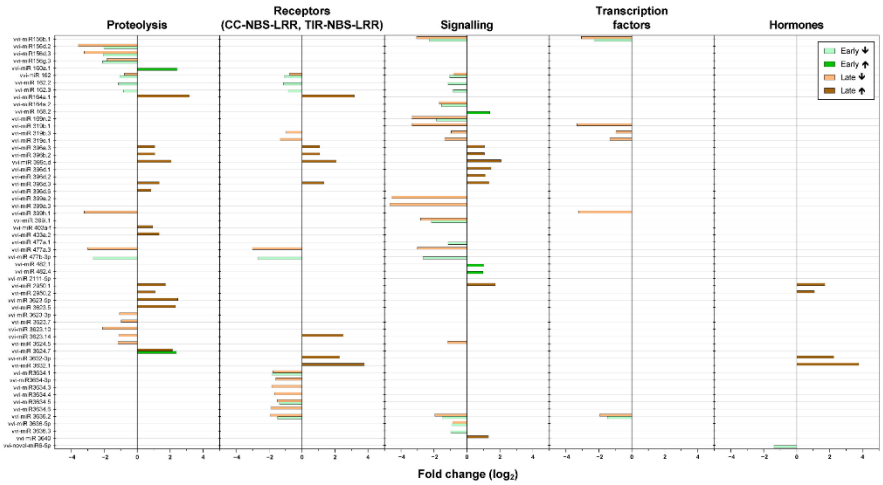


Figure 5. Differential expression of miRNAs associated with biotic stress in the grapevines infected with '*Ca.* P. solani' compared to the uninfected grapevines over the growing seasons. ↑, up-regulated; ↓, down-regulated.

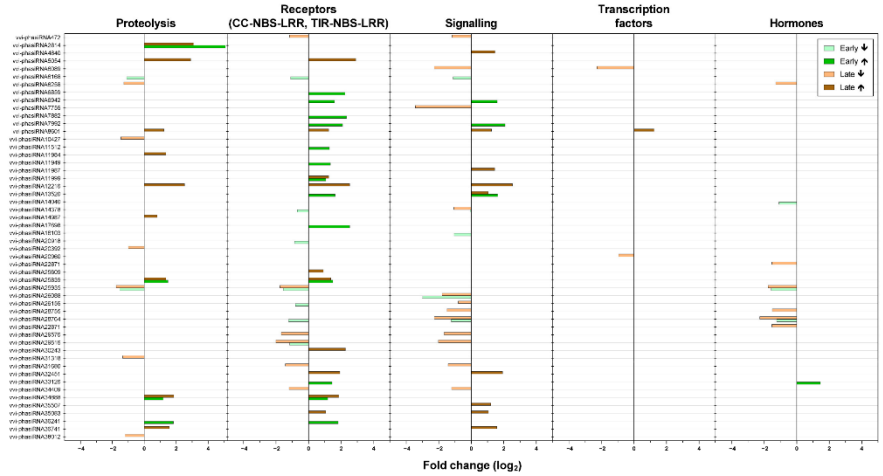


Figure 6. Differential expression of the phasiRNAs associated with biotic stress in the grapevines infected with '*Ca.* P. solani' compared to uninfected grapevines over the growing seasons. ↑, up-regulated; ↓, down-regulated.

2.4.1. Important Involvement of Genes Associated with the Cell Wall in Bois Noir Pathogenesis

The *cell wall* bin is putatively associated with biotic stress and comprised 262 differentially expressed genes. Early in the growing season, the second most up-regulated gene with an almost 8-fold increase was *Vitvi02g00653*, which encodes expansin-like B1 (Table S1) and is involved in cell wall processes [46]. On the other hand, a gene with the ID *Vitvi09g00767* that encodes expansin B2 was down-regulated (Table S1) and targeted by

vvi-miR166d.2 (Table S5). This miRNA was among those, which differentiated the samples by their sanitary status (Table 2). In the same *cell wall* bin was also the most down-regulated gene of all detected differentially expressed genes with a 9-fold decrease. This was the gene with the ID *Vitvi13g00172*, which encodes expansin A8 and was expressed in the late growing season (Table S1). Although the functional roles of the expansin-like A and B family members remain unclear, a recent study of the expansin-like B1 ortholog in *Brassica rapa* [49] suggested their involvement in stress. In addition, the second most down-regulated gene ($\log_2$ FC = −6.42) in the early growing season was from the same bin and encodes cellulose synthase-like G3 (*Vitvi02g00165*) (Table S1).

2.4.2. Hormonal Balance Is Disturbed Already in Pre-Symptomatic Phase of Phytoplasma Infection

Infection of grapevine cv. Zweigelt with '*Ca.* P. solani' confirmed the importance of hormonal regulation in phytoplasma diseases [27].

Salicylic acid studies have revealed that several salicylate biosynthetic, signaling, or marker genes are up-regulated in leaf-vein-enriched samples and whole leaves of grapevines infected with phytoplasma '*Ca.* P. solani' [27]. Infection with this phytoplasma for grapevine cv. Zweigelt here showed high induction of three transcripts of several genes that encode for S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase, which catalyzes the formation of the volatile ester methyl salicylate from salicylic acid (*Vitvi04g02117*, *Vitvi04g02118*, *Vitvi04g02122*); these showed a peak in the early growing season (Table S1). In agreement with the results from a study of grapevine cv. Chardonnay infected with '*Ca.* P. solani' [50], the relative expression of non-expressor of pathogenesis-related protein 1 (*NPR1*), the receptor for salicylic acid, did not differ across the infected grapevines of this cv. Zweigelt (Table S1). Among the-several genes that encode pathogenesis-related proteins, PR-1, PR-2, and PR-5 are induced by salicylic acid and are commonly used as molecular markers for the salicylic-acid-dependent systemic acquired resistance signaling; they have also been shown to be induced after phytoplasma infection [27]. Out of 14 genes that encode PR-1, the transcription of 12 was induced here in grapevines cv. Zweigelt infected with '*Ca.* P. solani' (Table S1). The same up-regulation pattern was shown for 19 genes that encode PR-5 and also for one gene that encodes PR-2 (Table S1). The *DMR6* gene [27] was also highly induced in the infected grapevine cv. Zweigelt. Similarly, as shown in grapevines infected with Flavescence dorée phytoplasma [14], *DMR6* expression was higher for the early growing season compared to the late season (Table S1).

Several genes that encode the major enzymes involved in jasmonate biosynthesis and modification were differentially expressed (Table S1). Their involvement in phytoplasma infections has been shown for grapevines infected with '*Ca.* P. solani' before, although their role in pathogenicity is not clear yet [27]. In general, genes that encode lipoxygenases are up-regulated upon infection with '*Ca.* P. solani', and they are suppressed upon infection with Flavescence dorée phytoplasma [43]. Lipoxygenase genes have also been shown as down-regulated in Arabidopsis infected with '*Ca.* P. asteri' strain witches' broom as a result of secreted effector protein11 ($SAP11_{AYWB}$), which destabilizes transcription factor promoting the expression of these genes [51]. A homolog of $SAP11_{AYWB}$ has also been found in the genome of '*Ca.* P. solani' strain SA-1 originally infecting grapevine [52]. However, our RNA-Seq here revealed three genes that encode lipoxygenases (*Vitvi14g00234*, *Vitvi14g02539*, and *Vitvi06g00149*) that were either unaffected or were down-regulated. On the other hand, the transcripts of the genes *Vitvi06g00155* and *Vitvi06g00158* increased from the beginning to the end of the growing seasons, with the same shown for genes that encode the PLAT/LH2 domain of plant lipoxygenase-related proteins (*Vitvi01g01562*, *Vitvi05g00472*, *Vitvi09g00096*, *Vitvi01g01562*, and *Vitvi05g00472*). Similar to a previous study [50], the allene oxide synthase gene (*Vitvi03g00395*) was up-regulated at both sampling times, although a transcript of *Vitvi18g00886* was up-regulated only later in the growing season. In contradiction with the previous reports [50], gene expression of jasmonic acid carboxyl methyltransferases (*Vitvi18g02762*, *Vitvi18g02763*, and

Vitvi18g02761) was suppressed in infected grapevines. The opposite grapevine response when infected with '*Ca.* P. solani' or Flavescence dorée phytoplasma was additionally supported by induction of gene expression of several 12-oxophytodienoate reductase genes (*Vitvi18g02485*, *Vitvi18g02138*, *Vitvi18g02139*, *Vitvi18g03161*, and *Vitvi18g03162*) early in the growing season in grapevine cv. Zweigelt infected with '*Ca.* P. solani'. Gene expression of jasmonic acid ZIM (zinc-finger inflorescence meristem) domain-containing protein, which was shown to be down-regulated upon infection with Flavescence dorée phytoplasma, in grapevine cv. Zweigelt depended here on the specific gene: while *Vitvi09g00064* and *Vitvi01g02293* were up-regulated in both growing seasons, *Vitvi01g00473* and *Vitvi17g00189* were down-regulated, and *Vitvi10g01879* was down-regulated in the early growing season and up-regulated in the late season. The PR3/4 marker genes for jasmonic acid metabolism were up-regulated only in grapevines that recovered from the infection of grapevine cv. Chardonnay with '*Ca.* P. solani' [50]. On the other hand, in grapevine cv. Zweigelt, four of the genes that encode PR3/4 (*Vitvi04g01049*, *Vitvi05g00094*, *Vitvi05g01366*, and *Vitvi05g01575*) were not affected, while the transcripts of *Vitvi05g02250*, *Vitvi15g01035*, and *Vitvi15g01037* (which also encode PR3/4) increased, especially during the early growing season. The transcript levels of a third jasmonic acid metabolism marker PR6 (*Vitvi02g01273*, *Vitvi05g01910*, *Vitvi18g00852*, *Vitvi18g03048*, *Vitvi09g00071*, *Vitvi11g00061*, *Vitvi17g01613*, and *Vitvi17g01121*) were higher later in the growing season compared to the early season. The families of jasmonic acid metabolism genes revealed here included at least one gene for which expression has been shown previously, which indicates the importance of high-throughput analysis for interpretation of the not always straightforward functions in infected plants.

The role of auxins in phytoplasma infections has already been documented [27], and the present data supported this, with differential expression of more than 200 auxin-associated genes (Table S1). *Small auxin-up RNAs* (*SAURs*) comprise a large multigene family that is involved in primary auxin responses in plants. These influence nearly all aspects of plant growth and development through the regulation of cell division, expansion, differentiation, and patterning. However, the functions of the SAUR proteins have remained elusive, presumably due to extensive genetic redundancy [53]. The present analysis revealed four down-regulated SAUR genes (*Vitvi02g00507*, *Vitvi03g01350*, *Vitvi04g02073*, and *Vitvi09g00046*). This is in agreement with a study of a single virulence factor, tengu-su inducer (TENGU), that is associated with the phytoplasma-infected plant phenotype [54,55]. Among the genes that directly influence the homeostasis of auxins, there is the auxin-responsive GH3 gene family. The related *Vitvi03g00586*, *Vitvi03g00586*, and *Vitvi07g01644* genes were up-regulated in the infected grapevines of cv. Zweigelt. To date, their role in phytoplasma infections has not been evaluated.

Fifty-nine of the genes revealed as associated with ethylene metabolism were affected by infection of grapevine cv. Zweigelt (Table S1). Among these, four were up-regulated for a more than 3-fold difference in the early growing season: *Vitvi07g02070*, which encodes ERF098; Vitvi08g01502, which encodes a transmembrane protein; and *Vitvi09g00834* and *Vitvi09g00837*, which encode two integrase-type DNA-binding superfamily proteins. *Vitvi08g01502* was also induced later in the growing season, together with *Vitvi04g00533*, which is also an integrase-type DNA-binding superfamily protein. On the other hand, two genes from the ethylene signaling bin were down-regulated later in the growing season, with a $\log_2$ FC greater than 3: *Vitvi05g01924*, which encodes a 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; and *Vitvi04g01895*, which encodes a PPPDE (permuted papain fold peptidases of DsRNA viruses and eukaryotes) thiol peptidase family protein. The significance of the expression of these genes for phytoplasma pathogenicity has not been explored.

Although less studied, plant hormones such as abscisic acid, gibberellic acid, cytokinins, brassinosteroids, and peptide hormones have important roles in plant defense against invading organisms through their fine-tuning of the plant responses to phytoplasmas [27].

In the early growing season, more than 3-fold increases were seen for gibberellin 3-oxidase 1 (*Vitvi04g00435*) and a proline-rich protein (*Vitvi14g01819*). In the later growing season, an overall decrease in gibberellin oxidase gene expression was seen (*Vitvi10g00020*, *Vitvi19g00432*, *Vitvi15g00782*, *Vitvi09g00448*, and *Vitvi17g00601*).

Although 45 genes associated with abscisic acid metabolism were affected by '*Ca*. P. solani' infection (Table S1), their differential expression did not exceed four-fold change, with the exceptions of the genes *Vitvi03g01727* and *Vitvi12g00015*, which encode HVA22 protein; these were shown to be involved in abiotic stress in *Citrus* spp. [56].

For brassinosteroids, high down-regulation in the late growing season was observed for two members, both of which are involved in sterol synthesis (*Vitvi19g00007* and *Vitvi01g00319*).

We also detected changes in both miRNAs and phasiRNAs associated with predicted targets in hormone metabolism (Figures 5 and 6). Early in the growing season, one sRNA associated with genes related to jasmonic acid was up-regulated (*vvi-phasiRNA33126*), while *vvi-phasiRNA33126* was down-regulated. On the other hand, *vvi-phasiRNA25935* and *vvi-phasiRNA28764*, which are associated with auxins, and *vvi-novel-miR6-5p*, which is associated with abscisic acid, were down-regulated. Early in the growing season, there were no differentially expressed sRNAs associated with brassinosteroids or ethylene. While 71% of detected sRNAs that targeted genes in the bin of auxins were down-regulated late in the growing season (*vvi-phasiRNA25935*, *vvi-phasiRNA28755*, *vvi-phasiRNA6258*, and *vvi-phasiRNA22871*), two miRNAs were up-regulated (*vvi-miR2950.1* and *vvi-miR2950.2*). Later in the growing season, there was down-regulation of *vvi-miR3624.5* and *vvi-phasiRNA31318*, which are associated with brassinosteroids and ethylene, respectively. In addition, two miRNAs that are associated with jasmonic acid were up-regulated (*vvi-miR3632-3p* and *vvi-miR3632.1*), while *vvi-phasiRNA14040* was down-regulated.

2.5. Known Pathways with the Involvement of Novel Genes

As photosynthesis and secondary metabolism are tightly intertwined with symptom development in the later growing season [3,43], they have been most studied in plants infected with phytoplasmas. However, their role in pre-symptomatic plants has been less considered. Of note, phasiRNAs were up-regulated in the bins associated with photosynthesis (Table S3), namely vvi-phasiRNA2814 in PS.lightreaction.photosystemII.PSII polypeptide subunits in the early and late growing season, as well as vvi-pahsiRNA12216 in PS.calvincycle.rubiscointeracting later in the growing season. These findings appear to be related to significant decreases in gene transcripts from the same bins.

Glycolysis and oxidative stress processes were also affected in grapevine cv. Zweigelt infected with phytoplasmas, with up-regulation of several aldolase genes that accelerate the reversible conversion of fructose-1,6-bisphosphate to dihydroxyacetone-phosphate and glyceraldehyde-3-phosphate (*Vitvi19g01724*, *Vitvi01g00360*, and *Vitvi08g01506*) (Table 1), and this correlated with detection of the activity increase of their encoded enzymes (Figure 7). The transcript of *Vitvi13g00241* that encodes dehydroascorbate reductase showed a small increase in the late growing season in response to infection (Table S1), but its enzymatic activity was significantly higher for both the early and late sampling times (Figure 7), indicating a possible posttranslational regulation. Lastly here, the expression of seven ascorbate peroxidase genes (*Vitvi04g02166*, *Vitvi06g00358*, *Vitvi18g00256*, *Vitvi08g01143*, *Vitvi03g00137*, *Vitvi04g00484*, and *Vitvi18g00445*) (Table S1) were differentially regulated in response to infection, likely contributing to a statistically insignificant difference in total ascorbate peroxidase activity in grapevine cv. Zweigelt (Figure 7). A role for ascorbate peroxidase in grapevine phytoplasma infections has been shown before, including in grapevines recovered of disease [57–61].

Previous studies have shown that phytoplasma infections affect carbohydrate metabolism [14,19,43]. Expression profiling analysis of the infected samples in the present study revealed several genes that encode starch synthase (Table S1). Three of these (*Vitvi10g02394*, *Vitvi10g00094*, and *Vitvi14g01968*) were down-regulated early in the grow-

ing season with small increases in the transcript levels later; two others (*Vitvi10g00739* and *Vitvi11g00903*) were down-regulated at both sampling times. As we have shown previously [14,19], sucrose synthase gene expression was higher in infected grapevines (*Vitvi04g00831*, *Vitvi07g00353*, *Vitvi11g00030*, and *Vitvi17g01221*), which indicated the important role of this enzyme in phytoplasma pathogenicity.

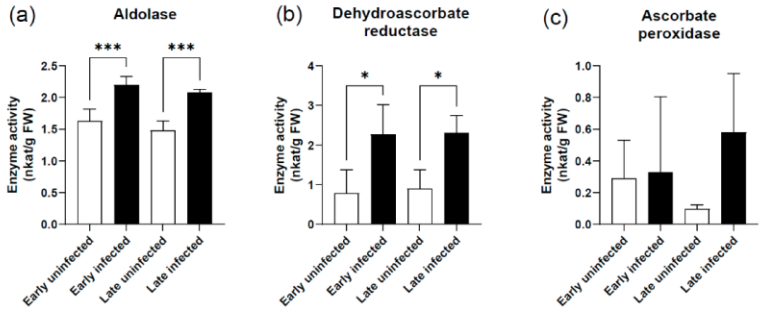


Figure 7. Enzyme activities of (**a**) aldolase, (**b**) dehydroascorbate, and (**c**) ascorbate peroxidase for the grapevine cv. Zweigelt. *, $p < 0.05$; ***, $p < 0.001$ (one-way ANOVA with Tukey's post-tests).

In agreement with the symptom development of phytoplasma-infected grapevines, the secondary metabolism changes are also pronounced. Previously observed changes in flavonoid synthesis genes [3,43] were also noticeable here for grapevine cv. Zweigelt. However, comparisons of the responses in the early and late growing seasons showed higher up-regulation in the early growing season for anthocyanins, chalcones, and dihydroflavonols, while isoprenoids were more commonly up-regulated in the late growing season (Table 1).

Abnormal lignification of canes is a prominent symptom of bois noir disease [62], and accordingly, several genes related to lignin biosynthesis are affected in these infected grapevines. Among these, there was up-regulation of phenylalanine ammonia-lyase in infected plants, which has also been shown in grapevine cv. Chardonnay infected with the same phytoplasma [26]. Our analysis additionally revealed several down-regulated laccase genes in the late growing season (*Vitvi18g01438*, *Vitvi18g02906*, *Vitvi04g01984*, *Vitvi08g01335*, *Vitvi08g01223*, and *Vitvi08g01031*). The biological functions of a diverse superfamily of multicopper oxidases and laccases include the lignification responsible for maintenance of the cell wall structure and of mechanical rigidity [63].

2.6. Temporal Network Modeling Reveals New Cross-Talk between Pathways Involved in Grapevines Infected with 'Ca. P. solani'

A drawback of most transcriptomic studies on phytoplasma-infected grapevines growing in their natural environment is that sampling occurs only at one time point in the growing season, which is most often when disease symptoms are most pronounced [3]. Only a few studies have been conducted in the early and late growing seasons, which have resulted in a short list of genes and their protein products; some of these have been explored in more detail [14,26,50]. Using a novel integrated analysis of transcriptomic data known as network enrichment methodology, which is based directly on RNA-Seq data [39], we explored the grouping of genes in communities that: (i) formed at the same times (i.e., early or late in the growing season) in the same groups of grapevines (i.e., uninfected and infected with '*Ca.* P. solani'); and (ii) showed significant disintegration (i.e., separation with a significant dissipation index) between the two growing seasons and between each group of grapevines (Figure 8, Table S6). By analyzing community behavior with time,

we discovered several new genes that are involved in phytoplasma pathogenesis and are connected in yet unexplored networks.

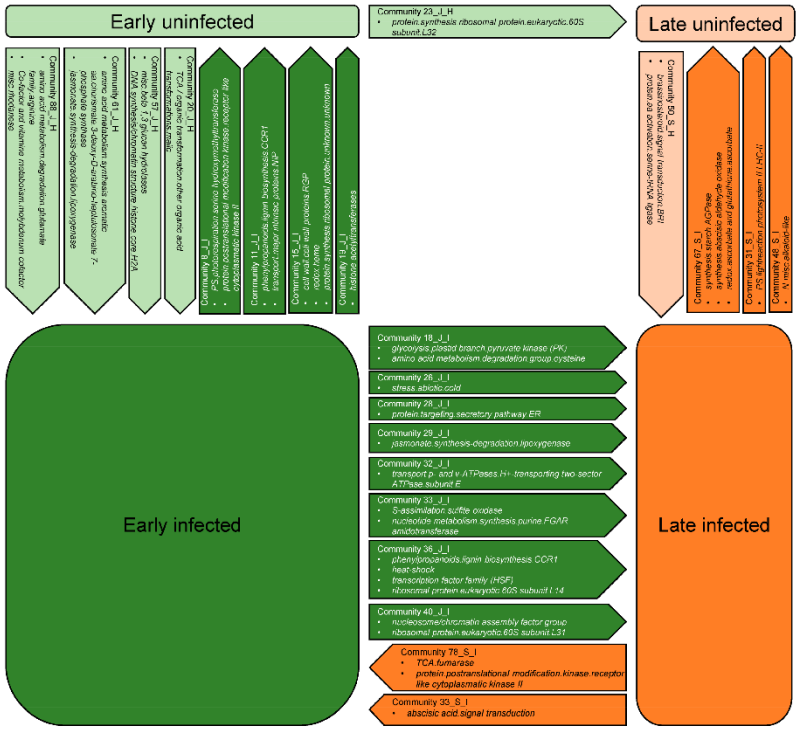


Figure 8. Schematic presentation of the communities with bins [46] that disintegrate between groups of uninfected (U) and grapevines infected (I) with '*Ca.* P. solani' in the early (E) and late (L) growing season.

2.6.1. Disintegrated Communities in Infected Grapevines

In the early growing season, in the group of uninfected grapevines, we detected four communities of genes that disintegrated with a high dissipation index in infected grapevines (Figure 8, Table S7). The first community consisted of genes functionally classified as β-1,3-glucan-hydrolases and H2A histone core proteins, which are involved in DNA synthesis. The latter included up-regulated plasmodesmata callose-binding protein 3 (*Vitvi01g00051*) in the infected grapevines in agreement with previous studies where callose was shown to be involved in phytoplasma pathogenicity [8,64,65].

The second community consisted of genes functionally classified as malic acid transformation enzymes, which included two induced genes in infected grapevines, the MATE (multidrug and toxic compound extrusion) efflux family protein (*Vitvi12g00101*), and the calcium-binding EF-hand family protein (*Vitvi14g00470*). Multidrug and toxic compound extrusion transporters are one of the largest secondary active transporter families in plants. These are involved in a wide variety of physiological functions throughout plant development for the transport of a broad range of substrates such as organic acids, plant hormones, and secondary metabolites [66,67]. They have been shown to be linked to disease resistance associated with salicylic acid [68], as well as to abscisic acid sensitivity and drought

tolerance [69]. Activation of salicylic acid metabolism in phytoplasma-infected grapevines in the early growing season has already been documented [14,50]. On the other hand, the calcium-binding EF-hand proteins have roles in the resistance mechanisms to various biotic and abiotic stresses [70,71].

Functional classification of the genes that comprise the third community consisted of arginine degradation, molybdenum involvement in vitamin metabolism, and rhodanese. Here, a molybdenum cofactor sulfurase (*Vitvi01g00087*) was significantly increased upon infection (Figure 9, Table S1). This enzyme is required for the enzymatic activity of the Mo enzymes (e.g., aldehyde oxidase) that are essential for the biosynthesis of the bioactive compounds abscisic acid and allantoin [72,73]. In Arabidopsis, this enzyme contributes to anthocyanin accumulation and oxidative stress tolerance in abscisic acid-dependent and -independent ways [74]. These processes have also been shown to be associated with phytoplasma infections previously [27,43,75].

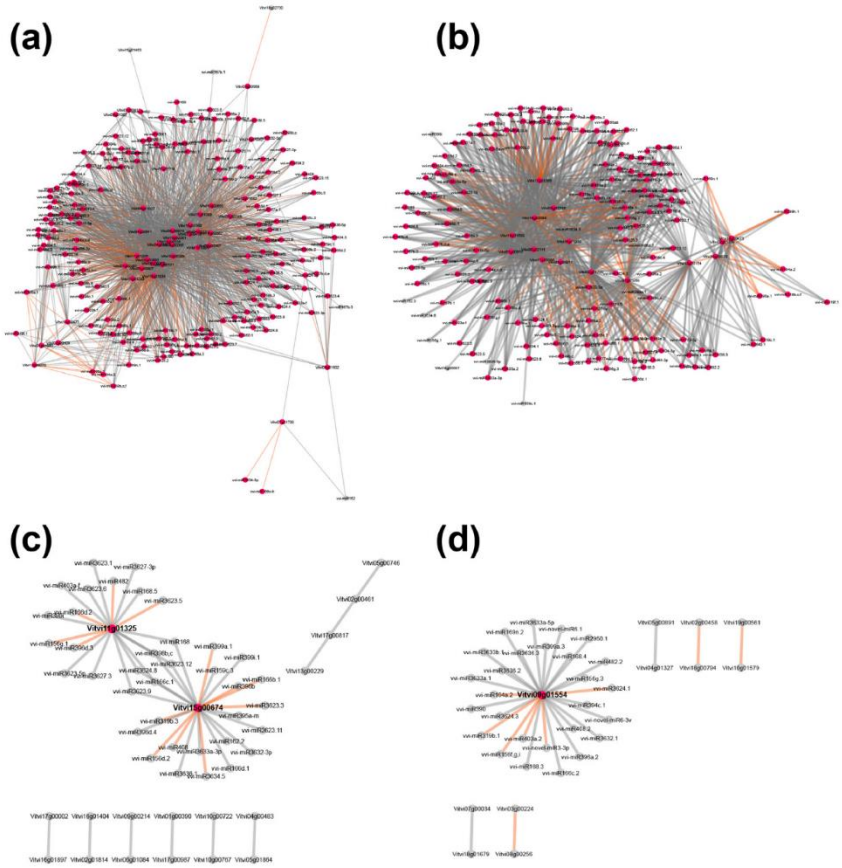


Figure 9. Visualization of networks that included mRNAs and miRNAs with Cytoscape: (**a**) uninfected, early growing season; (**b**) infected with '*Ca.* P. solani', early growing season; (**c**) uninfected, late growing season; (**d**) infected with '*Ca.* P. solani', late growing season.

The last community contained some genes that encode jasmonic acid lipoxygenases, which were discussed above in Section 2.4.2. (Table S1).

2.6.2. Early Growing Season Communities with a High Dissipation Index in Uninfected Grapevines

It is nevertheless worth comparing the communities that formed early in the growing season in infected grapevines, which disintegrated in uninfected grapevines with a very high dissipation index. With the applied methodology, we detected four such communities (Figure 8, Table S7). A high dissipation index of 0.62 was seen for a community that involved the genes from bins *photosynthesis.photorespiration.serine hydroxymethyltransferase* and *protein.postranslational modification.kinase.receptor-like cytoplasmatic kinase II*. Although photosynthesis is among the most studied processes during phytoplasma infection, in general, the genes revealed in this community have never before been described as responsive to phytoplasma infection. A possible reason for this oversight might be their low fold-changes between the uninfected and infected grapevines.

In the second community, there were two bins: *secondary metabolism.phenylpropanoids. lignin biosynthesis.CCR1* and *transport.major intrinsic proteins.NIP*. It is becoming increasingly evident that secondary metabolism is greatly affected during bois noir disease, which has been shown at the transcriptional, proteome, and metabolome levels [3]. A gene from this community (*Vitvi06g01762*) encodes a UDP-glycosyltransferase superfamily protein and its transcript increased in infected grapevines compared to uninfected grapevines (Table S6). It is known that glycosylation with UDP-glycosyltransferases is a major regulator of phenylpropanoid availability and biological activity in plants [76], and this protein family indicates stress-responsive regulation in Arabidopsis and *Brassica* species [77]. Although reports of cytokinin involvement in phytoplasma–plant interactions are scarce and mainly indirect, some symptoms of phytoplasma-infected plants hint at cytokinin participation (e.g., witches' broom, leaf yellowing, and fasciations) [27]. Thus, the observed induced transcript of CycD3 might be related to plants with phytoplasma infection phenotypes.

In the community with a dissipation index of 0.66 that was composed of the bins *cell wall.cell wall proteins.RGP*, *redox.heme*, and *protein.synthesis.ribosomal protein.unknown.unknown*, there was prominent differential expression of *Vitvi14g02435*, a gene that encodes a germin-like protein 10. This protein has been induced in soybean by methyl jasmonate, ethylene, and salicylic acid [78], which are all involved in defense responses and in phytoplasma infections [27]. In addition, overexpression of germin-like protein 10 in transgenic tobacco significantly enhanced tolerance to *Sclerotiorum* infection [78].

2.6.3. Early versus Late Growing Season Communities in Infected Grapevines

A comparison of communities that disintegrated from the early growing season toward the late growing season, and vice versa, showed eight communities with a high dissipation index in the former and only two in the latter group (Figure 8, Table S7). However, communities in the first group included genes with a wide array of functional classifications. This finding suggests a large transcriptional dynamic prior to symptom development, which previously used methods failed to capture.

2.6.4. Exploring mRNA-mRNA-miRNA Interaction Networks

We extended the idea of analyzing mRNA-mRNA expression similarity networks to also include miRNA molecules for each time and grapevine state point, which resulted in four individual networks. During community detection, the majority of the miRNA molecules were pruned due to low correlation values, which resulted from large expression differences between mRNA and miRNA. To remedy this situation, a novel type of connection between miRNA and mRNA molecules was introduced, based on the shortest path searches.

Following this approach, the complex network of connections both in uninfected and infected grapevines early in the growing season were defined (Figure 9). In particular, we detected only a few connections late in the growing season in uninfected grapevines, and

even fewer in the infected grapevines, such as the gene with the ID *Vitvi15g00674* that encodes SPX (SYG1/Pho81/XPR1) domain protein 3, which has been shown to have a role in signaling and homeostasis of cellular phosphate [79]. In addition, the transcript of this gene was induced in transgenic plants that expressed the effector $SAP11_{AYWB}$ [80]. It was shown in that study that SPX domain protein 3 is suppressed when $SAP11_{AYWB}$ is expressed under a *phosphate starvation response1* (*phr1*) mutant background, which suggested that *PHR1* is required for $SAP11_{AYWB}$-triggered cellular phosphate starvation responses. Of note, *PHR1* encodes the MYB (myeloblastosis) transcription factor that has a key role in responding to cellular phosphate deficiency in Arabidopsis [81]. However, the transcript of SPX domain protein 3 was down-regulated upon '*Ca.* P. solani' infection of grapevine cv. Zweigelt, although the transcripts of *PHR1* (*Vitvi14g00736* and *Vitvi07g00666*) were slightly increased. As has been shown before, there is no universal response to phytoplasma infections [3,27,43].

In infected grapevines later in the growing season, several miRNAs were connected to only one gene, namely *Vitvi09g01554* (Figure 9, Table S1). This gene has not been assigned to any bins and is annotated as PTHR36328:SF1, a cryptdin protein-like protein. Its mammal ortholog is an α-defensin known as cryptdin, which is a major microbicidal constituent of Paneth cell granules in mouse intestinal epithelial cells. Cryptdin is actively involved in innate enteric immunity and maintains intestinal homeostasis through the control of the intestinal microbiota [82–84]. Although its role in phytoplasma pathogenicity is entirely unknown, it might be a suitable candidate for exploring its presumed similar role in plant defense response against phytoplasmas.

3. Conclusions

This is the first comprehensive high-throughput RNA-Seq and sRNA-Seq study of grapevine infected with '*Ca.* P. solani' for two consecutive growing seasons, and it has resulted in several striking findings. We show a very dynamic transcriptional activity in infected grapevines in the early growing season (i.e., prior to symptom development). Grapevine plants reacted to the infection with two distinct sets of genes that responded at each growing phase. This reprogramming was not visible for the sRNA levels during the early time point of infection. However, the number of differentially expressed sRNA significantly increased with the annual development of bois noir disease, the majority of which were associated with biotic stress processes. This study has also revealed the importance of less-studied aspects of hormonal involvement in phytoplasma pathogenicity. The interaction network methodology introduced here enabled us to explore the temporal interaction network dynamics, and hence to discover new mRNAs and sRNAs, which might be crucial for an understanding of bois noir development over the growing seasons and might open new routes for research into phytoplasma diseases.

4. Materials and Methods

4.1. Plant Material

Four phytoplasma-infected grapevines (*Vitis vinifera*) cv. Zweigelt (syn. Rotburger,) and their four healthy neighbors from a vineyard in Klosterneuburg (Austria) were selected for the study, according to the development of their symptoms in the previous season (2016). The grapevines were planted in 2006, and the rootstock is Kober 5 BB. The first fully developed leaves were sampled from several canes of each vine in June and September 2017 and then in June 2020. The leaf veins were cut out of the leaves sampled, flash-frozen in liquid nitrogen, ground to a fine powder, and stored at −80 °C. In 2017, all four of the selected phytoplasma-infected grapevines showed severe symptoms in September, while three of these four phytoplasma-infected grapevines recovered in 2020.

4.2. Detection of Phytoplasma

In September 2017, all four phytoplasma-infected grapevines showed typical symptoms of infection, including reddening and curling of leaves and no or incomplete fruit

development. The presence of '*Ca*. P. solani' was confirmed by PCR with *SecY*, *Stamp*, *tuf*, and *vmp1* specific primers as described previously (Aryan et al., 2014). Based on these four markers, the previously described nettle type CPsM4_At1 was detected for all four of the infected grapevines. For the four control grapevines, none of the markers provided any PCR signal. All of the grapevines were free of symptoms of powdery and downy mildew. ESCA symptoms had not been recorded for any of the tested grapevines since 2006. The same grapevines were additionally sampled in 2020 and were again tested for the phytoplasmas in the late growing season by PCR and by visual characterization. However, the presence of phytoplasmas could not be detected in three out of the four infected grapevines, which clearly showed remission of bois noir disease.

4.3. RNA Extraction and Sequencing

Total RNAs, including small RNAs, were extracted from vein-enriched leaf samples from each plant separately, using an optimized cetyltrimethylammonium bromide (CTAB) method (adapted from [85]) combined with RNA purification on columns (Zymo-Spin; Direct-zol RNA MiniPrep Plus kits, Zymo Research, Irvine, CA, USA). About 50 mg frozen and powdered tissue was further homogenized with steel beads for 10 min at maximum speed in 800 µL CTAB buffer (100 mM Tris-HCl, pH = 8, 2 M NaCl, 25 mM EDTA, 2.0% (*w/v*) CTAB, 2.5% (*w/v*) PVP40, 2.0% (*v/v*) β-mercaptoethanol) using TissueLyser (Qiagen, Hilden, Germany). After the addition of an equal volume of chloroform-isoamyl alcohol (24:1), the sample was vortexed and centrifuged for 10 min at 10,000× *g* at 4 °C. The upper aqueous phase was recovered, to which 1.5 volume of pure ethanol was added. After a 30 min precipitation at 4 °C, the mixture was transferred into the columns (Zymo-Spin, Direct-zol RNA MiniPrep Plus kits, Zymo Research). The RNA was purified according to the manufacturer's instructions, with an additional washing step and a second prewashing step added to the beginning of the purification process. To elute the RNA, 30 µL preheated (80 °C) DNase/RNase-free water was added to the column and incubated for 10 min at room temperature. This was followed by 1 min centrifugation at 14,000× *g*. The isolated RNA was subjected to DNase digestion (DNase I Set; Zymo Research) and cleaned up using RNA Clean & Concentrator kits (Zymo Research). RNA concentration, integrity, and purity were assessed using a Bioanalyser (2100) and RNA 6000 Nano kits (Agilent Technologies, Santa Clara, CA, USA). Library preparations for the mRNAs and sRNAs, sequencing services (HiSeq 4000, Illumina, San Diego, CA, USA), and preprocessing to remove the adapter sequences and low-quality reads were provided by Novogene (Hong Kong). The raw data (in fastq format) have been deposited with the European Nucleotide Archive (ENA) under project accession number PRJEB42777.

4.4. mRNA Data Analysis

The 150 bp paired-end reads obtained were trimmed to remove low-quality bases (Phred < 20), clipped to remove the remaining adapter sequences, and mapped to the 12X.2 version of the PN40024 grapevine reference genome (https://urgi.versailles.inra.fr/files/Vini/Vitis%2012X.2%20annotations/, 29 February 2021), using CLC Genomics Workbench 12.0 (Qiagen), with the following parameters: mismatch cost, 2; insertion or deletion cost, 3; length fraction, 1.0; similarity fraction, 0.95; and maximum number of hits for a read, 1. The reads were annotated using the VCost.v2 annotation.

The differential expression analysis was performed in R v3.4.2 [86], using the limma package v3.34.9 [87]. mRNA counts with a baseline expression level of >50 reads mapped in at least 4 samples were TMM-normalized in edgeR v3.20.9 [88] and transformed using voom [89]. To identify differentially expressed mRNAs, the empirical Bayes approach was used, with Benjamini and Hochberg's (FDR) *p*-value adjustment. Genes with adjusted *p*-values <0.05 were considered statistically significantly differentially expressed.

4.5. sRNA Data Analysis

sRNA reads were filtered to exclude reads shorter than 18 nt and longer than 26 nt, as well as reads that were matching to rRNAs, tRNAs, snRNAs, and snoRNAs in the RNACentral database (https://rnacentral.org, 29 February 2021) [90], and using a CLC Genomics Workbench v8 (Qiagen). To identify known grapevine miRNAs, the remaining preprocessed sRNA reads were compared to the grapevine miRNAs registered in the miRBase database, release 22 (http://www.mirbase.org, 29 February 2021) [91], which did not allow mismatches. To identify novel unannotated miRNAs and their loci of origin (i.e., MIR loci), the reads were submitted to the two plant miRNA prediction tools ShortStack and miR-PREFeR [92,93]. Predictions were performed using the default parameters, except that no mismatches were allowed during mapping on reference 12X.2 version of the PN40024 grapevine reference genome (https://urgi.versailles.inra.fr/files/Vini/Vitis%2012X.2%20annotations/, 29 February 2021). MiRNAs were considered novel miRNAs only if they had >5 raw reads in at least two sRNA libraries and the miRNA sequence, and the corresponding miRNA* and MIR locus should have been predicted with both miRNA prediction tools. Reads that mapped to more than 30 locations in the grapevine genome were also discarded as being too repetitive to be miRNAs. As the miRNA prediction tools output also contained predictions of already annotated grapevine MIR loci, annotated grapevine miRNA precursors (pre-miRNA) from miRBase (v22) were mapped to the reference grapevine genome using bowtie2 [94]. Next, genome locations were extracted and compared with the predicted MIR loci using our internally developed script. If no overlap was detected, the predicted MIR loci were regarded as novel MIR loci. Novel grapevine miRNAs were further classified into known or novel miRNA families by clustering their predicted pre-miRNA sequences with sequences of known plant pre-miRNAs from miRBase using CD-HIT-EST, with an identity threshold of 0.8 [95]. The sequences with similarities with annotated pre-miRNAs were grouped into the corresponding known miRNA families, and sequences that did not show similarities with known plant miRNAs were classified as novel miRNA families.

Additionally, sequence miRNA variants (isomiRs) of known and novel miRNAs were identified using isomiRID [96]. Only sRNAs that matched perfectly to known or novel pre-miRNA sequences were considered (i.e., templated isomiRs). Prediction of PHA S (phasiRNA-producing) loci was performed using unitas [97]. The PHA S loci were detected by mapping preprocessed sRNA reads to the grapevine transcriptome sequences (Vitis 12X.2 annotations; https://urgi.versailles.inra.fr/files/Vini/Vitis%2012X.2%20annotations/, 29 February 2021) at 21-nt and 24-nt intervals and default settings.

The preprocessed reads from sRNA-Seq samples were mapped (with no mismatches allowed) to all known grapevine miRNAs, novel miRNAs, and isomiRs and counted using our internally developed script. Raw counts were exported and deposited with the European Nucleotide Archive under the project accession number PRJEB42777. Differential expression analysis was performed in R v3.2.2 [86] using the limma package v3.34.9 [87]. sRNA counts with a baseline expression level of >50 reads mapped in at least four samples were TMM-normalized in edgeR v3.20.9 [88] and transformed using voom [89]. To identify differentially expressed mRNAs, the empirical Bayes approach was used with the Benjamini and Hochberg (FDR) *p*-value adjustment. Genes with adjusted *p*-values <0.05 were considered statistically significantly differentially expressed.

4.6. sRNA Target Prediction

In-silico identification of grapevine transcripts targeted by sRNAs was carried out using the psRNATarget [98] and grapevine transcriptome sequences (Vitis 12X.2 annotations; https://urgi.versailles.inra.fr/files/Vini/Vitis%2012X.2%20annotations/, 29 February 2021), with the maximum expectation parameter set to 3 (Expectation), and otherwise using the default parameters.

4.7. Differential Expression Analysis and Visualization of mRNA and sRNA Expression

Samples were visualized for the level of their similarity using multidimensional scaling, as calculated from normalized mRNA and sRNA expression values separately. Principal component analysis in R v3.4.2 [86] was used to extract the genes that contributed most to the first two leading dimensions (LD1, LD2). To compare seasonal effects of infections, gene set enrichment analysis [99] was performed on normalized expression values using a grapevine MapMan mapping file [100]. The collapse/remap to gene symbols option was enabled (No_collapse), permutation was set to gene sets, and otherwise, the default settings were used. Functional categories with a false discovery rate corrected $q \leq 0.05$ were considered significant.

4.8. Network Community Analyses

The similarity between mRNA expression data was calculated as the reciprocal value of similarities between the mRNA expression data were calculated as the reciprocal value of the Euclidean distance for a given pair of expression vectors for each individual time, as the grapevine state points. The application of automated thresholding on these similarity networks allowed easier exploration of the space of possible networks. Scale-free networks ($2 > \alpha > 3$; α being the exponent of the fitted power-law function) were used for community detection with Infomap [101]. To generate mRNA–miRNA networks, connections were added by using shortest path searches of maximum length 3 [102]. Networks were visualized using Cytoscape [103].

4.9. Targeted Grapevine Gene Expression Analysis by qPCR

Differential expression of four genes was confirmed with qPCR (Table S7): for *DMR6* (*Vitvi16g01336*), *OLP* (*Vitvi02g01404*), *SAMT* (*Vitvi04g02122*), and *LOX* (*Vitvi06g00158*), with UBI_CF as the reference (*Vitvi19g00744*). These genes were chosen based on our previous results that showed their involvement in bois noir pathogenesis [8,20]. The complete list of primers and probes used is given in Table S7 (MIQE). Reverse transcription was performed with the High-Capacity RNA-to-cDNA kits (Applied Biosystems, Foster City, CA, USA). FastStart Universal Probe Master (Roche, Penzberg, Germany) was used for qPCR. The following thermal cycle conditions were applied: 95 °C for 10 min, 40 cycles of 95 °C for 15 s, and 60 °C for 1 min for PCR, and a climb in increments of 0.05 °C from 60 °C to 95 °C for the high-resolution melting curve. The Cq values were used for the relative calculation of initial target numbers from a serial dilution curve using quantGenius [104].

4.10. Enzymatic Activities

Extraction of enzymes from the grapevine material was performed as previously described by Jammer et al. (2015) [105] and adjusted according to Anžič (2019) [106]. The grapevine material was ground in liquid nitrogen using a mortar and pestle. About 0.5 g of material was collected into 2 mL microcentrifuge tubes (Eppendorf, Germany). Then 1 mL extraction buffer was added (0.5 M MOPS, 5 mM $MgCl_2$, 0.5 mg/mL BSA, 0.05% Triton X-100, 25 µM dithiothreitol, 1 mM benzamidine, 3% PEG-4000, 0.1 mM phenylmethylsulphonyl fluoride, 1% *polyvinylpyrrolidone*). The samples were then mixed on a rotary shaker for 40 min at 4 °C, and then they were centrifuged at 20,000× *g* for 10 min at 4 °C. The supernatant contained the cytoplasmic fraction of the enzymes, and it was dialyzed overnight against 20 mM KPO_4 buffer, pH 7.4. After dialysis, the protein-extract aliquots were pipetted into 96-well plates and stored at −20 °C. All of the enzymatic activity assays were performed in UV-transmissive, flat-bottomed, 96-well plates (UV-Star Greiner BioOne; Kremsmünster, Austria). Protein extract volumes from 1 µL to 20 µL were used for the reactions. The total reaction volume was 160 µL. Reaction mixes were incubated in a plate reader (Ascent Multiskan; Thermo Fisher Scientific, Waltham, MA, USA) for 40 min at 25 °C or 30 °C, according to the optimized protocol for each enzyme. All of the assays were carried out in triplicate, and for the control assays, the substrate was not added to the reaction mixes. Changes in absorbance per second were used to calculate

the activities of the enzymes, as nkat/g fresh weight (FW). These enzymatic activity assays were performed according to [107]. For measurement of ascorbate peroxidase activity, the samples were incubated with 0.025 mM ascorbate and 0.5 mM H_2O_2 in 50 mM K_2HPO_4/KH_2PO_4 buffer, pH 7.6. The H_2O_2 was omitted for the control reactions. The absorbance of H_2O_2 was measured at 290 nm. Glutathione S-transferase activity was measured in samples incubated with 1 mM 2,4-dinitrochlorobenzene and 1 mM reduced glutathione in 100 mM K_2HPO_4/KH_2PO_4 buffer, pH 7.4. 2,4-Dinitrochlorobenzene was omitted for the control reactions. Absorbance was measured at 334 nm, with the formation of (2,4-dinitrophenyl) glutathione. The enzymatic activities were further corrected by subtracting the non-enzymatic formation of (2,4-dinitrophenyl) glutathione by including in the 96-well plate a column without any extract added.

Supplementary Materials: Supplementary materials can be found at https://www.mdpi.com/article/10.3390/ijms22073531/s1. Table S1. High-throughput mRNA-Seq of grapevine cv. Zweigelt uninfected and infected with '*Candidatus* Phytoplasma solani' as vein-enriched leaf samples. BINCODE and NAME refer to [45]. For each mRNA sequence and sample, the differences in expression between the uninfected and phytoplasma-infected grapevines were calculated as $\log_2$ FC. Only mRNAs with a false discovery rate (FDR) adjusted *p*-value < 0.05 were considered as differentially expressed (Red, up-regulated; green, down-regulated). U, uninfected samples; I, samples infected with '*Ca.* P. solani'; R, recovered; E, early growing season; L, late growing season. Table S2. Known and novel grapevine miRNAs together with their expression levels in grapevine cv. Zweigelt uninfected and infected with '*Candidatus* Phytoplasma solani' as vein-enriched leaf samples. Details of all of the identified known and novel miRNA sequences are given. C, conserved in plants; G, grapevine specific; N, novel unannotated. For previously described miRNAs, the identifiers (ID) of the identical miRNAs in miRBase are also given. For each miRNA sequence and sample, the differences in expression between uninfected and phytoplasma-infected grapevines were calculated as $\log_2$ FC. Only miRNAs with a false discovery rate (FDR) adjusted *p*-value < 0.05 were considered as differentially expressed. Red, up-regulated; green, down-regulated; U, uninfected samples; I, samples infected with '*Ca.* P. solani'; E, early growing season; L, late growing season; NO, no miRNA homologs in other plant species. Table S3. List of phasiRNAs identified in grapevines, together with their expression levels in grapevine cv. Zweigelt uninfected and infected with '*Candidatus* Phytoplasma solani' as vein-enriched leaf samples. For each unique phasiRNA identifier (ID), the sequence and PHAS-producing locus are given. For each sample, the differences in expression between phytoplasma-uninfected and -infected grapevines were calculated as $\log_2$ FC. Only miRNAs with a false discovery rate (FDR) adjusted *p*-value <0.05 were considered as differentially expressed. Red, up-regulated; green, down-regulated; U, uninfected samples; I, samples infected with '*Ca.* P. solani'; E, early growing season; L, late growing season. Table S4. The novel miRNAs identified in grapevine cv. Zweigelt uninfected and infected with '*Candidatus* Phytoplasma solani' as vein-enriched leaf samples. Table S5. Genes that are regulated by the 10 sRNAs that contributed the largest amount of variance explained by the first leading dimensions (LD1) according to the health status of the samples, and by the 10 sRNAs that contributed the largest amount of variance explained by the second leading dimensions (LD2) according to the early or late growing season for the samples in multidimensional scaling for sRNA expression in grapevine cv. Zweigelt uninfected and infected with '*Candidatus* Phytoplasma solani' as vein-enriched leaf samples. Table S6. The grouping of genes in communities that: (i) form at the same times (i.e., early or late growing season) in the same group of grapevines (i.e., uninfected or infected with '*Ca.* P. solani'); and (ii) show significant disintegration (i.e., separation with significant dissipation index). U, uninfected grapevines; I, grapevines infected with '*Ca.* P. solani'; E, early growing season; L, late growing season. Table S7. Primers and probes used in this grapevine gene expression analysis.

Author Contributions: Conceptualization, M.D. and M.P.-N.; methodology, B.Š., R.S., Š.T., M.K., Ž.R., M.P., G.B., N.L., K.G., C.S., and M.R.-B.; validation, M.P.-N., T.R., and B.A.; resources, M.R.-B. and G.B.; visualization, A.K., writing—original draft preparation, M.D.; writing—review and editing, M.D., G.B., M.P.-N., Ž.R., M.R.-B., T.R., K.G., and N.L.; project administration, M.D.; funding acquisition, M.D. and G.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Slovenian Research Agency (ARRS) grant numbers: J4-2544, N2-0078, P4-0165, P2-0103; the young researcher grant of B.Š; and the Austrian Science Fund (FWF) grant numbers: I 5042-B and I 2763-B29.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Project accession: PRJEB42777. mRNA samples accessions: ERS5673290, ERS5673291, ERS5673292, ERS5673293, ERS5673294, ERS5673295, ERS5673296, ERS5673297, ERS5673298, ERS5673299, ERS5673300, ERS5673301, ERS5673302, ERS5673303, ERS5673304, ERS5673305, ERS5673306, ERS5673307, ERS5673308, ERS5673309, ERS5673310, ERS5673311. sRNA samples accessions: ERS5672105, ERS5672104, ERS5672103, ERS5672102, ERS5672101, ERS5672100, ERS5672099, ERS5672098, ERS5672097, ERS5672096, ERS5672095, ERS5672094, ERS5672093, ERS5672092, ERS5672091, ERS5672090.

Acknowledgments: Open Access Funding by the Austrian Fund Science (FWF).The authors wish to thank Karel Hanak for his support regarding the collection of the grapevine samples and Christopher Berrie for his linguistic touch.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Starý, M.; Válová, P.; Šafářová, D.; Lauterer, P.; Ackermann, P.; Navrátil, M. Survey and molecular detection of Bois noir in vineyards of the Czech Republic. *Hortic. Sci.* **2013**, *40*, 83–87. [CrossRef]
2. Quaglino, F.; Zhao, Y.; Casati, P.; Bulgari, D.; Bianco, P.A.; Wei, W.; Davis, R.E. '*Candidatus* Phytoplasma solani', a novel taxon associated with stolbur- and bois noir-related diseases of plants. *Int. J. Syst. Evol. Microbiol.* **2013**, *63*, 2879–2894. [CrossRef] [PubMed]
3. Dermastia, M. Interactions between grapevines and grapevine yellows phytoplasmas BN and FD. In *Grapevine Yellows Diseases and Their Phytoplasma Agents*; Dermastia, M., Bertaccini, A., Constable, F., Mehle, N., Eds.; Springer: Cham, Switzerland, 2017; pp. 47–67.
4. Albertazzi, G.; Milc, J.; Caffagni, A.; Francia, E.; Roncaglia, E.; Ferrari, F.; Tagliafico, E.; Stefani, E.; Pecchioni, N. Gene expression in grapevine cultivars in response to Bois Noir phytoplasma infection. *Plant Sci.* **2009**, *176*, 792–804. [CrossRef]
5. Bertamini, M.; Nedunchezhian, N. Decline of photosynthetic pigments, ribulose-1,5-bisphosphate carboxylase and soluble protein contents, nitrate reductase and photosynthetic activities, and changes in tylakoid membrane protein pattern in canopy shade grapevine (*Vitis vinifera* L.). *Photosynthetica* **2001**, *39*, 529–537. [CrossRef]
6. Bertamini, M.; Nedunchezhian, N. Effects of phytoplasma [stolbur-subgroup (Bois noir-BN)] on photosynthetic pigments, saccharides, ribulose 1,5-bisphosphate carboxylase, nitrate and nitrite reductases, and photosynthetic activities in field-grown grapevine (*Vitis vinifera* L.) cv. *Photosynthetica* **2001**, *39*, 119–122. [CrossRef]
7. Bertamini, M.; Nedunchezhian, N.; Tomasi, F.; Grando, M. Phytoplasma [Stolbur-subgroup (Bois Noir-BN)] infection inhibits photosynthetic pigments, ribulose-1,5-bisphosphate carboxylase and photosynthetic activities in field grown grapevine (*Vitis vinifera* L. cv. Chardonnay) leaves. *Physiol. Mol. Plant Pathol.* **2002**, *61*, 357–366. [CrossRef]
8. Hren, M.; Nikolić, P.; Rotter, A.; Blejec, A.; Terrier, N.; Ravnikar, M.; Dermastia, M.; Gruden, K. "Bois noir" phytoplasma induces significant reprogramming of the leaf transcriptome in the field grown grapevine. *BMC Genom.* **2009**, *10*, 460. [CrossRef]
9. Lepka, P.; Stitt, M.; Moll, E.; Seemmüller, E. Effect of phytoplasmal infection on concentration and translocation of carbohydrates and amino acids in periwinkle and tobacco. *Physiol. Mol. Plant Pathol.* **1999**, *55*, 59–68. [CrossRef]
10. Guthrie, J.N.; Walsh, K.B.; Scott, P.T.; Rasmussen, T.S.; Group, P.S.; Industries, P.; Centre, R.S.; Road, M.; Guthrie, J.N.; Walsh, K.B.; et al. The phytopathology of Australian papaya dieback: A proposed role for the phytoplasma. *Physiol. Mol. Plant Pathol.* **2001**, *58*, 23–30. [CrossRef]
11. Maust, B.E.; Espadas, F.; Talavera, C.; Aguilar, M.; Santamaría, J.M.; Oropeza, C. Changes in carbohydrate metabolism in coconut palms infected with the lethal yellowing phytoplasma. *Phytopathology* **2003**, *93*, 976–981. [CrossRef]
12. Gai, Y.P.; Han, X.J.; Li, Y.Q.; Yuan, C.Z.; Mo, Y.Y.; Guo, F.Y.; Liu, Q.X.; Ji, X.L. Metabolomic analysis reveals the potential metabolites and pathogenesis involved in mulberry yellow dwarf disease. *Plant Cell Environ.* **2014**, *37*, 1474–1490. [CrossRef] [PubMed]
13. Prezelj, N. *Molecular Interactions between Phytoplasmal Causal Agents of Grapevine Yellows Disease and Grapevine (Vitis vinifera L.)*; University of Ljubljana: Ljubljana, Slovenia, 2014.
14. Prezelj, N.; Covington, E.; Roitsch, T.; Gruden, K.; Fragner, L.; Weckwerth, W.; Chersicola, M.; Vodopivec, M.; Dermastia, M. Metabolic consequences of infection of grapevine (*Vitis vinifera* L.) cv. "Modra frankinja" with flavescence dorée phytoplasma. *Front. Plant Sci.* **2016**, *7*, 711. [CrossRef] [PubMed]
15. Luge, T.; Kube, M.; Freiwald, A.; Meierhofer, D.; Seemüller, E.; Sauer, S. Transcriptomics assisted proteomic analysis of *Nicotiana occidentalis* infected by *Candidatus* Phytoplasma mali strain AT. *Proteomics* **2014**, *14*, 1882–1889. [CrossRef] [PubMed]
16. Santi, S.; De Marco, F.; Polizzotto, R.; Grisan, S.; Musetti, R. Recovery from stolbur disease in grapevine involves changes in sugar transport and metabolism. *Front. Plant Sci.* **2013**, *4*, 171. [CrossRef]

17. Prezelj, N.; Fragener, L.; Weckwerth, W.; Dermastia, M. Metabolome of grapevine leaf vein-enriched tissue infected with '*Candidatus* Phytoplasma solani'. *Mitt. Klost. Rebe Wein Obs. Früchteverwertung* **2016**, *66*, 74–78.
18. Santi, S.; Grisan, S.; Pierasco, A.; De Marco, F.; Musetti, R. Laser microdissection of grapevine leaf phloem infected by stolbur reveals site-specific gene responses associated to sucrose transport and metabolism. *Plant Cell Environ.* **2013**, *36*, 343–355. [CrossRef] [PubMed]
19. Covington Dunn, E.; Roitsch, T.; Dermastia, M.; Covington, E.D.; Roitsch, T.; Dermastia, M. Determination of the activity signature of key carbohydrate metabolism enzymes in phenolic-rich grapevine tissues. *Acta Chim. Slov.* **2016**, *63*, 757–762. [CrossRef]
20. Rotter, A.; Nikolič, P.; Turnšek, N.; Kogovšek, P.; Blejec, A.; Gruden, K.; Dermastia, M. Statistical modeling of long-term grapevine response to '*Candidatus* Phytoplasma solani' infection in the field. *Eur. J. Plant Pathol.* **2018**, *150*, 653–668. [CrossRef]
21. Hren, M.; Ravnikar, M.; Brzin, J.; Ermacora, P.; Carraro, L.; Bianco, P.A.; Casati, P.; Borgo, M.; Angelini, E.; Rotter, A.; et al. Induced expression of sucrose synthase and alcohol dehydrogenase I genes in phytoplasma-infected grapevine plants grown in the field. *Plant Pathol.* **2009**, *58*, 170–180. [CrossRef]
22. Dermastia, M.; Nikolic, P.; Chersicola, M.; Gruden, K. Transcriptional profiling in infected and recovered grapevine plant responses to '*Candidatus* Phytoplasma solani'. *Phytopathogenic Mollicutes* **2015**, *5*, S123. [CrossRef]
23. Rusjan, D.; Mikulic-Petkovsek, M. Phenolic responses in 1-year-old canes of V itis vinifera cv. Chardonnay induced by grapevine yellows (Bois noir). *Aust. J. Grape Wine Res.* **2015**, *21*, 123–134. [CrossRef]
24. Rusjan, D.; Halbwirth, H.; Stich, K.; Mikulič-Petkovšek, M.; Veberič, R. Biochemical response of grapevine variety "Chardonnay" (Vitis vinifera L.) to infection with grapevine yellows (Bois noir). *Eur. J. Plant Pathol.* **2012**, *134*, 231–237. [CrossRef]
25. Rusjan, D.; Veberič, R.; Mikulič-Petkovšek, M. The response of phenolic compounds in grapes of the variety "Chardonnay" (*Vitis vinifera* L.) to the infection by phytoplasma Bois noir. *Eur. J. Plant Pathol.* **2012**, *133*, 965–974. [CrossRef]
26. Landi, L.; Romanazzi, G. Seasonal variation of defense-related gene expression in leaves from bois noir affected and recovered grapevines. *J. Agric. Food Chem.* **2011**, *59*, 6628–6637. [CrossRef]
27. Dermastia, M. Plant hormones in phytoplasma infected plants. *Front. Plant Sci.* **2019**, *10*, 477. [CrossRef]
28. Cora, D.; Re, A.; Caselle, M.; Bussolino, F. MicroRNA-mediated regulatory circuits: Outlook and perspectives. *Phys. Biol.* **2017**, *14*. [CrossRef] [PubMed]
29. Megraw, M.; Cumbie, J.S.; Ivanchenko, M.G.; Filichkin, S.A. Small genetic circuits and microRNAs: Big players in polymerase II transcriptional control in plants. *Plant Cell* **2015**, *28*, 286–303. [CrossRef]
30. Wang, H.L.V.; Chekanova, J.A. Small RNAs: Essential regulators of gene expression and defenses against environmental stresses in plants. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 356–381. [CrossRef]
31. Belli Kullan, J.; Lopes Paim Pinto, D.; Bertolini, E.; Fasoli, M.; Zenoni, S.; Tornielli, G.B.; Pezzotti, M.; Meyers, B.C.; Farina, L.; Pè, M.E.; et al. miRVine: A microRNA expression atlas of grapevine based on small RNA sequencing. *BMC Genom.* **2015**, *16*, 393. [CrossRef]
32. Ehya, F.; Monavarfeshani, A.; Mohseni Fard, E.; Karimi Farsad, L.; Khayam Nekouei, M.; Mardi, M.; Salekdeh, G.H. Phytoplasma-responsive microRNAs modulate hormonal, nutritional, and stress signalling pathways in Mexican lime trees. *PLoS ONE* **2013**, *8*. [CrossRef]
33. Shao, F.; Zhang, Q.; Liu, H.; Lu, S.; Qiu, D. Genome-wide identification and analysis of microRNAs involved in witches'-broom phytoplasma response in *Ziziphus jujuba*. *PLoS ONE* **2016**, *11*, e0166099. [CrossRef]
34. Gai, Y.-P.; Zhao, H.-N.; Zhao, Y.-N.; Zhu, B.-S.; Yuan, S.-S.; Li, S.; Guo, F.-Y.; Ji, X.-L. MiRNA-seq-based profiles of miRNAs in mulberry phloem sap provide insight into the pathogenic mechanisms of mulberry yellow dwarf disease. *Sci. Rep.* **2018**, *8*, 812. [CrossRef] [PubMed]
35. Fan, G.; Niu, S.; Zhao, Z.; Deng, M.; Xu, E.; Wang, Y.; Yang, L. Identification of microRNAs and their targets in Paulownia fortunei plants free from phytoplasma pathogen after methyl methane sulfonate treatment. *Biochimie* **2016**, *127*, 271–280. [CrossRef] [PubMed]
36. Chitarra, W.; Pagliarani, C.; Abbà, S.; Boccacci, P.; Birello, G.; Rossi, M.; Palmano, S.; Marzachì, C.; Perrone, I.; Gambino, G. miRVIT: A novel miRNA database and its application to uncover *Vitis* responses to favescence dorée infection. *Front. Plant Sci.* **2018**, *9*, 1034. [CrossRef] [PubMed]
37. Snyman, M.C.; Solofoharivelo, M.-C.; Souza-Richards, R.; Stephan, D.; Murray, S.; Burger, J.T. The use of high-throughput small RNA sequencing reveals differentially expressed microRNAs in response to aster yellows phytoplasma-infection in *Vitis vinifera* cv. 'Chardonnay'. *PLoS ONE* **2017**, *12*, e0182629. [CrossRef] [PubMed]
38. Emamjomeh, A.; Saboori Robat, E.; Zahiri, J.; Solouki, M.; Khosravi, P. Gene co-expression network reconstruction: A review on computational methods for inferring functional information from plant-based expression data. *Plant Biotechnol. Rep.* **2017**, *11*, 71–86. [CrossRef]
39. Škrlj, B.; Pompe Novak, M.; Brader, G.; Anžič, B.; Ramšak, Ž.; Gruden, K.; Kralj, J.; Lavrač, N.; Roitsch, T.; Dermastia, M. New cross-talks between pathways involved in grapevine infection with '*Candidatus* Phytoplasma solani' revealed by temporal network modelling. *Plants* **2021**, in press.
40. Aryan, A.; Brader, G.; Mörtel, J.; Pastar, M.; Riedle-Bauer, M. An abundant '*Candidatus* Phytoplasma solani' tuf b strain is associated with grapevine, stinging nettle and Hyalesthes obsoletus. *Eur. J. Plant Pathol. Eur. Found. Plant Pathol.* **2014**, *140*, 213–227. [CrossRef]

41. Murolo, S.; Garbarino, M.; Mancini, V.; Romanazzi, G. Spatial pattern of Bois noir: Case study of a delicate balance between disease progression and recovery. *Sci. Rep.* **2020**, *10*. [CrossRef] [PubMed]
42. Riedle-Bauer, M.; Hanak, K.; Regner, F.; Tiefenbrunner, W. Influence of pruning measures on recovery of Bois Noir-infected grapevines. *J. Phytopathol.* **2010**, *158*, 628–632. [CrossRef]
43. Dermastia, M.; Kube, M.; Šeruga-Musić, M. Transcriptomic and proteomic studies of phytoplasma-infected plants. In *Phytoplasmas: Plant Pathogenic Bacteria—III*; Springer: Singapore, 2019; pp. 35–55.
44. Williams, M.; Cheng, Y.Y.; Blenkiron, C.; Reid, G. Exploring Mechanisms of MicroRNA Downregulation in Cancer. *MicroRNA* **2017**, *6*, 2–16. [CrossRef]
45. Zombardo, A.; Crosatti, C.; Bagnaresi, P.; Bassolino, L.; Reshef, N.; Puccioni, S.; Faccioli, P.; Tafuri, A.; Delledonne, M.; Fait, A.; et al. Transcriptomic and biochemical investigations support the role of rootstock-scion interaction in grapevine berry quality. *BMC Genom.* **2020**, *21*, 468. [CrossRef]
46. Thimm, O.; Bläsing, O.; Gibon, Y.; Nagel, A.; Meyer, S.; Krüger, P.; Selbig, J.; Müller, L.A.; Rhee, S.Y.; Stitt, M. MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* **2004**, *37*, 914–939. [CrossRef] [PubMed]
47. Ballicora, M.A.; Iglesias, A.A.; Preiss, J. ADP-Glucose Pyrophosphorylase: A Regulatory Enzyme for Plant Starch Synthesis. *Photosynth. Res.* **2004**, *79*, 1–24. [CrossRef] [PubMed]
48. Mardi, M.; Karimi Farsad, L.; Gharechahi, J.; Salekdeh, G.H. In-depth transcriptome sequencing of Mexican lime trees infected with "*Candidatus* Phytoplasma aurantifolia". *PLoS ONE* **2015**, *10*, e0130425. [CrossRef] [PubMed]
49. Krishnamurthy, P.; Muthusamy, M.; Kim, J.A.; Jeong, M.J.; Lee, S.I. Brassica rapa expansin-like B1 gene (BrEXLB1) regulate growth and development in transgenic Arabidopsis and elicits response to abiotic stresses. *J. Plant Biochem. Biotechnol.* **2019**, *28*, 437–446. [CrossRef]
50. Paolacci, A.R.; Catarcione, G.; Ederli, L.; Zadra, C.; Pasqualini, S.; Badiani, M.; Musetti, R.; Santi, S.; Ciaffi, M. Jasmonate-mediated defence responses, unlike salicylate-mediated responses, are involved in the recovery of grapevine from bois noir disease. *BMC Plant Biol.* **2017**, *17*, 118. [CrossRef]
51. Sugio, A.; Kingdom, H.N.; MacLean, A.M.; Grieve, V.M.; Hogenhout, S.A. Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E1254–E1263. [CrossRef] [PubMed]
52. Music, M.S.; Samarzija, I.; Hogenhout, S.A.; Haryono, M.; Cho, S.T.; Kuo, C.H. The genome of '*Candidatus* Phytoplasma solani' strain SA-1 is highly dynamic and prone to adopting foreign sequences. *Syst. Appl. Microbiol.* **2019**, *42*, 117–127. [CrossRef]
53. Ren, H.; Gray, W.M. SAUR Proteins as Effectors of Hormonal and Environmental Signals in Plant Growth. *Mol. Plant* **2015**, *8*, 1153–1164. [CrossRef]
54. Hoshi, A.; Oshima, K.; Kakizawa, S.; Ishii, Y.; Ozeki, J.; Hashimoto, M. A unique virulence factor for proliferation and dwarfism in plants identified from a phytopathogenic bacterium. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6416–6421. [CrossRef] [PubMed]
55. Minato, N.; Himeno, M.; Hoshi, A.; Maejima, K.; Komatsu, K.; Takebayashi, Y.; Kasahara, H.; Yusa, A.; Yamaji, Y.; Oshima, K.; et al. The phytoplasmal virulence factor TENGU causes plant sterility by downregulating of the jasmonic acid and auxin pathways. *Sci. Rep.* **2014**, *4*, 7399. [CrossRef] [PubMed]
56. Gomes Ferreira, M.D.; Araújo Castro, J.; Santana Silva, R.J.; Micheli, F. HVA22 from citrus: A small gene family whose some members are involved in plant response to abiotic stress. *Plant Physiol. Biochem.* **2019**, *142*, 395–404. [CrossRef]
57. Monavarfeshani, A.; Mirzaei, M.; Sarhadi, E.; Amirkhani, A.; Khayam Nekouei, M.; Haynes, P.A.; Mardi, M.; Salekdeh, G.H. Shotgun proteomic analysis of the Mexican lime tree infected with "*Candidatus* Phytoplasma aurantifolia". *J. Proteome Res.* **2013**, *12*, 785–795. [CrossRef] [PubMed]
58. Musetti, R.; Paolacci, A.; Ciaffi, M.; Tanzarella, O.A.; Polizzotto, R.; Tubaro, F.; Mizzau, M.; Ermacora, P.; Badiani, M.; Osler, R. Phloem cytochemical modification and gene expression following the recovery of apple plants from apple proliferation disease. *Phytopathology* **2010**, *100*, 390–399. [CrossRef]
59. Margaria, P.; Abbà, S.; Palmano, S. Novel aspects of grapevine response to phytoplasma infection investigated by a proteomic and phospho-proteomic approach with data integration into functional networks. *BMC Genom.* **2013**, *14*, 38. [CrossRef]
60. Margaria, P.; Palmano, S. Response of the *Vitis vinifera* L. cv. "Nebbiolo" proteome to Flavescence dorée phytoplasma infection. *Proteomics* **2011**, *11*, 212–224. [CrossRef]
61. Musetti, R.; Marabottini, R.; Badiani, M.; Martini, M.; Sanit, L.; Sanita di Toppi, L.; Borselli, S.; Borgo, M.; Osler, R. On the role of H 2 O 2 in the recovery of grapevine (Vitis vinifera cv. Prosecco) from Flavescence dorée disease. *Funct. Plant Biol.* **2007**, *34*, 750–758. [CrossRef] [PubMed]
62. Dermastia, M.; Bertaccini, A.; Constable, F.; Mehle, N. *Grapevine Yellows Diseases and Their Phytoplasma Agents: Biology and Detection*; Springer: Cham, Switzerland, 2017; ISBN 9783319506487.
63. Janusz, G.; Pawlik, A.; Świderska-Burek, U.; Polak, J.; Sulej, J.; Jarosz-Wilkołazka, A.; Paszczyński, A. Laccase properties, physiological functions, and evolution. *Int. J. Mol. Sci.* **2020**, *21*, 966. [CrossRef]
64. Cao, Y.; Fan, G.; Wang, Z.; Gu, Z. Phytoplasma-induced changes in the acetylome and succinylome of paulownia tomentosa provide evidence for involvement of acetylated proteins in witches' broom disease. *Mol. Cell. Proteom.* **2019**, *18*, 1210–1226. [CrossRef] [PubMed]

65. Musetti, R.; Buxa, S.V.; De Marco, F.; Loschi, A.; Polizzotto, R.; Kogel, K.-H.; van Bel, A.J.E. Phytoplasma-triggered Ca(2+) influx is involved in sieve-tube blockage. *Mol. Plant. Microbe. Interact.* **2013**, *26*, 379–386. [CrossRef] [PubMed]
66. Takanashi, K.; Shitan, N.; Yazaki, K. The multidrug and toxic compound extrusion (MATE) family in plants. *Plant Biotechnol.* **2014**, *31*, 417–430. [CrossRef]
67. Kusakizako, T.; Miyauchi, H.; Ishitani, R.; Nureki, O. Structural biology of the multidrug and toxic compound extrusion superfamily transporters. *Biochim. Biophys. Acta Biomembr.* **2020**, *1862*, 183154. [CrossRef]
68. Sun, X.; Gilroy, E.M.; Chini, A.; Nurmberg, P.L.; Hein, I.; Lacomme, C.; Birch, P.R.J.; Hussain, A.; Yun, B.-W.; Loake, G.J. ADS1 encodes a MATE-transporter that negatively regulates plant disease resistance. *New Phytol.* **2011**, *192*, 471–482. [CrossRef]
69. Zhang, H.; Zhu, H.; Pan, Y.; Yu, Y.; Luan, S.; Li, L. A DTX/MATE-type transporter facilitates abscisic acid efflux and modulates ABA sensitivity and drought tolerance in Arabidopsis. *Mol. Plant* **2014**, *7*, 1522–1532. [CrossRef] [PubMed]
70. Shi, J.; Du, X. Identification, characterization and expression analysis of calmodulin and calmodulin-like proteins in Solanum pennellii. *Sci. Rep.* **2020**, *10*. [CrossRef]
71. Poovaiah, B.W.; Du, L.; Wang, H.; Yang, T. Recent advances in calcium/calmodulin-mediated signaling with an emphasis on plant-microbe interactions. *Plant Physiol.* **2013**, *163*, 531–542. [CrossRef] [PubMed]
72. Koshiba, T.; Saito, E.; Ono, N.; Yamamoto, N.; Satô, M. Purification and properties of flavin- and molybdenum-containing aldehyde oxidase from coleoptiles of maize. *Plant Physiol.* **1996**, *110*, 781–789. [CrossRef]
73. Mendel, R.-R.; Müller, A.J. A Common Genetic Determinant of Xanthine Dehydrogenase and Nitrate Reductase in Nicotiana tabacum. *Biochem. Physiol. Pflanz.* **1976**, *170*, 538–541. [CrossRef]
74. Watanabe, S.; Sato, M.; Sawada, Y.; Tanaka, M.; Matsui, A.; Kanno, Y.; Hirai, M.Y.; Seki, M.; Sakamoto, A.; Seo, M. Arabidopsis molybdenum cofactor sulfurase ABA3 contributes to anthocyanin accumulation and oxidative stress tolerance in ABA-dependent and independent ways. *Sci. Rep.* **2018**, *8*. [CrossRef] [PubMed]
75. Mehle, N.; Dermastia, M.; Brus, R.; Jurc, D. First report of '*Candidatus Phytoplasma ulmi*' in *Ulmus minor* and *Ulmus glabra* in Slovenia. *Plant Dis.* **2017**, *101*. [CrossRef]
76. Le Roy, J.; Huss, B.; Creach, A.; Hawkins, S.; Neutelings, G. Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Front. Plant Sci.* **2016**, *7*, 735. [CrossRef]
77. Rehman, H.M.; Nawaz, M.A.; Shah, Z.H.; Ludwig-Müller, J.; Chung, G.; Ahmad, M.Q.; Yang, S.H.; Lee, S.I. Comparative genomic and transcriptomic analyses of Family-1 UDP glycosyltransferase in three Brassica species and Arabidopsis indicates stress-responsive regulation. *Sci. Rep.* **2018**, *8*. [CrossRef] [PubMed]
78. Zhang, Y.; Wang, X.; Chang, X.; Sun, M.; Zhang, Y.; Li, W.; Li, Y. Overexpression of germin-like protein GmGLP10 enhances resistance to Sclerotinia sclerotiorum in transgenic tobacco. *Biochem. Biophys. Res. Commun.* **2018**, *497*, 160–166. [CrossRef]
79. Liu, N.; Shang, W.; Li, C.; Jia, L.; Wang, X.; Xing, G.; Zheng, W.M. Evolution of the SPX gene family in plants and its role in the response mechanism to phosphorus stress. *Open Biol.* **2018**, *8*, 170231. [CrossRef] [PubMed]
80. Lu, Y.-T.; Li, M.-Y.; Cheng, K.-T.; Tan, C.M.; Su, L.-W.; Lin, W.-Y.; Shih, H.-T.; Chiou, T.-J.; Yang, J.-Y. Transgenic plants that express the phytoplasma effector SAP11 show altered phosphate starvation and defense responses. *Plant Physiol.* **2014**, *164*, 1456–1469. [CrossRef] [PubMed]
81. Rubio, V.; Linhares, F.; Solano, R.; Martín, A.C.; Iglesias, J.; Leyva, A.; Paz-Ares, J. A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes Dev.* **2001**, *15*, 2122–2133. [CrossRef]
82. Ouellette, A.J.; Hsieh, M.M.; Nosek, M.T.; Cano-Gauci, D.F.; Huttner, K.M.; Buick, R.N.; Selsted, M.E. Mouse Paneth cell defensins: Primary structures and antibacterial activities of numerous cryptdin isoforms. *Infect. Immun.* **1994**, *62*, 5040–5047. [CrossRef]
83. Ouellette, A.J.; Satchell, D.P.; Hsieh, M.M.; Hagen, S.J.; Selsted, M.E. Characterization of luminal Paneth cell α-defensins in mouse small intestine: Attenuated antimicrobial activities of peptides with truncated amino termini. *J. Biol. Chem.* **2000**, *275*, 33969–33973. [CrossRef]
84. Ayabe, T.; Satchell, D.P.; Wilson, C.L.; Parks, W.C.; Selsted, M.E.; Ouellette, A.J. Secretion of microbicidal α-defensins by intestinal Paneth cells in response to bacteria. *Nat. Immunol.* **2000**, *1*, 113–118. [CrossRef]
85. Carra, A.; Gambino, G.; Schubert, A. A cetyltrimethylammonium bromide-based method to extract low-molecular-weight RNA from polysaccharide-rich plant tissues. *Anal. Biochem.* **2007**, *360*, 318–320. [CrossRef]
86. R Core Team. R A Language and Environment for Statistical Computing. Scientific Research Publishing. 2017. Available online: https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=2144573 (accessed on 21 January 2021).
87. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [CrossRef]
88. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2009**, *26*, 139–140. [CrossRef] [PubMed]
89. Law, C.W.; Chen, Y.; Shi, W.; Smyth, G.K. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **2014**, *15*, R29. [CrossRef] [PubMed]
90. Bateman, A.; Agrawal, S.; Birney, E.; Bruford, E.A.; Bujnicki, J.M.; Cochrane, G.; Cole, J.R.; Dinger, M.E.; Enright, A.J.; Gardner, P.P.; et al. RNAcentral: A vision for an international database of RNA sequences. *RNA* **2011**, *17*, 1941–1946. [CrossRef] [PubMed]
91. Kozomara, A.; Griffiths-Jones, S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **2014**, *42*, D68. [CrossRef] [PubMed]

92. Shahid, S.; Axtell, M.J. Identification and annotation of small RNA genes using ShortStack. *Methods* **2014**, *67*, 20–27. [CrossRef]
93. Lei, J.; Sun, Y. MiR-PREFeR: An accurate, fast and easy-to-use plant miRNA prediction tool using small RNA-Seq data. *Bioinformatics* **2014**, *30*, 2837–2839. [CrossRef]
94. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef] [PubMed]
95. Huang, Y.; Niu, B.; Gao, Y.; Fu, L.; Li, W. CD-HIT Suite: A web server for clustering and comparing biological sequences. *Bioinformatics* **2010**, *26*, 680–682. [CrossRef]
96. De Oliveira, L.F.V.; Christoff, A.P.; Margis, R. isomiRID: A framework to identify microRNA isoforms. *Bioinformatics* **2013**, *29*, 2521–2523. [CrossRef]
97. Gebert, D.; Hewel, C.; Rosenkranz, D. Unitas: The universal tool for annotation of small RNAs. *BMC Genom.* **2017**, *18*, 644. [CrossRef]
98. Dai, X.; Zhao, P.X. PsRNATarget: A plant small RNA target analysis server. *Nucleic Acids Res.* **2011**, *39*. [CrossRef] [PubMed]
99. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef]
100. Ramšak, Ž.; Baebler, Š.; Rotter, A.; Korbar, M.; Mozetič, I.; Usadel, B.; Gruden, K. GoMapMan: Integration, consolidation and visualization of plant gene annotations within the MapMan ontology. *Nucleic Acids Res.* **2014**, *42*, D1167–D1175. [CrossRef] [PubMed]
101. Rosvall, M.; Axelsson, D.; Bergstrom, C.T. The map equation. *Eur. Phys. J. Spec. Top.* **2009**, *178*, 13–23. [CrossRef]
102. Hagberg, A.; Schult, D.; Swart, P. Exploring network structure, dynamics, and function using networkX. In Proceedings of the 7th Python in Science Conference, Pasadena, CA, USA, 19–24 August 2008; pp. 11–15.
103. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef]
104. Baebler, Š.; Svalina, M.; Petek, M.; Stare, K.; Rotter, A.; Pompe-Novak, M.; Gruden, K. QuantGenius: Implementation of a decision support system for qPCR-based gene quantification. *BMC Bioinform.* **2017**, *18*, 276. [CrossRef] [PubMed]
105. Jammer, A.; Gasperl, A.; Luschin-Ebengreuth, N.; Heyneke, E.; Chu, H.; Cantero-Navarro, E.; Großkinsky, D.K.; Albacete, A.A.; Stabentheiner, E.; Franzaring, J.; et al. Simple and robust determination of the activity signature of key carbohydrate metabolism enzymes for physiological phenotyping in model and crop plants. *J. Exp. Bot.* **2015**, *66*, 5531–5542. [CrossRef]
106. Anžič, B. Transformation of Periwinkle (Catharanthus roseus) for Studying Infection with 'Candidatus Phytoplasma Solani' and Enzymatic Activities in Infectedon of Madagascar Plants, University of Ljubljana. Available online: https://repozitorij.uni-lj.si/IzpisGradiva.php?id=111641&lang=slv (accessed on 29 February 2021).
107. Fimognari, L.; Dölker, R.; Kaselyte, G.; Jensen, C.N.G.; Akhtar, S.S.; Großkinsky, D.K.; Roitsch, T. Simple semi-high throughput determination of activity signatures of key antioxidant enzymes for physiological phenotyping. *Plant Methods* **2020**, *16*, 42. [CrossRef]

2.2 Candidate Effector Proteins of '*Candidatus* Phytoplasma solani' are Associated with Modulation of Plant Carbohydrate Metabolism Toward Effective Glycolysis, with Accelerated Ascorbate-Glutathione Cycle, and with Induction of Autophagosomes

Marina Dermastia, Špela Tomaž, Rebeka Strah, Timotej Čepin, Tjaša Lukan, Anna Coll, Barbara Dušak, Barbara Anžič, Thomas Roitsch, Stefanie Wienkoop, Wolfram Weckwert, Kristina Gruden, Maruša Pompe Novak and Günter Brader

Manuscript draft

In our second publication, we used a bioinformatics approach to identify six candidate '*Ca.* P. solani' effector proteins, which we further characterized using functional analysis. Using enzyme activity assays, we showed different candidate effectors can affect the activity of several host enzymes involved in different steps of carbohydrate metabolism, as well as enzymes related to regulation of oxidative stress. We used a proteomics approach to identify potential plant protein targets of the effector PoSTOSP28 and confirmed its interaction with two grapevine phosphoglucomutase enzymes (VvPGM). Finally, our localization studies in tobacco revealed a connection between the effector PoSTOSP28 and formation of autophagosomes.

The PhD candidate adapted and performed a protein pull-down assay from tobacco leaves expressing a potential '*Ca.* P. solani' effector protein PoSTOSP28. She prepared the samples for mass spectrometry analysis, searched for orthologues of identified tobacco targets in grapevine, and confirmed the interactions between the effector and grapevine proteins with a co-immunoprecipitation assay. The PhD candidate subsequently wrote the results and methods section connected to these experiments in the resulting manuscript. Co-immunoprecipitation assay results not included in the manuscript are included in Appendix A.

Candidate effector proteins of '*Candidatus* Phytoplasma solani' are associated with modulation of plant carbohydrate metabolism toward effective glycolysis, with accelerated ascorbate-glutathione cycle, and with induction of autophagosomes

Marina Dermastia[1*], Špela Tomaž[1,2], Rebeka Strah[1,2], Timotej Čepin[1], Tjaša Lukan[1], Anna Coll[1], Barbara Dušak [1], Barbara Anžič[1], Thomas Roitsch[3], Stefanie Wienkoop[4], Wolfram Weckwert[4], Kristina Gruden[1], Maruša Pompe Novak[1,5], Günter Brader[6]

[1] National Institute of Biology, 1000 Ljubljana, Slovenia;
[2] Jožef Stefan International Postgraduate School, 1000 Ljubljana, Slovenia;
[3] Department of Plant and Environmental Sciences, University of Copenhagen, 2630 Taastrup, Denmark;
[4] Department of Ecogenomics and Systems Biology, Faculty of Life Sciences University of Vienna, Vienna, Austria;
[5] Faculty of Viticulture and Enology, University of Nova Gorica, 5271 Vipava, Slovenia;
[6] Austrian Institute of Technology, 3430 Tulln, Austria;
* Correspondence: marina.dermastia@nib.si

Abstract

'*Candidatus* Phytoplasma solani' is a member of the ribosomal group 16SrXII-A of intracellular plant pathogenic bacteria in the class Mollicutes. It causes phloem-feeding, insect-borne diseases in grapevine and solanaceous crops with significant economic impact on yield. By deciphering the '*Ca.* P. solani' genome derived from the infected grapevine (*Vitis vinifera*) cultivar Zweigelt, we discovered up to 20 effector proteins, six of which were selected for further study. This first analysis of the physiological role of candidate effectors of '*Ca.* P. solani' in *Nicotiana benthamiana* leaves after *Agrobacterium*-mediated transient transformation with the effector constructs confirms their involvement in reprogramming the host carbohydrate metabolism to favor its own growth and infection. They were specifically associated with three different metabolic pathways leading to fructose-6-phosphate as an input substrate for glycolysis, the upper part of which is encoded in the phytoplasma genome. The most active effector candidate was PoSTOSP06, which shares some similarities with a putative effector from '*Ca.* P. solani'. PoSTOSP28, previously shown to be an antigenic membrane protein required for phytoplasma-insect interaction, specifically interacted with phosphoglucomutase and increased its enzyme activity. PoSTOSP04, the only candidate effector that was identical to one previously identified in '*Ca.* P. solani', was associated with the induction of the ascorbate-glutathione cycle along with other effector candidates. Interestingly, PoSTOSP06, PoSTOSP013 and PoSTOSP28 were localized in the nucleus. It was shown here for the first time in phytoplasma-plant interaction that especially PoSTOSP28 induces autophagosomes, and the potential significance of this finding is discussed.

Introduction

Phytoplasmas are a diverse genus of plant pathogenic bacteria in the class Mollicutes that cause several economically important insect-transmitted diseases (EPPO Global Database, no date). Since it has not been possible to routinely culture phytoplasmas under *in vitro* conditions (Contaldo, Bertaccini, Paltrinieri, Windsor & Windsor 2012; Contaldo, Satta, Zambon, Paltrinieri & Bertaccini 2016), our knowledge of their biology and pathogenicity is still very rudimentary. However, with appropriate bioinformatic analyzes and molecular biology approaches, we are gradually deciphering the life of these plant pathogens.

Phytoplasmas exclusively inhabit the sieve cells of the phloem and lack many genes involved in metabolic pathways important for free-living cells (Namba 2019). In acquiring nutrients from their plant host, phytoplasmas compete with the hosts for the same nutrient substrates. Under conditions of phytoplasma infection, the metabolic networks of the invading pathogen and its host should be closely linked, and even a minor change in metabolism could significantly affect the outcome of the phytoplasma-host interaction. Therefore, phytoplasmas are thought to need to modulate host plant cell metabolism to provide nutrients, energy, and metabolites for successful replication and infection in plants (Prezelj *et al.* 2016a; Dermastia 2019; Dermastia, Kube & Šeruga-Musić 2019).

The genomes of phytoplasmas lack the homologs of the III type secretion system (Kube, Mitrovic, Duduk, Rabus & Seemüller 2012) that form hollow tubes through which effectors pass from the bacterial cytosol directly into the cytosol of host cells and is essential for the virulence of most Gram-negative bacterial pathogens (Abramovitch, Anderson & Martin 2006). Instead, phytoplasmas have a functional Sec-dependent pathway for the export of proteins across the cytoplasmic membrane to the bacterial periplasm and outer membrane (Green & Mecsas 2016). Phytoplasmas use the Sec-dependent system for the secretion of virulence factors or effectors directly into the cytoplasm of plant host phloem sieve cells (Bai *et al.* 2009; Hoshi *et al.* 2009; MacLean *et al.* 2014; Huang *et al.* 2021). Phytoplasma candidate effector proteins can be identified by finding their cleavable signal peptides, which are associated with a functional Sec-dependent pathway (MacLean *et al.* 2011). Once inside the host cell, the effectors often target key proteins to hijack the host cellular machinery and remodel the signaling cascade (Tomkins, Kliot, Marée & Hogenhout 2018). Some known phytoplasma effectors target developmental and immunity-related host transcription factors

(Marrero *et al.* 2020). The secreted AY-WB protein (SAP) 11 (SAP11) is involved in the development of witches broom symptoms by interacting with the transcription factors TEOSINTE BRANCHED 1, CYCLOIDEA, and PCF1. The action of TENGU is related to auxin and jasmonic acid signaling by repressing the expression of auxin response factor 6 and 8 genes, resulting in dwarfism and floral sterility, and the development of leaf-like flowers is related to the degradation of MADS-box transcription factors by SAP54 (Tomkins *et al.* 2018). In addition, SAP05 can hijack the plant ubiquitin receptor RPN10 independently of substrate ubiquitination, allowing phytoplasmas to alter plant architecture and reproduction (Huang *et al.* 2021). Most of the phytoplasma effectors that have been studied in detail are derived from '*Candidatus* Phytoplasma asteris', namely the SAPs from the witches broom strain (Bai *et al.* 2009) and TENGU from the OY strain (Hoshi *et al.* 2009). Several homologs of these effectors have been discovered in other phytoplasma genomes (Siewert *et al.* 2014; Lu *et al.* 2014; Janik *et al.* 2017; Anabestani *et al.* 2017; Pecher *et al.* 2019; Zhou *et al.* 2021).

The molecular mechanisms of other putative secreted virulence factors/effectors by which phytoplasmas manipulate their hosts are not known, thus our understanding of the general mechanisms of phytoplasma interaction with host plants remains elusive. However, a recent study suggests that phytoplasmas have a unique infection strategy and that their effectors have unusual targets (Marrero *et al.* 2020).

'*Ca.* P. solani' from the ribosomal group 16SrXII-A (Quaglino *et al.* 2013) is associated with the most widespread phytoplasma disease of grapevine (*Vitis vinifera*) in Europe, namely bois noir (Dermastia, Bertaccini, Constable & Mehle 2017), and with stolbur disease of plants from the Solanaceae family, including potato (*Solanum tuberosum*), tomato (*Solanum lycopersicum*), and tobacco (*Nicotiana* sp.). The disease in the Solanaceae is already a growing problem in Serbia (Mitrović & Duduk 2011; Mitrović *et al.* 2016), Romania (Ember, Acs, Munyaneza, Crosslin & Kolber 2011) and Turkey (Çağlar *et al.* 2010), and has been reported sporadically in Central Europe including Germany (Preiss, Keuck & Albert 2008). Recently, severe outbreaks have been reported, especially in potatoes in eastern Austria (Brader *et al.* 2017; Riedle-Bauer *et al.* 2017).

Studies on the interactions between host grapevines and '*Ca.* P. solani' at the levels of transcription, regulation, proteins, and metabolism show that infection affects biotic stress signaling, hormonal balance, photosynthesis, glycolysis, oxidative stress processes and

secondary metabolism (Hren *et al.* 2009; Dermastia 2017; Rotter *et al.* 2018; Dermastia *et al.* 2019, 2021; Škrlj *et al.* 2021). Moreover, all these studies indicated that carbohydrate metabolism is severely impaired in infected grapevines, as evidenced by the increased carbohydrate content, changed activity of the corresponding enzymes and/or genes encoding them, and the expression of regulated genes related to carbohydrate synthesis.

Recent sequencing and assembly of the '*Ca.* P. solani' strain SA-1 genome identified 38 putative secreted protein/effector genes (Music Seruga *et al.* 2019). Twenty of these genes are located within phytoplasma potential mobile units. These regions exhibit variations in gene order intermixed with genes of unknown function and a lack of similarity to other phytoplasma genes, suggesting that they are prone to rearrangements and acquisition of new sequences by recombination. Five of the putative effectors were unique to the sequenced strain and two of them resembled proteins secreted from AY-WB (Music Seruga *et al.* 2019). However, the role of none of the identified putative effectors has been elucidated.

Here we report the identification of six potential virulence factors/effectors of the '*Ca.* P. solani' strain isolated from the infected grapevine cv. Zweigelt and their possible roles in the pathogenicity of this phytoplasma.

Material and Methods

Bioinformatic analysis

Ten different genomic DNA sequences from bois noir (BN) phytoplasma (one *tuf* a strain, one *tuf* b2 strain and eight *tuf* b strains) were aligned and annotated. To visualize protein sequence similarity, the sequences were aligned using Geneious Alignment in Geneious Prime 2020.1.1 (https://www.geneious.com/). The presence and location of signal peptide cleavage sites were predicted by SignalP 4.1 (Nielsen 2017) that discriminates signal peptides from transmembrane regions and by LipoP 1.0 (Rahman, Cummings, Harrington & Sutcliffe 2008) that discriminates between lipoprotein signal peptides, other signal peptides and N-terminal membrane helices. Although LipoP 1.0 has been trained on sequences from Gram-negative bacteria only, it also has a good performance on sequences from Gram-positive bacteria (Rahman *et al.* 2008) including mollicute proteins (Bai *et al.* 2009). Transmembrane helices in proteins were predicted using TMHMM 2.0 (Krogh, Larsson, von Heijne & Sonnhammer 2001). The expression profile of phytoplasma proteins *in planta* was studied for BN601.

Plasmid constructs

Full length coding sequences of the six putative effector proteins (PoSTOSP04, PoSTOSP06, PoSTOSP13, PoSTOSP14, PoSTOSP18 and PoSTOSP28), as well as two grapevine phosphoglucomutases (*Vitvi16g00891*, *Vitvi01g00455*), were amplified from BN601-infected grapevine material of cv. Zweigelt using the Phusion® High-Fidelity PCR kit (New England Biolabs, USA) and the primer pairs listed in Supplemental Table 1. The products of correct size were selected on an agarose gel and purified using the Wizard® SV Gel and PRC Clean-Up System kit (Promega, USA). The phosphoglucomutase sequences were first cloned into the pJET1.2/blunt vector using the CloneJET PCR Cloning Kit (Thermo Scientific, USA). The amplified sequences were then cloned into pENTR D-TOPO vector using pENTR™ Directional TOPO® Cloning Kit (Invitrogen, USA) (Supplemental Table 1) and SANGER sequenced (Eurofins Genomics, Germany). The effector sequences were recombined through LR reaction with the Gateway® LR Clonase TM II Enzyme Mix (Invitrogen, USA) into pH7WGY2 Gateway destination vectors containing the enhanced yellow fluorescent protein (YFP) (VIB, Belgium). For co-immunoprecipitation assay, the phosphoglucomutase sequences were cloned into the pJCV52 Gateway vector (VIB, Belgium) (Karimi, Inzé & Depicker 2002) containing the hemagglutinin A (HA) tag.

Besides the vectors containing effector sequences, three additional constructs were used for co-localization studies, i.e. histone 2B (H2B) fused to monomeric red fluorescent protein 1 (mRFP1) (Lukan *et al.* 2018) was used for nucleus localization; a construct containing *Arabidopsis thaliana* remorin 1.3 (REM1.3) fused with mCherry (Marín, Thallmair & Ott 2012) was used for plasmalemma localization; and a construct with autophagy-related protein (ATG8CL) fused with mRFP (Dagdas *et al.* 2016) was used for autophagosome localization.

Agrobacteria-mediated transformation

Transient transformation of three to four week old *Nicotiana benthamiana* leaves was performed using *Agrobacterium tumefaciens* strain GV3101 carrying the constructs described above. Briefly, the cultures were grown overnight, pelleted, washed with LB medium, and re-pelleted. Afterwards, they were re-suspended in infiltration solution consisting of 10 mM MES, 10 mM $MgCl_2$, and 0.2 mM acetosyringone to A_{600} of 0.5. Agrobacteria cultures carrying the appropriate vector constructs were mixed in equal ratios. An equal volume of agrobacteria culture carrying the p19 silencing suppressor was added in all cases. Leaves of *N. benthamiana* plants were infiltrated with the agrobacteria mix using a syringe. The plants were kept at 21 ± 2 °C in growth chambers, with illumination at 70 $\mu Mm^{-2}s^{-1}$ (Osram L36W/77 lamp), a photoperiod of 16 h and relative humidity 70%.

Confocal imaging

Expression of fluorescent proteins was followed three and four days after agroinfiltration using two different laser scanning confocal microscopes. Leica TCS SP5 laser scanning confocal microscope mounted on a Leica DMI 6000 CS inverted microscope (Leica Microsystems) with a 63x objective with zoom factor 1 was used for localization studies. Sequential scanning was performed (sequential setting 1: 543 nm laser, emission window 555 – 637 nm, sequential setting 2: 488 nm and 514 nm lasers, emission windows 529 – 594 nm, 703 – 800 nm and transmission). Two to five regions of interest per one agroinfiltrated area were scanned unidirectionally with a zoom factor 1, line average 2 and scan speed 400 Hz. Leica TCS LSI laser scanning confocal microscope mounted on a Leica Z6 APOA microscope (Leica Microsystems) with an 20x objective was used for co-localization with ATG8 CL marker. The 488 nm and 532 nm lasers were used for the excitation of the enhanced yellow fluorescent protein (YFP), enhanced green fluorescent protein (GFP) and mRFP1, respectively. The YFP emission was measured in the window from 525 to 550 nm, the GFP emission in the window

from 505 to 525 nm and the mRFP1 in the window from 600 to 620 nm. For localization, three regions of interest per one agroinfiltrated area were scanned unidirectionally with zoom factor 3, line average 3 and scan speed 600 Hz. The overlay of images obtained in different channels and maximum projections from Z-stacks were performed using Leica LAS AF Lite software (Leica Microsystems).

Protein pull-down and co-immunoprecipitation assay

For protein pull-down, the YFP-tagged PoSTOSP28 effector protein was transiently expressed in *N. benthamiana* leaves for five days and the fluorescence confirmed with confocal microscopy (effector sample). Leaves infiltrated with p19 silencing suppressor only were used as control (control sample). Total proteins were extracted from ~200 mg leaf material with immunoprecipitation (IP) buffer, containing 25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM dithiotreitol, 0.1 mM PMSF, 0.02% NP-40, 10% glycerol and cOmplete™ ULTRA Tablets, Mini, EDTA-free Protease Inhibitor Cocktail (Roche Switzerland). Protein extraction was followed by 1-2 h incubation with GFP-Trap® Magnetic Agarose beads (ChromoTek, Germany) at 4 °C. The beads were washed three times with IP buffer without NP-40 and eluted into the SDS-PAGE loading buffer, containing 100 mM Tris-HCl, pH 6.8, 4% SDS, 0.2% bromophenol blue, 20% glycerol and 200 mM dithiothreitol. The immunoprecipitated proteins in the effector and control samples were then analyzed by SDS-PAGE, followed by tryptic digestion.

For co-immunoprecipitation assay, the YFP-tagged PoSTOSP28 was co-expressed with either of the HA-tagged phosphoglucomutase proteins in *N. benthamiana* leaves for four days. As control, the phosphoglucomutases were co-expressed with GFP, encoded by the pB7WGF2 vector (Karimi *et al.* 2002), while each of the interactors was also expressed separately. The fluorescence of YFP-tagged PoSTOSP28 or GFP was confirmed with confocal microscopy. About 500 mg leaf material was used for protein extraction and immunoprecipitation, performed as described above, with an additional dilution step in IP buffer without NP-40 following protein extraction (extract to buffer ratio 1:2.5). The immunoprecipitated proteins and protein extracts were analyzed by SDS-PAGE and Western blot, using anti-GFP (diluted 1:3.000, Invitrogen, USA) and anti-HA (diluted 1:1.000, ChromoTek, Germany) antibodies.

Protein sequencing

Bands of about 1 × 10 mm in size were excised from the PageBlue™ Protein Staining Solution (Thermo Scientific, USA) stained SDS-PAGE gel, at approximately the same heights from the effector and control sample lanes. In gel protein reduction, alkylation and tryptic digestion procedures were performed as described by Olsen *et al.* (2004). In short, gel bands were cut into 1 mm^2 pieces and de-stained. Following protein reduction with dithiothreitol (20 min at 56 °C) and alkylation with iodoacetamide (20 min at room temperature in dark), the samples were digested with trypsin (Sigma-Aldrich, USA) overnight at 37 °C. Gel protein digests were extracted with 30% acetonitrile and 1% trifluoroacetic acid. Samples were desalted with Bond Elut OMIX C18 tips (Agilent, USA) and dried down completely.

Gel protein digests were re-dissolved in 2% acetonitrile and 0.1% formic acid, ultrasonicated for 15 s, and centrifuged (10.000 x g, 5 min, 4 °C). An uHPLC system (Dionex Ultimate 3000, Thermo Fischer Scientific) with a flow rate of 300 µl/min was used. The column (Thermo Scientific Easy Spray column) was loaded with the protein digests and peptides were eluted with a 90-min gradient from 2% to 90% acetonitrile containing 0.1% formic acid. Mass spectrometry (MS) analysis was carried out using a Thermo Scientific Orbitrap Elite mass spectrometer (Thermo Scientific) using a data-dependent top 20 method dynamically choosing the most abundant precursor ions from the survey scan (400-1.800 m/z and a resolution of 30.000). Default charge state was set to 2-fold charge; unassigned charge states as well as +1 charge states were rejected. Minimal required signal was set to 1.000, size of exclusion mass list was set to 500 (with a duration of 60 s), and exclusion mass width was set to 4 ppm with one repeated count of 30 s.

Analysis of MS data was performed using MaxQuant (1.6.5.0). Raw files were searched against a FASTA file of the uniprot-proteome 7360 (AUP000084051 4097, *N. tabacum*, July2019). Tryptic peptides were allowed a maximum of 2 missed cleavages as well as a maximum of three modifications per peptide (oxidation M and N-terminal acetylation). Precursor mass tolerance was set to 4.5 ppm (FTMS) and 0.6 Da (ITMS). To eliminate matching by chance, data was searched against a database of revert sequences in a target-decoy approach. Only high confidence peptides (FDR < 0.01%), as well as proteins with at least two distinct identified peptides, passed the criteria for identification. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner

repository with the dataset identifier PXD015374. Target proteins were selected based on peptide identification in gel bands from effector sample, but not in the control sample. The grapevine orthologues of *N. tabacum* target proteins were identified with protein BLAST, using the VCost.v3 version of the 12X.v2 version of the grapevine genome assembly (Canaguier *et al.* 2017).

Enzymatic assays

Enzymatic assays were performed either in grapevine or tobacco (*N. benthamiana*) leaf material. Grapevine material was collected from grapevine cv. Zweigelt ('Rotburger') naturally infected with '*Ca.* P. solani' in the vineyard in Klosterneuburg (Austria). Transiently transformed *N. benthamiana* leaves were infiltrated with a mixture of agrobacteria, harboring either the effector expression cassette and RNA silencing suppressor p19 (effector expressing samples) or only p19 (control samples). Leaves of non-transformed plants were also analyzed (intact control samples). In transformed plants, the enzyme activity was determined in the agroinfiltrated leaf and in leaf above, here named systemic leaf. Leaves were sampled 3 and 14 days after agroinfiltration. Extraction of enzymes from plant leaf material was performed exactly as previously described (Jammer *et al.* 2015; Covington Dunn *et al.* 2016; Fimognari *et al.* 2020).

All enzymatic activity assays were performed in either UV-transmissive flat bottom 96-well plates (UV-Star Greiner Bio One, Kremsmünster, Austria). Protein extract volumes from 1 to up to 20 µl were used for the reactions. Total reaction volume was 160 µl. Reaction mixes were incubated in a plate reader (Ascent Multiskan; Thermo Fisher Scientific) for 40 min at 25 or 30°C according to the optimized protocol for each enzyme. All assays were carried out in triplicates and for control assays the substrate was not added to the reaction mixes. Change of absorbance per second was used to calculate activity of enzyme in nkat/g protein. Enzymatic activity assay of carbohydrate metabolism related enzymes was performed according to (Jammer *et al.* 2015). Antioxidant metabolism related enzymatic activity assay were performed according to (Fimognari *et al.* 2020).

Results

Prediction of candidate effector proteins

Bioinformatic analysis of ten genomic DNA sequences from '*Ca.* P. solani' revealed the presence of signal peptide cleavage sites characteristic of candidate effector proteins in 11 to 20 proteins per genomic phytoplasma DNA and the presence of transmembrane helices in proteins in 15 to 34 proteins per genomic phytoplasma DNA. Six of these proteins, that met the criteria of signal peptide cleavage site, presence of transmembrane helices, annotation, and unique expression *in planta*, were selected for further investigation as six putative effector proteins: PoSTOSP04, PoSTOSP06, PoSTOSP13, PoSTOSP14, PoSTOSP18, and PoSTOSP28 (Table 1).

Table 1. List of candidate effectors of '*Candidatus* Phytoplasma solani'.

Candidate effector name	Annotation	Signal cleaving	Number of predicted transmembrane helices	Number out of ten gDNA phytoplasma sequences, in which the putative effector is present	Length (amino acids)	Variation in length between phytoplasma genotypes	Amino acid sequence similarity, if a putative effector is present	Localization prediction
PoSTOSP04	Putative effector ('*Ca.* P. solani')	Mature chain: 33–258	1	8	260	53 amino acids missing in one of the genotypes	100 %	No predicted locations
PoSTOSP06	Putative effector ('*Ca.* P. solani')	Mature chain: 32–506	1	9	506	109 amino acids missing in one of the genotypes	97 – 99 %	Chloroplast, nucleus
PoSTOSP13	Putative effector ('*Ca.* P. solani')	Mature chain: 35–107	1	7	372	80 amino acids missing in two of the genotypes	93 %	No predicted locations
PoSTOSP14	Putative effector ('*Ca.* P. solani')	Mature chain: 34–629	2	6	630	80 amino acids missing in two of the genotypes	96 –100 %	Nucleus
PoSTOSP18	ABC-type sugar transport system, substrate binding protein ('*Ca.* P. solani')	Mature chain: 32–540	1	10	540		99 %	No predicted locations
PoSTOSP28	Antigenic membrane protein ('*Ca.* P. solani')	Mature chain: 32–157	2	10	159	66 amino acids missing in one of the genotypes and 26 additional amino acids is missing in another genotype	91 – 100 %	No predicted locations

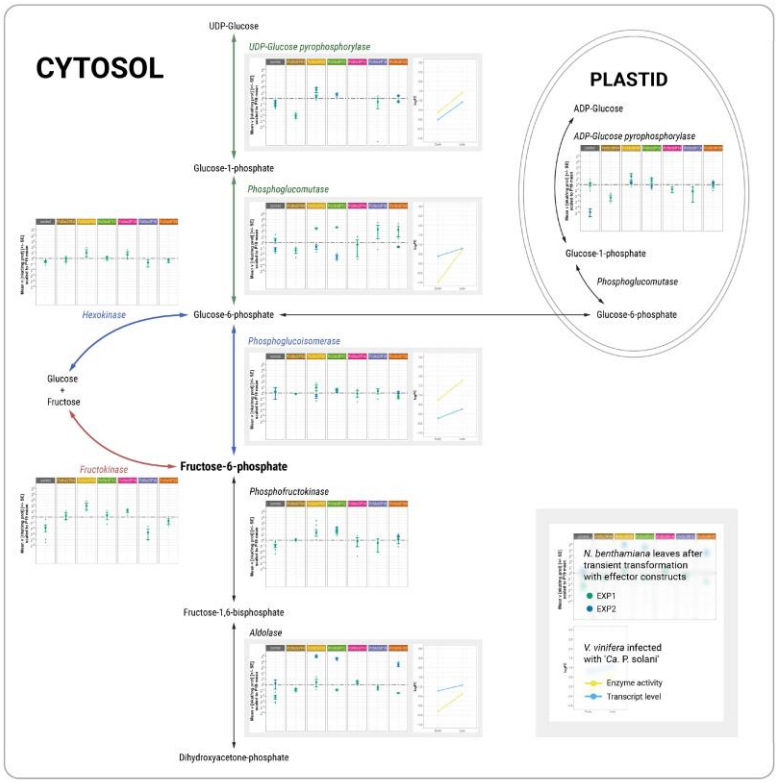
PoSTOSP18 and PoSTOSP28 have known functions

Although the genomic sequences of the putative effectors were revealed here according to the predicted effector criteria, i.e. presence and location of signal peptide cleavage sites and transmembrane helices, two of them have been annotated previously. PoSTOSP18 was annotated as a substrate-binding protein within the ABC-type sugar transport system (Table 1). PoSTOSP28, on the other hand, is 91-100% identical to an antigenic membrane protein StAMP of '*Ca.* P. solani' (Fabre, Danet & Foissac 2011). It encodes a 157 amino acid long protein with a predicted signal peptide cleavage site at amino acid 32 and a C-terminal hydrophobic α-helix detected in eight of ten genotypes examined (Table 1).

Toward deciphering the mechanisms of '*Ca.* P. solani' manipulation of carbohydrate metabolism

Hypothesizing that effectors of '*Ca.* P. solani' are directly involved in the altered carbohydrate metabolism previously demonstrated in '*Ca.* P. solani', several key enzymes of the carbohydrate metabolism were tested in leaves of *N. benthamiana* after Agrobacterium-mediated transient transformation with each effector construct. The enzyme activity was determined in the leaf, agroinfiltrated with the effector construct, and in a systemic leaf, 3 and 14 days after agroinfiltration. In general, there were no significant differences in the activity determined in the control and inoculated leaves 3 days after transformation. There were also no differences in activity between the control and systemic leaves, indicating that the effectors did not move away from the inoculation site, probably due to the size of the protein fused with YFP (see section on subcellular localization). However, in leaves inoculated with some of the effector constructs, differences in enzyme activities were detected after 14 days compared with control (Fig. 1). The increased activity of UDP-glucose pyrophosphorylase was associated with PoSTOSP06 and PoSTOSP13; PoSTOSP18 and PoSTOSP28 had effects on phosphoglucomutase activity; and PoSTOSP04, PoSTOSP06, PoSTOSP13, and slightly PoSTOSP28 on phosphofructokinase activity. The activity of aldolase was affected by PoSTOSP06, PoSTOSP13, and PoSTOSP28. In addition, the higher activities of hexokinase and fructokinase compared with the control were associated with PoSTOSP06, PoSTOSP13, and PoSTOSP14 and PoSTOSP04, PoSTOSP06, PoSTOSP13, and PoSTOSP14, respectively. A minor effect on ADP-glucose pyrophosphorylase activity was found in *N. benthamiana* leaves producing effectors PoSTOSP06 and PoSTOSP13.

309 In symptomatic grapevines cv. Zweigelt infected with '*Ca.* P. solani' compared to uninfected
310 grapevines, the metabolic pathways associated with the production of phosphorylated sugars
311 were induced both at the transcriptional level (i.e. *Vitvi04g01633* encoding UDP-glucose
312 pyrophosphorylase, *Vitvi01g00455* encoding phosphoglucomutase, *Vitvi18g00504* encoding
313 phosphoglucoisomerase, *Vitvi14g01938* encoding phosphofructokinase, and *Vitvi09g01500*
314 encoding aldolase) and at the level of activity of the corresponding enzymes (Fig. 1,
315 Supplemental Table 2).



316

317 **Figure 1. Activities of key enzymes involved in carbohydrate metabolism associated with sugar**
318 **phosphorylation pathways.** Enzyme activities are shown after transient transformation of *N.*
319 *benthamiana* with different effector constructs in inoculated leaves 14 days after transformation.
320 Enzyme activity levels were scaled using the corresponding mean value of the RNA silencing
321 suppressor p19. Scaled values, mean values and standard errors of the mean are shown for two
322 experiments. The grapevine enzyme activities together with the transcription level of corresponding
323 genes are shown as $\log_2$ fold change values given by the differences between grapevines infected and
324 uninfected with '*Ca*. P. solani' before (early) and after (late) symptom development. Transcription
325 values correspond to the following genes (Supplemental Table 2): *Vitvi04g01633* for UDP-glucose
326 pyrophosphorylase, *Vitvi01g00455* for phosphoglucomutase, *Vitvi18g00504* for
327 phosphoglucoisomerase, *Vitvi09g01500* for aldolase.

Effector PoSTOSP28 interacts with grapevine phosphoglucomutase enzymes

Protein partner determination is an important step in understanding protein function and identifying relevant biological pathways. We used a pull-down assay to detect possible physical interaction(s) between PoSTOSP28 and proteins from *N. benthamiana*. We identified potential interactors from 58 protein groups. These proteins were involved in various processes, including primary and secondary metabolism, redox regulation, photosynthesis, protein degradation, and processing (Supplemental Table 3) (Perez-Riverol et al., 2019). For further analysis, we selected phosphoglucomutase (PGM, UniProtKB identifiers A0A1S3XHG4, A0A1S4AIH6, and A0A1S4A6S4), which was associated with increased enzyme activity in *N. benthamiana* leaves after transient transformation with PoSTOSP28 (Fig. 1). Using an *in planta* co-immunoprecipitation assay of YFP-tagged PoSTOSP28 and HA-tagged *N. tabacum* phosphoglucomutase orthologues in grapevine *Vitvi16g00891* (VvPGM1) and *Vitvi01g00455* (VvPGM2) in *N. benthamiana* (Fig. 2, Supplemental Fig. 1), we confirmed that PoSTOSP28 interacts with both grapevine phosphoglucomutases.

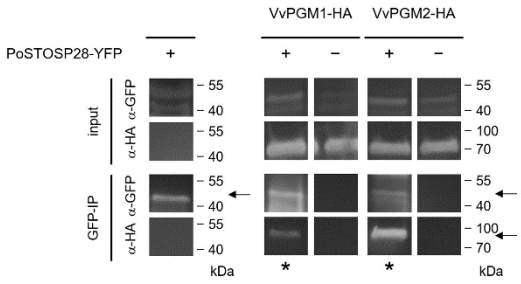
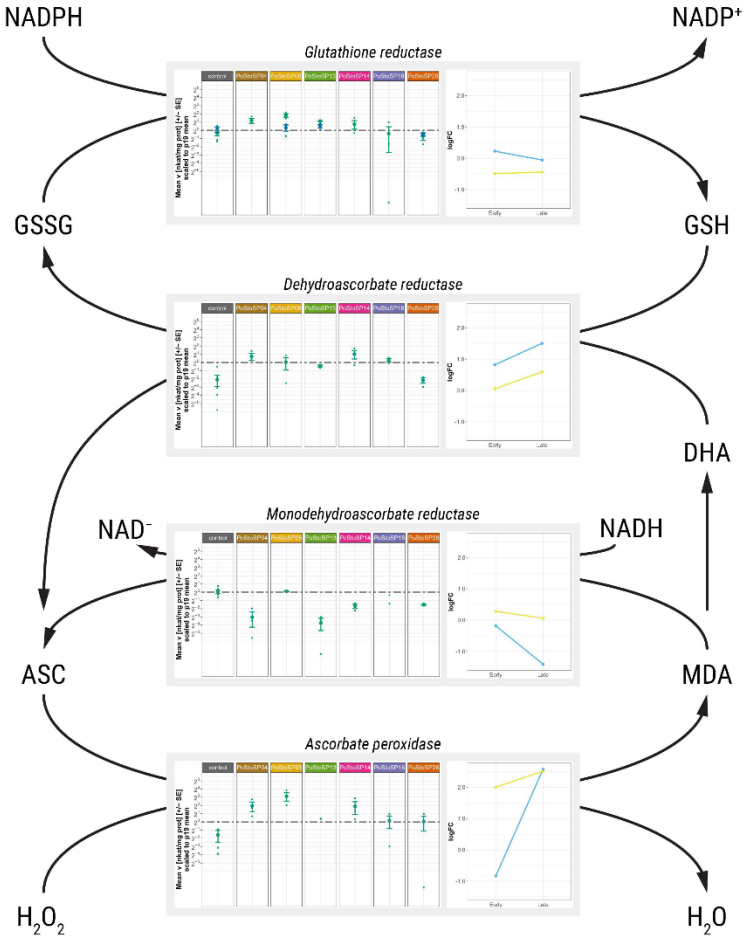


Figure 2. PoSTOSP28 interacts with grapevine phosphoglucomutases *in planta*. Co-immunoprecipitation assay results, showing the interaction between PoSTOSP28 and VvPGM1 and VvPGM2. The combination of YFP- and HA-tagged proteins expressed in *N. benthamiana* is indicated for each sample (+/−). Positive interactions were determined by detection of immunoprecipitated (GFP-IP) complexes with anti-HA antibodies (asterisks). Detection of YFP-tagged PoSTOSP28 with anti-GFP antibodies in GFP-IP samples and detection of proteins with anti-GFP and anti-HA antibodies in leaf protein extracts (input) are shown as controls. Negative controls for GFP are shown in Supplemental Fig. 1. The arrows indicate expected bands.

Association of '*Ca.* P. solani' infection with oxidative stress

It has been previously shown that transcripts of some enzymes related to oxidative stress in grapevine cv. Zweigelt infected with '*Ca.* P. solani' are increased (Dermastia et al., 2021). Here, we show that transient transformation of *N. benthamiana* leaves with the candidate effector constructs affects the activity of these enzymes. Transformation with all candidate effectors tested was associated with an increase in ascorbate peroxidase activity, which was also strongly induced in grapevine late in the growing season (Fig. 3). Interestingly, monodehydroascorbate reductase activity was lower in all effectors except PoSTOSP06 compared with control, which was associated with a decrease in the transcript of the two grapevine monodehydroascorbate reductase genes *Vitvi08g01483* and *Vitvi14g01751* (Fig. 3, Supplemental Table 2) and lower activity late in the growing season (Fig. 3). A transcript of *Vitvi13g00241* (Fig. 3, Supplemental Table 2), along with the activity of the associated dehydroascorbate reductase, was increased in grapevine late in the growing season, and this induction was consistent with increased activity in *N. benthamiana* leaves after transformation by all effectors tested, except PoSTOSP28 (Fig. 3). Similar effects of transformation with effector constructs were detected on glutathione reductase activity. In symptomatic grapevine plants late in the growing season the transcript level of *Vitvi07g00037* that encodes glutathione reductase was slightly lower in comparison with the uninfected grapevines (Fig. 3, Supplemental Table 2), but the activity of the corresponding enzyme remains similar (Fig. 3). In addition, the activity of *N. benthamiana* catalase was slightly increased by all effector constructs tested, similar to the activity in grapevine.


NADPH
NADP+
Glutathione reductase
GSSG
GSH
Dehydroascorbate reductase
DHA
NAD-
Monodehydroascorbate reductase
NADH
ASC
MDA
Ascorbate peroxidase
H2O2
H2O
N. benthamiana leaves after transient transformation with effector constructs
EXP1
EXP2
V. vinifera infected with 'Ca. P. solani'
Enzyme activity
Transcript level

374 **Figure 3. Activities of key enzymes involved in ascorbate-glutathione cycle.** Enzyme activities are
375 shown after transient transformation of *N. benthamiana* with different effector constructs in inoculated
376 leaves 14 days after transformation. Enzyme activity levels were scaled using the corresponding mean
377 value of the RNA silencing suppressor p19. Scaled values, mean values and standard errors of the
378 mean are shown for two experiments. The grapevine enzyme activities together with the transcription
379 level of corresponding genes are shown as $\log_2$ fold change values given by the differences between
380 grapevines infected and uninfected with '*Ca.* P. solani' before (early) and after (late) symptom
381 development. Transcription values correspond to the following genes (Supplemental Table 2):
382 *Vitvi04g02166* for ascorbate peroxidase, *Vitvi08g01483* for monodehydroascorbate reductase,
383 *Vitvi13g00241* for dehydroascorbate reductase, *Vitvi07g00037* for glutathione reductase.

384 **Subcellular localization of the candidate effector proteins**

385 To determine the subcellular localization of effector proteins, especially those with the highest
386 impact on enzymes involved in the phosphorylation of sugars (i.e. PoSTOSP06 ,
387 PoSTOSP013, and PoSTOSP28), Agrobacterium-mediated transient transformation of *N.*
388 *benthamiana* was performed with YFP-tagged effector sequences, histone 2B (H2B) tagged
389 with monomeric red fluorescent protein 1 (H2B-mRFP1) and *Arabidopsis thaliana* remorin
390 1.3 (REM1.3) fused with mCherry (REM1.3-mRFP1). Their expression showed that all
391 proteins examined were localized in the nucleus and cytoplasm (Fig. 4, Supplemental Fig. 2).
392 Out of all the effector proteins, PoSTOSP28 showed the highest expression in the thread-like
393 structures spanning the cells. Additionally, although it was also present in the nucleus, the
394 level of expression was usually lower compared with other effectors, accompanied by a
395 distinct edge at the periphery or just outside of the nucleus (Fig. 4, Supplemental Fig. 2).

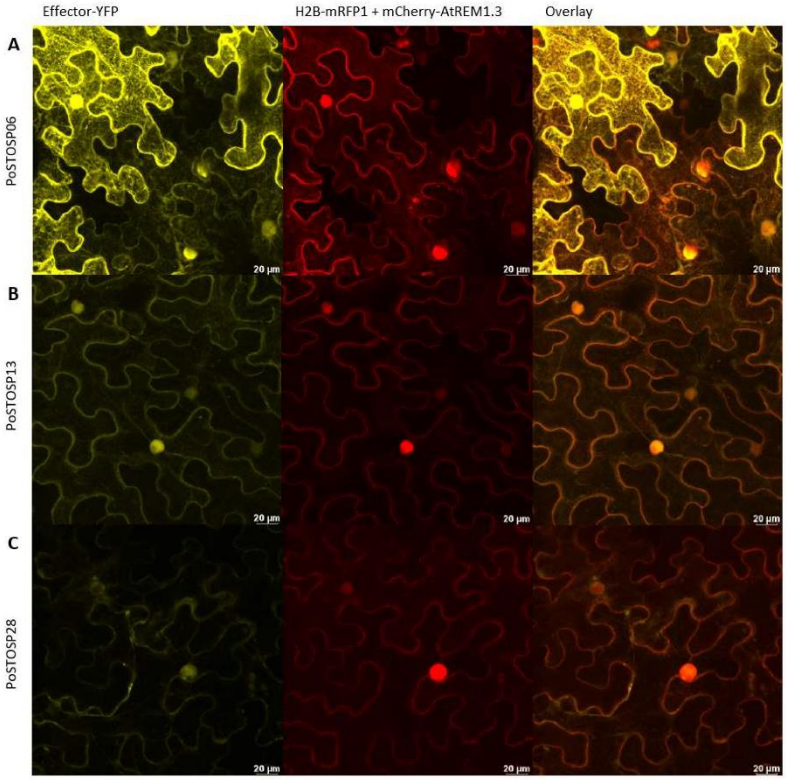
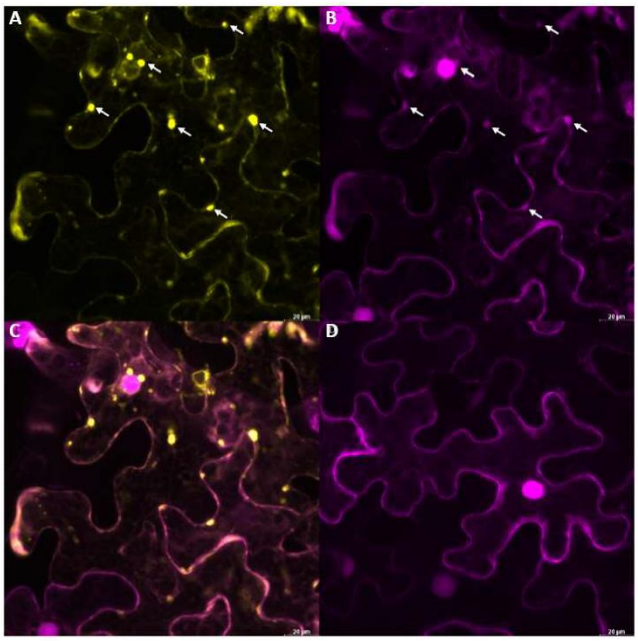


Figure 4. Subcellular localization of effectors PoSTOSP06 (A), PoSTOSP013 (B), and PoSTOSP28 (C). From left to right: effector tagged with YFP (effector-YFP), nuclear marker histone H2B tagged with mRFP (H2B RFP) and plasmalemma marker *A. thaliana* remorin 1.3 tagged with mCherry (mCherry-AtREM1.3), overlay of effector and organelle markers. Scale bars 20 µm.

401 **Effector PoSTOSP28 induces autophagosomes in *N. benthamiana***

402 On closer inspection, some round structures were found in the PoSTOSP28 effector-
403 expressing cells. To identify the round structures, the effector of interest was expressed
404 together with the mRFP1-tagged ATG8 CL autophagosome marker. The marker showed that
405 the observed structures were frequently co-localized with autophagosomes (Fig. 5). Moreover,
406 almost no autophagosomes were observed in the control plants expressing only the mRFP1-
407 tagged ATG8 CL autophagosome marker, suggesting that the effector PoSTOSP28 is not only
408 localized in the autophagosome but can also increase the occurrence of autophagosomes.



409

410 **Figure 5. Potential effector PoSTOSP28 induces autophagy in inoculated cells.** The YFP-tagged
411 effector PoSTOSP28 co-expressed with the mRFP1-tagged ATG8 CL autophagosome marker. White
412 arrows indicate co-localization of effector and autophagosome markers. (A) YFP fluorescence derived
413 from YFP-tagged effector PoSTOSP28 (yellow), (B) mRFP1-tagged nuclear marker histone H2B and
414 mRFP1-tagged ATG8 CL marker fluorescence (magenta), (C) overlay of images, (D) cell expressing
415 only mRFP1-tagged ATG8 CL. Scale bars 20 µm.

Discussion

Candidate effectors are phytoplasma species and strain specific

Candidate effectors from *Ca.* P. solani' strain BN601 originated from grapevine cv. Zweigelt were compared with effectors from the '*Ca.* P. solani' strain SA-1 derived from infected grapevines and transferred to and maintained in Madagascar periwinkle (Music Seruga *et al.* 2019). PoStoSP04 is identical to PSSA1_v1c2050. PoStoSP13 showed 45 % identity with PSSA1_v1c4770. The sequence of PoStoSP14 includes the sequence of PSSA1_v1c5140. This sequence comprises only one third of the PoStoSP14 sequence and shows 92 % identity with it (Supplemental Fig. 3). Considering the broad plant host range and polyphagous insect vectors of '*Ca.* P. solani', the presence of a significant number of strain-specific genes within the species has already been suggested (Music Seruga *et al.* 2019).

PoStoSP06, PoStoSP18, and PoStoSP28 are not included in the effector list of the strain SA-1 (Music Seruga *et al.* 2019). A gene encoding PoStoSP28 is present in the '*Ca.* P. solani' genome annotated as *stamp*, with a suggested function in phytoplasma-insect interaction (Fabre *et al.* 2011). The gene encoding StAMP is highly variable, as evidenced in several studies on grapevines infected with '*Ca.* P. solani' from different countries (Cvrković, Jović, Mitrović, Krstić & Toševski 2014; Murolo & Romanazzi 2015; Atanasova *et al.* 2015; Delić *et al.* 2016; Quaglino *et al.* 2021; Mehle *et al.* 2022). This variability suggests that StAMP is subject to positive diversifying selection pressure, probably related to its interaction with the insect vector (Fabre *et al.* 2011). StAMP has a similar structure to the cell-surface phytoplasma membrane antigen membrane proteins (AMP) of '*Ca.* P. asteris', which form a complex with microfilaments of insect cells only in the viscera and salivary glands of insect vectors (Suzuki *et al.* 2006; Galetto *et al.* 2011). A similar structure may indicate a similar function as AMP of '*Ca.* P. asteris' (Kakizawa *et al.* 2004). Since phytoplasmas can survive only in their insect vectors or in the phloem of their host plants, the possible biological role of StAMP as a virulence factor in plant pathogenesis was investigated here along with other identified candidate effectors.

On the other hand, several substrate-binding proteins associated with the ATP-binding cassette (ABC) transporters that are secreted through the Sec-dependent pathway have been characterized in '*Ca.* P. asteris' strain AY-WB (Bai *et al.* 2009) and are also present in '*Ca.* P.

446 solani' strain SA-1 (Music Seruga *et al.* 2019). In this study, we examined a substrate-binding
447 protein of ABC-type sugar transport system as a candidate effector PoStoSP18. Although
448 plant pathogenic bacteria encode numerous ABC transporters that enable them to import a
449 greater number of molecules compared to saprophytes and animal pathogens, they have been
450 little examined in plant pathogens (Zeng & Charkowski 2021).

451 Phytoplasma genomes include potential mobile units, putative transposon-like elements, that
452 are thought to contribute to the genome instability observed in these bacteria (Bai *et al.* 2009).
453 They are also present in the '*Ca.* P. solani' strain SA-1 genome (Music Seruga *et al.* 2019).
454 These potential mobile units and DNA elements often contain genes for candidate virulence
455 proteins effectors (Bai *et al.* 2009). At least twenty of the proteins annotated as putative
456 effectors in '*Ca.* P. solani' strain SA-1 genome are located within these regions (Music
457 Seruga *et al.* 2019). However, a comparison of candidate effectors studied in this work with
458 putative effectors from the strain SA-1 indicate that only PoStoSP14, which is related to
459 PSSA1_v1c5140, might be associated with the potential mobile units. Among the candidate
460 effectors of SA-1, PSSA1_v1c5140 is shown to be unique to that strain (Music Seruga *et al.*
461 2019).

462 **Modulation of phosphorylated sugars by the candidate effectors**

463 There is growing evidence that infection with '*Ca.* P. solani' affects carbohydrate metabolism
464 in grapevine, probably leading to symptoms such as leaf curling, yellowing or reddening of
465 leaf margins, and hard, brittle texture of leaves (Dermastia *et al.* 2017). However, it is not
466 always easy to distinguish whether the primary event is the generation of a metabolic sugar
467 signal or a direct effect of a corresponding pathogen signal (Prezelj *et al.* 2016a). By directly
468 tracking the consequences of transient transformation of *N. benthamiana* with putative
469 phytoplasma effectors on the host plant, this study at least partially resolves a dilemma.
470 Although the molecular mechanism behind the effect of the analyzed effector candidates on
471 plant carbohydrate metabolism remains unclear, the localization of the effector in plant nuclei
472 suggests a possible perturbation of this metabolism at the transcriptional level (Fig. 4), besides
473 direct interactions with enzymes, as shown for phosphoglucomutase (Fig. 2).

474 The general upper part of glycolysis is encoded in all known phytoplasma genomes. The
475 genes encoding the upper part of glycolysis (i.e. phosphoglucose isomerase,
476 phosphofructokinase, aldolase) were found in two strains of '*Ca.* P. solani' originated from

477 tobacco plants as well in SA-1 strain originated from grapevine (Mitrović *et al.* 2014; Music
478 Seruga *et al.* 2019). However, phytoplasmas lack the hexokinase and sugar-specific
479 phosphotransferase systems to mediate the entry of phosphorylated hexose into glycolysis
480 (Kube *et al.* 2012), which is the first step in the breakdown of glucose to produce energy in
481 the form of ATP. Our previous metabolome analysis showed that the amount of fructose-6-
482 phosphate and other hexose-6-phosphates in the middle leaf veins, containing the phloem sap
483 from grapevines infected with '*Ca.* P. solani', was 2.7- to 2.9-fold higher than in uninfected
484 grapevines (Prezelj *et al.*, 2016). This finding confirmed the suggestion that there should be a
485 yet unclear phytoplasma pathway for importing sugars and generating the hexose-6-phosphate
486 to support glycolysis (Tan *et al.* 2021). The results of this study show for the first time that
487 candidate phytoplasma effector proteins may be involved in the manipulation of plant
488 metabolism through these processes. Moreover, they suggest that effector proteins may even
489 mediate more than one pathway that could supply phytoplasmas with phosphorylated hexoses
490 (Fig. 1). Although the molecular mechanism of how the candidate effectors of '*Ca.* P. solani'
491 alter plant metabolism to increase the activity of specific enzymes in the glycolytic pathway is
492 currently unknown, it appears that the coordinated action of multiple candidate effectors is
493 involved (Fig. 1). This is consistent with the finding that several effectors of '*Ca.* P. asteris'
494 AY-WB have at least two interactions with host proteins and some of them with more than
495 240 host proteins (Marrero *et al.* 2020).

496 The first possible pathway for providing '*Ca.* P. solani' with fructose-6-phosphate includes
497 the activation of the grapevine hexokinase system by the candidate effectors (Fig. 1 - blue
498 pathway) and further conversion of the resulting glucose-6-phosphate by phytoplasma-
499 encoded phosphoglucose isomerase (Kube *et al.* 2012; Music Seruga *et al.* 2019). Fructose
500 can also be directly converted to fructose-6-phosphate by the action of fructokinase, shown to
501 be affected by all but PoSTOSP18 and PoSTOSP28 effector constructs (Fig. 1 - red pathway).
502 Another alternative pathway leading to fructose-6-phosphate includes induction of the
503 enzyme UDP-glucose pyrophosphorylase leading to glucose-1-phosphate, followed by
504 conversion of the resulting glucose-1-phosphate to glucose-6-phosphate by
505 phosphoglucomutase (Fig. 1 – green pathway). Its activity was induced in transformed *N.*
506 *benthamiana* by PoSTOSP18 and PoSTOSP28 (Fig. 1 – green pathway), and the product may
507 be a substrate for phosphoglucoisomerase or for ADP -glucose pyrophosphorylase involved in
508 starch biosynthesis (Fig. 1 - black pathway). Moreover, the results suggest a role of

509 PoSTOSP28 in the pathogenicity of phytoplasma in grapevine by interacting with grapevine
510 phosphoglucomutase enzymes (Fig. 2).

511 It is worth noting that in this analysis PoSTOSP18, as a putative substrate-binding protein
512 within the ABC-type sugar transport system, was associated only with the increase in
513 phosphoglucomutase activity (Fig. 1 – green pathway). At the moment it is not known if
514 PoSTOSP18 binds to any of the phytoplasma ABC transporters present in phytoplasma
515 genomes (Bai *et al.* 2009; Music Seruga *et al.* 2019) or to the host ABC transporters, which
516 are particularly abundant in plants and are involved in many essential physiological processes
517 as well as responses to biotic and abiotic stresses and interactions with the environment
518 (Lefèvre & Boutry 2018). Since some ABC transporters can transport many structurally
519 unrelated substances, the list of their possible substrates is far from clear (Lefèvre & Boutry
520 2018).

521 The conversion of fructose-6-phosphate to fructose-1,6-bisphosphate by phosphofructokinase
522 and of fructose-1,6-bisphosphate to dihydroxyacetone phosphate by aldolase was induced by
523 PoSTOSP06 and PoSTOSP13. The latter was also affected by PoSTOSP28 and was
524 significantly higher at transcript level and enzyme activity in infected grapevines (Fig. 1),
525 indicating a potential regulation of its activity through transcriptional regulation.

526 **Candidate effectors induce ascorbate-glutathione cycle**

527 Conditions that exceed the sensory or pathogen defense capacities of plant cells disrupt the
528 balance between oxidants and antioxidants in favor of oxidants, resulting in disruption of
529 redox signaling and control and/or molecular damage (Sies, Berndt & Jones 2017). As a
530 result, plants generate large numbers of reactive oxygen species (ROS) in cells. To avoid
531 metabolic disorders, plants must have systems for ROS elimination. On the other hand, since
532 ROS are a common feature of plant defenses, it should also be advantageous for any pathogen
533 to be able to block the effects of ROS. A common first line of defense against ROS is the use
534 of antioxidant enzymes (Fones & Preston 2012). The most important antioxidant system for
535 the efficient removal of ROS and maintenance of cellular homeostasis is the ascorbate-
536 glutathione cycle with three interdependent redox pairs: ascorbate/dehydroascorbate, reduced
537 form of glutathione (GSH)/glutathione disulfide (GSSG), NADPH/NADP, and three
538 enzymes: ascorbate peroxidase, monodehydroascorbate reductase, dehydroascorbate
539 reductase, and glutathione reductase. Remarkably, the induction of the entire ascorbate-

glutathione cycle in infected grapevines cv. Zweigelt toward the end of the growing season correlated positively with the induced enzyme activity in *N. benthamiana* leaves after transient transformation with different candidate effectors (Fig. 3). It appears that all effector candidates tested contribute to the induced activity of ascorbate peroxide after transient transformation of *N. benthamiana* leaves, with a very low contribution from PoSTOSP18 and PoSTOSP28 (Fig. 3). Interestingly, PoSTOSP04, which was the only candidate effector identical to one found in the strain SA-1 (Music Seruga *et al.* 2019), contributed only to the enzymes associated with oxidative stress.

The possible involvement of candidate effectors in induced autophagosome formation in '*Ca.* P. solani' pathogenicity

An interesting observation in this study was the co-localization of autophagosomes with the effector candidate PoSTOSP28 (Fig. 5), after transient transformation of *N. benthamiana* with effector constructs. The formation of these *de novo* double-membrane vesicles is part of an evolutionarily conserved degradation process of macroautophagy, which is required for the maintenance of cellular homeostasis. Plants use autophagosomes to engulf damaged organelles, nonfunctional proteins, and pathogens. Autophagosomes then fuse with the plant vacuole, where their contents are degraded (Wun, Quan & Zhuang 2020). Plant pathogens have evolved ways to bypass or modulate plant autophagy and use it to their own advantage (Leary *et al.* 2018). Macroautophagy is divided into several steps involving different autophagy-related proteins (ATGs). Some of them initiate the process, while ATG9 is required for the final formation of autophagosomes (Feng & Klionsky 2017). In grapevine cv. Zweigelt, several autophagosome-related genes were upregulated in plants infected with '*Ca.* P. solani', especially late in the growth period (Supplemental Table 2). Of note are the increased transcript levels of Atg18 (i.e. *Vitvi07g00210*, *Vitvi14g00317*, *Vitvi11g00813*, *Vitvi19g01985*) and *Vitvi03g00492*, which encodes ATG2. It has been suggested that the ATG2-ATG18 complex binds the pre-autophagosomal structure to ER to initiate expansion during autophagosome formation (Kotani, Kirisako, Koizumi, Ohsumi & Nakatogawa 2018). In addition, the upregulated Atg9 gene *Vitvi07g00580* may be involved in further autophagosome expansion (Feng & Klionsky 2017). Transcriptome analysis also revealed upregulation of *Vitvi16g01309* and *Vitvi04g01617*, which encode ATG13, shown to bind ATG9 (Suzuki *et al.* 2015). The mechanism linking these data to the induction of autophagosomes in transformed *N. benthamiana* leaves is not known. However, there are

reports of mammalian systems in which impaired regulation of redox signaling can result in autophagic activity leading to a variety of diseases (Yun *et al.* 2020). Due to the low activity of the ascorbate-glutathione cycle in *N. benthamiana* leaves transformed with the PoSTOSP28 construct, its role in autophagy is less likely in oxidative stress.

Conclusions

This is the first comprehensive study on the possible physiological roles of candidate effectors or virulence factors of the plant pathogen '*Ca.* P. solani'. Bioinformatic search for their sequences in the genomes of phytoplasmas isolated from infected grapevines revealed several candidates. One of them was identical to and some of them resembled those already found in '*Ca.* P. solani' strain SA-1. Two others already suspected of playing a role in phytoplasmas, were added to the list. According to the hypothesis that phytoplasmas obtain most nutrients from the plant and therefore compete with the host for similar or identical nutrient substrates, it is hypothesized that a minor change in metabolism could significantly affect the outcome of pathogen-host interactions in providing nutrients, energy, and metabolites for successful replication and infection in plants. Since carbohydrate metabolism in plants has been shown to change upon phytoplasma infection, this study confirms the modulation of this metabolism by potential effectors. Moreover, most of the effectors tested contributed to the enzyme activities of the ascorbate-glutathione cycle. In addition, an effector showing some similarities with the effector of the strain SA-1 and the effector previously shown to interact with insect vectors as StAMP were associated with the induction of autophagosomes in plant cells. These results confirm that phytoplasmas interfere with plant metabolism and pave the way for further elucidation of the molecular mechanisms involved in these interactions.

Acknowledgements

The authors thank Prof. Dr. Jim Haseloff and Dr. Fernán Federici (University of Cambridge, UK) for providing the plasmid containing H2B-RFP, Alexandre Leary and Tolga Bozkurt (Imperial College London, UK) for the autophagosome marker, and Claudia Ribas and Brina Dragar for technical support.

599 **References**

600 Abramovitch R.B., Anderson J.C. & Martin G.B. (2006) Bacterial elicitation and evasion of
601 plant innate immunity. *Nature reviews. Molecular cell biology* **7**, 601–611.
602 Anabestani A., Izadpanah K., Abbà S., Galetto L., Ghorbani A., Palmano S., … Marzachì C.
603 (2017) Identification of putative effector genes and their transcripts in three strains
604 related to '*Candidatus* Phytoplasma aurantifolia.' *Microbiological Research* **199**, 57–66.
605 Atanasova B., Jakovljević M., Spasov D., Jović J., Mitrović M., Toševski I. & Cvrković T.
606 (2015) The molecular epidemiology of bois noir grapevine yellows caused by
607 '*Candidatus* Phytoplasma solani' in the Republic of Macedonia. *European Journal of*
608 *Plant Pathology* **142**, 759–770.
609 Bai X., Correa V.R., Toruño T.Y., Ammar E.D., Kamoun S. & Hogenhout S.A. (2009) AY-
610 WB phytoplasma secretes a protein that targets plant cell nuclei. *Molecular Plant-*
611 *Microbe Interactions* **22**, 18–30.
612 Brader G., Riedle-Bauer M., Aryan A., Kirchmaier S., Wieland L., Schönhuber C. & Fuchs F.
613 (2017) Zum Auftreten von Candidatus Phytoplasma solani an Kartoffelbeständen in
614 Niederösterreich. In *58th Österreichische Pflanzenschutztage*. St. Pölten, Austria.
615 Çağlar B.K., Elbeaino T., Küsek M., Pehlivan D., Fidan H. & Portakaldali M. (2010) Stolbur
616 phytoplasma infections in potato and tomato plants from different locations in Turkey.
617 *THE JOURNAL OF TURKISH PHYTOPATHOLOGY* **39**, 1–8.
618 Canaguier A., Grimplet J., Di Gaspero G., Scalabrin S., Duchêne E., Choisne N., … Adam-
619 Blondon A.F. (2017) A new version of the grapevine reference genome assembly
620 (12X.v2) and of its annotation (VCost.v3). *Genomics data* **14**, 56–62.
621 Contaldo N., Bertaccini A., Paltrinieri S., Windsor H.M. & Windsor G.D. (2012) Axenic
622 culture of plant pathogenic phytoplasmas. *Phytopathologia Mediterranea* **51**, 607–617.
623 Contaldo N., Satta E., Zambon Y., Paltrinieri S. & Bertaccini A. (2016) Development and
624 evaluation of different complex media for phytoplasma isolation and growth. *Journal of*
625 *microbiological methods* **127**, 105–110.
626 Covington Dunn E., Roitsch T., Dermastia M., Covington E.D., Roitsch T. & Dermastia M.
627 (2016) Determination of the activity signature of key carbohydrate metabolism enzymes
628 in phenolic-rich grapevine tissues. *Acta Chimica Slovenica* **63**, 757–762.
629 Cvrković T., Jović J., Mitrović M., Krstić O. & Toševski I. (2014) Experimental and
630 molecular evidence of *Reptalus panzeri* as a natural vector of bois noir. *Plant Pathology*

631 **63**, 42–53.
632 Dagdas Y.F., Beihaj K., Maqbool A., Chaparro-Garcia A., Pandey P., Petre B., … Bozkurt
633 T.O. (2016) An effector of the irish potato famine pathogen antagonizes a host
634 autophagy cargo receptor. *eLife* **5**.
635 Delić D., Balech B., Radulović M., Lolić B., Karačić A., Vukosavljević V., … Cvetković T.J.
636 (2016) Vmp1 and stamp genes variability of `Candidatus phytoplasma solani' in Bosnian
637 and Herzegovinian grapevine. *European Journal of Plant Pathology* **145**, 221–225.
638 Dermastia M. (2017) Interactions between grapevines and grapevine yellows phytoplasmas
639 BN and FD. In *Grapevine Yellows Diseases and Their Phytoplasma Agents*. (eds M.
640 Dermastia, A. Bertaccini, F. Constable & N. Mehle), pp. 47–67. SpringerBriefs in
641 Agriculture. Springer, Cham.
642 Dermastia M. (2019) Plant hormones in phytoplasma infected plants. *Frontiers in Plant*
643 *Science l* **10**, 477.
644 Dermastia M., Bertaccini A., Constable F. & Mehle N. (2017) *Grapevine yellows diseases*
645 *and their phytoplasma agents: biology and detection*, 1st ed. (eds M. Dermastia, A.
646 Bertaccini, F. Constable & N. Mehle), Springer, Singapore.
647 Dermastia M., Kube M. & Šeruga-Musić M. (2019) Transcriptomic and proteomic studies of
648 phytoplasma-infected plants. In *Phytoplasmas: Plant Pathogenic Bacteria - III*. pp. 35–
649 55. Springer Singapore, Singapore.
650 Dermastia M., Škrlj B., Strah R., Anžič B., Tomaž Š., Križnik M., … Pompe-novak M.
651 (2021) Differential response of grapevine to infection with 'candidatus phytoplasma
652 solani' in early and late growing season through complex regulation of mRNA and small
653 RNA transcriptomes. *International Journal of Molecular Sciences* **22**, 1–28.
654 Ember I., Acs Z., Munyaneza J.E., Crosslin J.M. & Kolber M. (2011) Survey and molecular
655 detection of phytoplasmas associated with potato in Romania and southern Russia.
656 *European Journal of Plant Pathology* **130**, 367–377.
657 Fabre A., Danet J.-L. & Foissac X. (2011) The stolbur phytoplasma antigenic membrane
658 protein gene *stamp* is submitted to diversifying positive selection. *Gene* **472**, 37–41.
659 Feng Y. & Klionsky D.J. (2017) Autophagic membrane delivery through ATG9. *Cell*
660 *Research* **27**, 161.
661 Fimognari L., Dölker R., Kaselyte G., Jensen C.N.G., Akhtar S.S., Großkinsky D.K. &
662 Roitsch T. (2020) Simple semi-high throughput determination of activity signatures of
663 key antioxidant enzymes for physiological phenotyping. *Plant Methods* **16**, 42.

664 Fones H. & Preston G.M. (2012) Reactive oxygen and oxidative stress tolerance in plant
665 pathogenic Pseudomonas. *FEMS microbiology letters* **327**, 1–8.
666 Galetto L., Bosco D., Balestrini R., Genre A., Fletcher J. & Marzach?? C. (2011) The major
667 antigenic membrane protein of "Candidatus Phytoplasma asteris" selectively interacts
668 with ATP synthase and actin of leafhopper vectors. *PLoS ONE* **6**.
669 Green E.R. & Mecsas J. (2016) Bacterial Secretion Systems – An overview. *Microbiology*
670 *spectrum* **4**.
671 Hoshi A., Oshima K., Kakizawa S., Ishii Y., Ozeki J. & Hashimoto M. (2009) A unique
672 virulence factor for proliferation and dwarfism in plants identified from a
673 phytopathogenic bacterium. *Proceedings of the National Academy of Sciences of the*
674 *United States of America* **106**, 6416–6421.
675 Hren M., Nikolić P., Rotter A., Blejec A., Terrier N., Ravnikar M., … Gruden K. (2009)
676 "Bois noir" phytoplasma induces significant reprogramming of the leaf transcriptome in
677 the field grown grapevine. *BMC Genomics* **10**, 460.
678 Huang W., MacLean A.M., Sugio A., Maqbool A., Busscher M., Cho S.T., … Hogenhout
679 S.A. (2021) Parasitic modulation of host development by ubiquitin-independent protein
680 degradation. *Cell* **184**, 5201–5214.
681 Jammer A., Gasperl A., Luschin-Ebengreuth N., Heyneke E., Chu H., Cantero-Navarro E., …
682 Roitsch T. (2015) Simple and robust determination of the activity signature of key
683 carbohydrate metabolism enzymes for physiological phenotyping in model and crop
684 plants. *Journal of experimental botany* **66**, 5531–42.
685 Janik K., Mithöfer A., Raffeiner M., Stellmach H., Hause B. & Schlink K. (2017) An effector
686 of apple proliferation phytoplasma targets TCP transcription factors-a generalized
687 virulence strategy of phytoplasma? *Molecular Plant Pathology* **18**, 435–442.
688 Kakizawa S., Oshima K., Nishigawa H., Jung H.Y., Wei W., Suzuki S., … Namba S. (2004)
689 Secretion of immunodominant membrane protein from onion yellows phytoplasma
690 through the Sec protein-translocation system in Escherichia coli. *Microbiology* **150**, 135–
691 142.
692 Karimi M., Inzé D. & Depicker A. (2002) GATEWAY^{TM} vectors for Agrobacterium-
693 mediated plant transformation. *Trends in Plant Science* **7**, 193–195.
694 Kotani T., Kirisako H., Koizumi M., Ohsumi Y. & Nakatogawa H. (2018) The Atg2-Atg18
695 complex tethers pre-autophagosomal membranes to the endoplasmic reticulum for
696 autophagosome formation. *Proceedings of the National Academy of Sciences of the*

United States of America **115**, 10363–10368.

Krogh A., Larsson B., von Heijne G. & Sonnhammer E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology* **305**, 567–580.

Kube M., Mitrovic J., Duduk B., Rabus R. & Seemüller E. (2012) Current view on phytoplasma genomes and encoded metabolism. *The Scientific World Journal* **2012**, 1–25.

Leary A.Y., Sanguankiattichai N., Duggan C., Tumtas Y., Pandey P., Segretin M.E., … Bozkurt T.O. (2018) Modulation of plant autophagy during pathogen attack. *Journal of Experimental Botany* **69**, 1325–1333.

Lefèvre F. & Boutry M. (2018) Towards Identification of the Substrates of ATP-Binding Cassette Transporters. *Plant physiology* **178**, 18–39.

Lu Y.-T., Li M.-Y., Cheng K.-T., Tan C.M., Su L.-W., Lin W.-Y., … Yang J.-Y. (2014) Transgenic plants that express the phytoplasma effector SAP11 show altered phosphate starvation and defense responses. *Plant physiology* **164**, 1456–69.

Lukan T., Machens F., Coll A., Baebler Š., Messerschmidt K. & Gruden K. (2018) Plant X-tender: An extension of the AssemblX system for the assembly and expression of multigene constructs in plants. *PLoS ONE* **13**, e0190526.

MacLean A.M., Orlovskis Z., Kowitwanich K., Zdziarska A.M., Angenent G.C., Immink R.G.H. & Hogenhout S.A. (2014) Phytoplasma Effector SAP54 Hijacks Plant Reproduction by Degrading MADS-box Proteins and Promotes Insect Colonization in a RAD23-Dependent Manner. *PLoS Biology* **12**.

MacLean A.M., Sugio A., Makarova O. V., Findlay K.C., Grieve V.M., Tóth R., … Hogenhout S.A. (2011) Phytoplasma effector SAP54 induces indeterminate leaf-like flower development in Arabidopsis plants. *Plant Physiology* **157**, 831–841.

Marín M., Thallmair V. & Ott T. (2012) The intrinsically disordered N-terminal region of AtREM1.3 remorin protein mediates protein-protein interactions. *The Journal of biological chemistry* **287**, 39982–39991.

Marrero M.C., Capdevielle S., Huang W., Busscher M., Busscher-Lange J., De Ridder D., … Immink R.G.H. (2020) Widespread targeting of development-related host transcription factors by phytoplasma effectors.

Mehle N., Kavčič S., Mermal S., Vidmar S., Pompe Novak M., Riedle-Bauer M., … Dermastia M. (2022) Geographical and Temporal Diversity of ‘ Candidatus Phytoplasma

730 solani ' in Wine-Growing Regions in Slovenia and Austria. *Frontiers in Plant Science*
731 **13**, 889675.
732 Mitrović J. & Duduk B. (2011) Occurrence of a new stolbur strain in tobacco in Serbia.
733 *Bulletin of Insectology* **64**, 107–108.
734 Mitrović J., Siewert C., Duduk B., Hecht J., Molling K., Broecker F., … Kube M. (2014)
735 Generation and analysis of draft sequences of "stolbur" phytoplasma from multiple
736 displacement amplification templates. *Journal of Molecular Microbiology and*
737 *Biotechnology* **24**, 1–11.
738 Mitrović M., Jakovljević M., Jović J., Krstić O., Kosovac A., Trivellone V., … Cvrković T.
739 (2016) '*Candidatus* Phytoplasma solani' genotypes associated with potato stolbur in
740 Serbia and the role of *Hyalesthes obsoletus* and *Reptalus panzeri* (hemiptera, cixiidae) as
741 natural vectors. *European Journal of Plant Pathology* **144**, 619–630.
742 Murolo S. & Romanazzi G. (2015) In-vineyard population structure of '*Candidatus*
743 Phytoplasma solani' using multilocus sequence typing analysis. *Infection, Genetics and*
744 *Evolution* **31**, 221–230.
745 Music Seruga M., Samarzija I., Hogenhout S.A., Haryono M., Cho S.T. & Kuo C.H. (2019)
746 The genome of "Candidatus Phytoplasma solani" strain SA-1 is highly dynamic and
747 prone to adopting foreign sequences. *Systematic and applied microbiology* **42**, 117–127.
748 Namba S. (2019) Molecular and biological properties of phytoplasmas. *Proceedings of the*
749 *Japan Academy. Series B, Physical and biological sciences* **95**, 401–418.
750 Nielsen H. (2017) Protein function prediction: Methods and Protocols. *Methods in Molecular*
751 *Biology* **1611**, 1–243.
752 Pecher P., Moro G., Canale M.C., Capdevielle S., Singh A., MacLean A., … Hogenhout S.A.
753 (2019) Phytoplasma SAP11 effector destabilization of TCP transcription factors
754 differentially impact development and defence of Arabidopsis versus maize. *PLOS*
755 *Pathogens* **15**, e1008035.
756 Preiss U., Keuck A. & Albert G. (2008) First detection of stolbur phytoplasma in potato in
757 Rhineland Palatinate. In *Deutsche Pflanzenschutztagung*. Kiel, Germany.
758 Prezelj N., Covington E., Roitsch T., Gruden K., Fragner L., Weckwerth W., … Dermastia M.
759 (2016a) Metabolic consequences of infection of grapevine (*Vitis vinifera* L.) cv. "Modra
760 frankinja" with flavescence dorée phytoplasma. *Frontiers in Plant Science* **7**, 711.
761 Prezelj N., Fragener L., Weckwerth W. & Dermastia M. (2016b) Metabolome of grapevine
762 leaf vein-enriched tissue infected with '*Candidatus* Phytoplasma solani'. *Mitteilungen*

763 *Klosterneuburg Rebe und Wein, Obstbau und Früchteverwertung* **66**, 74–78.
764 Quaglino F., Passera A., Faccincani M., Moussa A., Pozzebon A., Sanna F., … Mori N.
765 (2021) Molecular and spatial analyses reveal new insights on Bois noir epidemiology in
766 Franciacorta vineyards. *Annals of Applied Biology* **179**, 151–168.
767 Quaglino F., Zhao Y., Casati P., Bulgari D., Bianco P.A., Wei W. & Davis R.E. (2013)
768 '*Candidatus* Phytoplasma solani', a novel taxon associated with stolbur- and bois noir-
769 related diseases of plants. *International journal of systematic and evolutionary*
770 *microbiology* **63**, 2879–94.
771 Rahman O., Cummings S.P., Harrington D.J. & Sutcliffe I.C. (2008) Methods for the
772 bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-
773 positive bacteria. *World Journal of Microbiology and Biotechnology* **24**, 2377–2382.
774 Riedle-Bauer M., Brader G., Wieland L., Schönhuber C., Aryan A. & Fuchs F. (2017)
775 Untersuchungen zur Rolle verschiedener Zikadenarten als Überträger von ‚Candidatus
776 Phytoplasma solani' an Kartoffeln und Tomaten in Österreich. In *58th Österreichische*
777 *Pflanzenschutztage*. St. Pölten, Austria.
778 Rotter A., Nikolić P., Turnšek N., Kogovšek P., Blejec A., Gruden K. & Dermastia M. (2018)
779 Statistical modeling of long-term grapevine response to '*Candidatus* Phytoplasma solani'
780 infection in the field. *European Journal of Plant Pathology* **150**, 653–668.
781 Sies H., Berndt C. & Jones D.P. (2017) Oxidative Stress. *https://doi.org/10.1146/annurev-*
782 *biochem-061516-045037* **86**, 715–748.
783 Siewert C., Luge T., Duduk B., Seem??ller E., B??ttner C., Sauer S. & Kube M. (2014)
784 Analysis of expressed genes of the bacterium "Candidatus phytoplasma mali" highlights
785 key features of virulence and metabolism. *PLoS ONE* **9**.
786 Škrlj B., Pompe Novak M., Brader G., Anžič B., Ramšak Ž., Gruden K., … Dermastia M.
787 (2021) New cross-talks between pathways involved in grapevine infection with
788 '*Candidatus* Phytoplasma solani' revealed by temporal network modelling. *Plants*
789 *(Basel, Switzerland)*.
790 Suzuki S., Oshima K., Kakizawa S., Arashida R., Jung H.-Y., Yamaji Y., … Namba S. (2006)
791 Interaction between the membrane protein of a pathogen and insect microfilament
792 complex determines insect-vector specificity.
793 Suzuki S.W., Yamamoto H., Oikawa Y., Kondo-Kakuta C., Kimurac Y., Hirano H. & Ohsumi
794 Y. (2015) Atg13 HORMA domain recruits Atg9 vesicles during autophagosome
795 formation. *Proceedings of the National Academy of Sciences of the United States of*

America **112**, 3350–3355.

Tan Y., Li Q., Zhao Y., Wei H., Wang J., Baker C.J., … Id W.W. (2021) Integration of metabolomics and existing omics data reveals new insights into phytoplasma-induced metabolic reprogramming in host plants. *Plos One* **16**, e0246203.

Tomkins M., Kliot A., Marée A.F. & Hogenhout S.A. (2018) A multi-layered mechanistic modelling approach to understand how effector genes extend beyond phytoplasma to modulate plant hosts, insect vectors and the environment. *Current Opinion in Plant Biology* **44**, 39–48.

Wun C.L., Quan Y. & Zhuang X. (2020) Recent Advances in Membrane Shaping for Plant Autophagosome Biogenesis. *Frontiers in Plant Science* **11**, 565.

Yun H.R., Jo Y.H., Kim J., Shin Y., Kim S.S. & Choi T.G. (2020) Roles of Autophagy in Oxidative Stress. *International journal of molecular sciences* **21**.

Zeng Y. & Charkowski A.O. (2021) The Role of ATP-Binding Cassette Transporters in Bacterial Phytopathogenesis.

Zhou J., Ma F., Yao Y., Deng M., Chen | Mengting, Zhang S., … Sun J. (2021) Jujube witches' broom phytoplasma effectors SJP1 and SJP2 induce lateral bud outgrowth by repressing the ZjBRC1-controlled auxin efflux channel.

2.3 TGA Transcription Factors – Structural Characteristics as Basis for Functional Variability

Špela Tomaž, Kristina Gruden and Anna Coll

Frontiers in Plant Science, 2022, 13:935819. DOI: 10.3389/fpls.2022.935819

This review paper provides a new perspective on the TGA transcription factor mechanism of action by focusing on the structural specifics of TGA proteins in Arabidopsis and tobacco. A comprehensive review of TGA function was last published in 2013, while the structural-functional features of individual TGAs have not yet been addressed. Our paper introduces the reader to the functional diversity of TGAs and describes each of their three main protein parts: the intrinsically disordered N-terminus, the conserved DNA-binding domain, and the C-terminus. By combining published data with *in silico* analyses of TGA sequences, we connect their structural characteristics with their functional roles.

The PhD candidate carefully examined the available literature on the different roles of TGA transcription factor structural parts and used the protein sequences of Arabidopsis and tobacco TGAs to create sequence alignments, perform phylogenetic analyses, and intrinsic disorder predictions. The PhD candidate wrote the first manuscript draft and prepared the figures.

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 26 July 2022
DOI 10.3389/fpls.2022.935819

OPEN ACCESS

EDITED BY
Ian T. Major,
Canadian Forest Service, Canada

REVIEWED BY
Mark Zander,
Rutgers, The State University of New Jersey, United States
Shelley Hepworth,
Carleton University, Canada

*CORRESPONDENCE
Špela Tomaž
spela.tomaz@nib.si

SPECIALTY SECTION
This article was submitted to Plant Cell Biology, a section of the journal Frontiers in Plant Science

RECEIVED 04 May 2022
ACCEPTED 04 July 2022
PUBLISHED 26 July 2022

CITATION
Tomaž Š, Gruden K and Coll A (2022) TGA transcription factors—Structural characteristics as basis for functional variability.
Front. Plant Sci. 13:935819.
doi: 10.3389/fpls.2022.935819

TGA transcription factors—Structural characteristics as basis for functional variability

Špela Tomaž[1,2]*, Kristina Gruden[1] and Anna Coll[1]

[1]Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia, [2]Jožef Stefan International Postgraduate School, Ljubljana, Slovenia

TGA transcription factors are essential regulators of various cellular processes, their activity connected to different hormonal pathways, interacting proteins and regulatory elements. Belonging to the basic region leucine zipper (bZIP) family, TGAs operate by binding to their target DNA sequence as dimers through a conserved bZIP domain. Despite sharing the core DNA-binding sequence, the TGA paralogues exert somewhat different DNA-binding preferences. Sequence variability of their N- and C-terminal protein parts indicates their importance in defining TGA functional specificity through interactions with diverse proteins, affecting their DNA-binding properties. In this review, we provide a short and concise summary on plant TGA transcription factors from a structural point of view, including the relation of their structural characteristics to their functional roles in transcription regulation.

KEYWORDS

DOG1 domain, functional variability, intrinsically disordered regions, plant transcription regulation, post-translational modifications, structural characteristics, TGA transcription factors

Introduction

The *Arabidopsis thaliana* genome encodes for over 2,200 transcription factor genes, according to the Plant Transcription Factor Database[1] yet few of them have been thoroughly characterized. The TGACG-binding (TGA) transcription factors were among the first plant transcription factors ever studied, their discovery dating back to the year of 1989 (Katagiri et al., 1989). Named after their hallmark binding site, the TGA factors became known for their regulation of defense-related genes through interaction with NON-EXPRESSOR OF PR-1 (NPR1) cofactor (Zhang et al., 1999), a salicylic acid receptor and master regulator of plant immunity (Wu et al., 2012; Backer et al., 2019; Wang W. et al., 2020). Among dicot plant species, the ten Arabidopsis TGA factors,

1 http://planttfdb.gao-lab.org

AtTGA1-7, AtPERIANTHIA (AtPAN), and AtTGA9-10, have been investigated most thoroughly, next to five tobacco (*Nicotiana tabacum*) members, NtTGA1A, NtPG13, NtTGA2.1, NtTGA2.2, and NtTGA10. They are distributed into five clades (Jakoby et al., 2002), which are phylogenetically divided into two branches (**Figure 1**). Functional analysis of TGAs from different clades revealed not only their importance in biotic stress response (Zhang et al., 2003; Kesarwani et al., 2007; Zander et al., 2010; Sun et al., 2018), but also in regulation of gene expression connected to abiotic stress responses (Zhong et al., 2015; Fang et al., 2017; Herrera-Vásquez et al., 2021), developmental processes (Murmu et al., 2010; Maier et al., 2011; Wang et al., 2019), circadian rhythm (Zhou et al., 2015), detoxification (Fode et al., 2008; Mueller et al., 2008; Herrera-Vásquez et al., 2021), nitrate signaling (Alvarez et al., 2014; Canales et al., 2017), flowering (Thurow et al., 2005; Song et al., 2008; Maier et al., 2011; Xu et al., 2021), and autophagy (Wang P. et al., 2020) (**Figure 1**). Initially, the function of TGAs from clades I, II, and III was mainly associated with plant immunity, whereas the first reports on clade IV and V members revealed their role in regulating developmental processes (reviewed in Gatz, 2013). However, this apparent functional division is becoming less evident as an increasing number of reports show that most clades are involved in a variety of processes (**Figure 1**). For example, clade I TGAs have also been shown to be involved in regulating growth and development (Li et al., 2019; Wang et al., 2019), while clade IV TGAs are also important in biotic stress (Noshi et al., 2016; Venturuzzi et al., 2021).

A detailed review, integrating TGA transcription factor research, was published last in 2013 (Gatz, 2013). The complexity of TGA involvement in various molecular processes, the lack of research supporting the proposed mechanisms of action and limited reports on the relation between structure and *in vivo* function have become characteristic features of TGA studies and are the reason for our lack of knowledge about their mechanism of action. In order to understand how transcription factors operate, we must consider their structural characteristics as the underlying basis of protein activity. Here, we consider the importance of reported and *in silico* determined TGA structural features in defining their functional specificity and variability, focusing on the characterized TGA factors from Arabidopsis and tobacco.

From TGA factor structure to their function

For more than 30 years after their discovery, the TGA protein three-dimensional (3D) structure remained a mystery and the first structural data have been published only recently in a breakthrough report providing a partial cryo-electron microscopy (cryo-EM) structure of AtTGA3 in complex with NPR1 (Kumar et al., 2022). Additionally, the novel artificial intelligence algorithm AlphaFold (AF) allows structure prediction without the availability of known similar structures (Jumper et al., 2021) and AF models for full-length Arabidopsis and tobacco TGAs are already available in the AF Protein Structure Database (**Figure 2A**; Varadi et al., 2022). In the following subchapters, we aim to connect literature reports with *in silico* analyses, to better understand the biological role of the three main structural parts of TGAs: The conserved basic region leucine zipper (bZIP) domain, the highly variable N-terminal part, from now on referred to as the N-terminus, and the C-terminal part (C-terminus), containing a putative Delay of Germination 1 (DOG1) domain (**Figures 2A,B**).

The highly conserved DNA-binding domain

TGAs are members of the bZIP protein superfamily and represent a plant-specific subgroup, found in different species, including mosses and liverworts (Gutsche and Zachgo, 2016; Gutsche et al., 2017). bZIP proteins are defined by their DNA-binding and dimerization region known as the bZIP domain, which is highly conserved among plants and even across kingdoms (Jindrich and Degnan, 2016). Multiple sequence alignment of Arabidopsis and tobacco TGAs shows the bZIP domain as the region of highest protein sequence identity, regardless of plant species (**Figure 2C**). The bZIP domain determines DNA-binding specificity and serves as a nuclear localization signal (**Figure 2B**; van der Krol and Chua, 1991; Deppmann et al., 2004). It consists of two regions, the basic region and the leucine zipper. Hydrogen bond formation with the major DNA groove is facilitated through the basic region, which contains an invariant N-x_7-R/K motif (Dröge-Laser et al., 2018). Nineteen out of 20 amino acids of the TGA basic region, including the N-x_7-R motif, remain identical in all aligned Arabidopsis and tobacco sequences, with a few clade/species-specific differences present only in the first residue (**Figure 2D**).

bZIP proteins bind target DNA as dimers, with combinatorial homo- or heterodimerization at the DNA-binding site granting them broad variability in regulation of physiological responses (Deppmann et al., 2006; Rodríguez-Martínez et al., 2017). The leucine zipper confers dimer formation and determines dimerization specificity. It consists of repetitive seven-amino acid units, called heptads. Each heptad contains a conserved leucine residue at its fifth position (Landschulz et al., 1988; Deppmann et al., 2006). TGA bZIP domains have three leucine zipper heptads, which show higher variability than the basic region, yet retain the conserved leucines. The only exception is AtTGA10, where the third leucine is replaced with isoleucine (**Figure 2D**). The number of heptads is among the lowest compared to other Arabidopsis bZIP proteins (Deppmann et al., 2004), rendering the 41 aa bZIP domains in TGAs considerably shorter from the typical 60–80

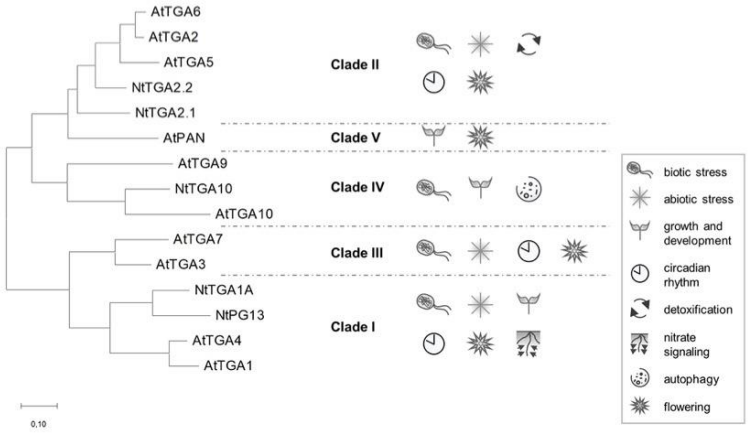


FIGURE 1
Unrooted phylogenetic tree of Arabidopsis and tobacco TGAs. Phylogenetic analysis of TGA factors shows an earlier separation of clades into two branches, one dividing into clades II, IV, and V, the other into clades I and III, indicating a closer evolutionary relationship between clade members in the same branch. TGA involvement in regulation of different processes, based on literature search, is represented for each clade. Sequence alignment by MUSCLE (Edgar, 2004) and phylogenetic analysis by the Maximum Likelihood method, based on the JTT matrix-based model (Jones et al., 1992), were conducted in MEGA7 (Kumar et al., 2016). The branch length scale represents the number of substitutions per site. Protein sequences with listed protein identification numbers were retrieved from UniProtKB (https://www.uniprot.org/): AtTGA1 (Q39237), AtTGA2 (P43273), AtTGA3 (Q39234), AtTGA4 (Q39162), AtTGA5 (Q39163), AtTGA6 (Q39140), AtTGA7 (Q93ZE2), AtPAN (Q9SX27), AtTGA9 (Q93XM6), AtTGA10 (E3VNM4), NtTGA1A (P14232), NtPG13 (Q05699), NtTGA2.1 (O24160), NtTGA2.2 (Q9SQK1), and NtTGA10 (Q52MZ2).

aa bZIP length (Jindrich and Degnan, 2016). Additionally, the leucine zippers of TGAs contain destabilizing residues at dimer contact sites, making the zipper formation less stable (Deppmann et al., 2004). The extent to which dimerization stability affects transcription factor binding time at its specific motif is not known and unstable interactions might shorten DNA-binding times, resulting in a lower number of generated transcripts (Swift and Coruzzi, 2017).

During DNA-binding the bZIP domains of two proteins grip the DNA segment in a scissor-like fashion, while attaining an alpha-helical fold (Vinson et al., 1989; Ellenberger et al., 1992), as shown in the human FosB-JunD-DNA complex 3D structure (**Figure 2E**; Yin et al., 2017). The role of the leucine zipper in TGA dimerization had been demonstrated by switching the leucine zipper in NtTGA2.2 for a zipper from the human Jun bZIP protein, which prevented heterodimerization with NtTGA2.1 (Thurow et al., 2005). However, deletion of 93 aa from AtTGA2 N-terminus, including the entire bZIP domain, still allowed homodimer formation (Boyle et al., 2009) and multiple reports have shown that TGA dimerization depends significantly on other protein parts as well. A dimer stabilization region had been identified in the C-terminus of tobacco NtTGA1A, located between 178 and 373 aa (Katagiri et al., 1992). Formation of stable contacts through the TGA C-terminus was recently confirmed with cryo-EM, which revealed homodimerization of AtTGA3 C-termini in the AtTGA3-NPR1 complex structure (Kumar et al., 2022). Moreover, deleting the residues from 146 to 330, spanning more than half of the protein C-terminus, abolished the DNA-binding activity of AtTGA2 (Johnson et al., 2008), which could be due to hindered dimerization. Additionally, protein interaction analyses *in vitro* and *in vivo* showed that TGAs can form homodimers, heterodimers as well as higher order complexes (Niggeweg et al., 2000; Schiermeyer et al., 2003; Boyle et al., 2009) and the oligomerization properties of AtTGA2 seem to be dependent on the region spanning its N-terminus and bZIP domain (Boyle et al., 2009).

Recognition of only a short DNA sequence is usually sufficient for transcription factor binding (Kribelbauer et al., 2019). The TGACG pentamer is the common TGA dimer binding site and sufficient for their binding (Katagiri et al., 1992; Schindler et al., 1992; Izawa et al., 1993). ChIP-seq analysis of AtTGA2 revealed that 55% of significantly enriched regions in the Arabidopsis genome contained the TGACGTCA palindrome, while all carried at least the TGACG core motif (Thibaud-Nissen et al., 2006). The palindrome was also determined as the representative binding motif of AtTGAs in DAP-seq data (O'Malley et al., 2016). In addition, tandem TGACG repeats, such as the *activating sequence-1* (*as-1*) or *as-1*-like elements, allow more options regarding

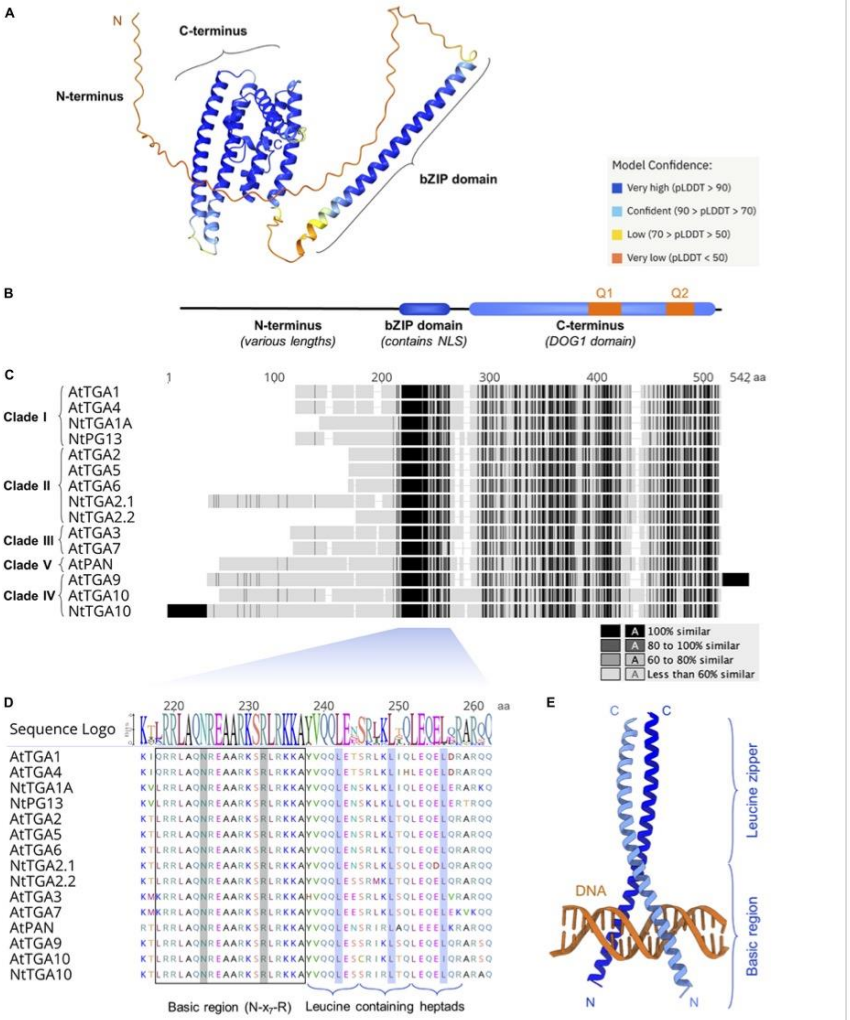


FIGURE 2

Among the three main TGA protein parts, the bZIP domain is the most highly conserved. **(A)** The AlphaFold generated 3D model of AtTGA1 (Jumper et al., 2021; Varadi et al., 2022) (pLDDT, AlphaFold per-residue confidence score) and **(B)** a schematic representation of TGA domain organization, showing the flexible N-terminus, the bZIP domain and the C-terminus, encompassing a putative Delay of Germination 1 (DOG1) domain. The nuclear localization signal (NLS) and glutamine rich regions Q1 and Q2 are indicated. **(C)** Multiple sequence alignment of ten Arabidopsis and five tobacco TGAs, with segments of high similarity or identity colored darkest and lowest similarity lightest, shows the bZIP domain retains the highest sequence identity throughout the whole protein sequence. In cases where sequence segments at N-terminal or C-terminal ends are not aligned to any of the other sequences, they are considered identical and are colored black. **(D)** A closer examination of the bZIP domain shows few variations in the basic region, while the three zipper heptads, with conserved leucine residue positions marked, show higher variability. The alignment and sequence logo were prepared and visualized with Geneious Prime 2020 (Kearse et al., 2012), using default parameters. **(E)** Structural model of the human FosB-JunD bZIP dimer in complex with DNA (5VPE entry in RCSB PDB) (Yin et al., 2017). The models in **(A,E)** were visualized in UCSF ChimeraX (Pettersen et al., 2021).

the binding stoichiometry. For example AtTGA2, AtTGA5, NtTGA2.1, NtTGA2.2, and NtTGA10, can bind tandem repeats in two-dimer complexes (tetramers). Although past reports indicated that AtTGA1, AtTGA3, and NtTGA1A prefer single-dimer formation (Lam and Lam, 1995; Niggeweg et al., 2000; Schiermeyer et al., 2003), Kumar et al. (2022) show single and double occupancy of tandem repeats in the *Pathogenesis related-1 (PR-1)* promoter by AtTGA3 in electrophoretic mobility shift assays. The single-occupancy band is depleted in the presence of NPR1, which supershifts the double-occupancy band. The AtTGA3-NPR1-DNA complex structure consists of four AtTGA3 and two NPR1 proteins, where an NPR1 dimer connects two DNA-bound AtTGA3 dimers (Kumar et al., 2022). The spacing between tandem repeats is also important, as it affects element recognition, binding affinity and TGA transcription activation ability (Krawczyk et al., 2002). Moreover, TGA paralogues have been shown to occupy the A-box (TACGTA), C-box (GACGTC), G-box (CACGTG), and T-box (AACGTT) motifs with different affinities (Izawa et al., 1993; Wang et al., 2019).

Heterodimerization of transcription factors with different binding preferences can, undoubtedly, result in distinct DNA-binding specificities and affinities (Rodríguez-Martínez et al., 2017). NtTGA2.1/NtTGA2.2-NtTGA1A heterodimers have been recruited to a single TGACG motif (Niggeweg et al., 2000). Homodimer binding can be stabilized by the presence of other TGA homodimers at the tandem occupancy site (Lam and Lam, 1995). Besides the motif sequence, adjacent sequence regions are also important for binding, as they determine intrinsic DNA properties and consequently the transcription factor binding affinity (Kribelbauer et al., 2019). For example, when compared to the *as-1*-like element, the affinity of AtPAN was stronger for the 33 bp long AAGAAT motif, characterized by an AAGAAT sequence upstream of a single TGACG pentamer (Gutsche and Zachgo, 2016). Additionally, local adjustments of TGA concentration, for example through specific subnuclear localization, may contribute to successful binding to suboptimal binding sites (Kribelbauer et al., 2019).

Staying flexible through N-terminus

The N-terminus is likely a major contributor in defining TGA functional specificity. It is the least conserved part of the TGA structure, with a high variability in amino acid sequence and length (**Figure 2C**). Studies analyzing the TGA N-terminus function indicate its various roles, but only a few of them have been corroborated by further analysis. Due to its relatively high acidic amino acid content (9–24%, calculated as the percentage of aspartic and glutamic acid residues), it was proposed that it likely takes on a transcription regulation function (Katagiri et al., 1989). Acidic regions are common to many transcription activation domains and, according to the model proposed by Staller et al. (2018), help exposing the hydrophobic residues to facilitate contact formation with coactivators. In support of this, removing the N-terminus of NtTGA1A diminished its transcription activation ability in tobacco cotyledons (Neuhaus et al., 1994).

Because activation domains often interact with a variety of structurally distinct coactivators, they are usually intrinsically disordered as well, thus their 3D structure is hard to determine (Staller et al., 2018). In line with this, the TGA N-terminal region is mostly unstructured in the AF models (**Figure 2A**) and has an overall high probability of disorder (>0.5) with variably interspaced more structured sections according to three intrinsic disorder predictors, IUPred2 (**Figure 3**; Mészáros et al., 2018), PrDOS (Ishida and Kinoshita, 2007) and SPOT-disorder (Hanson et al., 2017) (**Supplementary Figure 1**). The intrinsic disorder pattern as well as the N-terminus length seem largely clade-dependent. Clade I and III TGAs all harbor medium length N-termini of about 75–100 aa, which, according to two of the prediction algorithms, share a higher probability of disorder near the basic region, with disorder decreasing further away from the bZIP domain. Most clade II members exhibit short N-termini of about 40–50 aa, with a high disorder-probability spanning their whole length. Clade IV and clade V member N-termini are the longest, reaching 216 aa in NtTGA10. While intrinsic disorder is higher in clade IV TGA N-termini, the AtPAN N-terminus seems more structured based on IUPred2 and PrDOS predictions (**Figure 3** and **Supplementary Figure 1**).

Electrophoretic mobility shift assay results have shown that the long N-terminus of NtTGA2.1 enabled weak binding to the *as-1* element, while the binding of shorter NtTGA2.2 was stronger. Furthermore, shortening the NtTGA2.1 N-terminus increased the protein DNA-binding stability (Niggeweg et al., 2000). Modulation of protein-DNA interaction stability therefore may be one of the N-terminus features. Additional stabilizing elements in DNA-binding motif vicinity can stabilize the protein-DNA interactions through parts other than the DNA-binding domain. Motif recognition and binding stabilization through the N-terminus have been shown in other transcription factors. For instance, the N-terminal arm of the *Drosophila melanogaster* Hox homeodomain factor stabilizes the binding to DNA by inserting the positively charged amino acids within the arm into the minor DNA groove (Abe et al., 2015).

While several discrepancies regarding the mechanisms of AtTGA-mediated transcriptional regulation in cooperation with NPR1 remain (discussed previously in Gatz, 2013), the study from Boyle et al. (2009) indicated that the TGA N-terminus is important for determining its activation/repression function. In accordance with the results from Zhang et al. (2003) and Kesarwani et al. (2007), who show that AtTGA2 acts as a constitutive repressor, modulating basal promoter activity of the *PR-1* gene, the AtTGA2 N-terminus assumes a non-autonomous

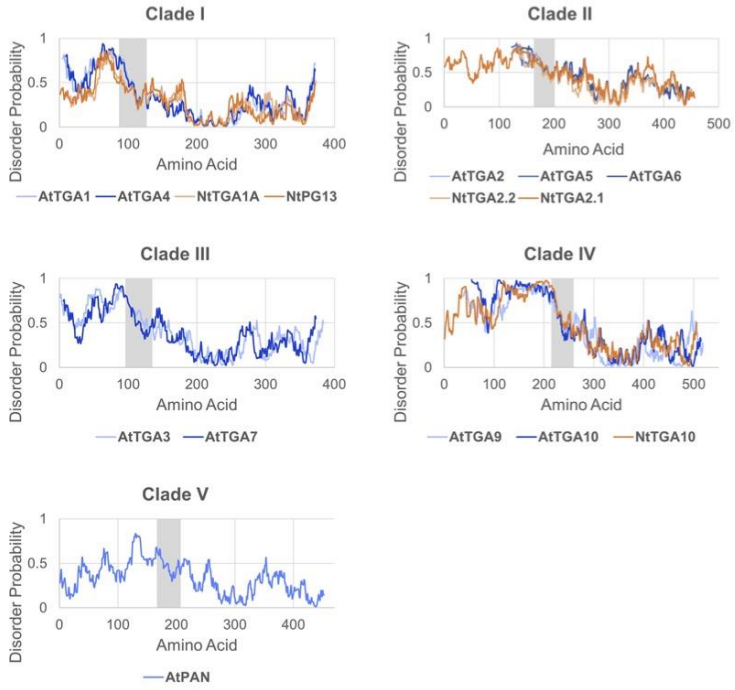


FIGURE 3
The N-termini of TGAs are intrinsically disordered. Representation of intrinsic disorder regions of full-length TGA amino acid sequences from Arabidopsis and tobacco, created based on IUPred2 prediction algorithm results (Mészáros et al., 2018). The N-termini of analyzed TGAs show generally high (>0.5), yet clade-specific pattern of intrinsic disorder probability, while the disorder probability is considerably lower in their C-termini. Charts representing TGAs from the same clade were aligned based on the conserved bZIP domain, which is shown as grey area.

repression function and proved important also for AtTGA2 oligomerization at the DNA-binding site (Boyle et al., 2009). However, AtTGA2 interaction with NPR1 in the presence of salicylic acid activates gene expression (Rochon et al., 2006). In the proposed NPR1-AtTGA2 activation complex, NPR1 prevents the DNA-binding of AtTGA2 oligomer and negates the AtTGA2 repression function through interaction with its N-terminus (Boyle et al., 2009).

Interestingly, Gutsche and Zachgo (2016) described the role of AtPAN N-terminus in connection to its redox-sensitive DNA-binding. Its unique feature is the presence of five cysteine residues, dispersed throughout the N-terminus. AtPAN binds the AAGAAT motif in reducing conditions, while an oxidizing environment diminishes this interaction. Mutations of all AtPAN cysteines, including Cys340 in the C-terminus or the complete removal of N-terminus, both prevented such redox-dependent motif-binding (Gutsche and Zachgo, 2016). The activity of the N-terminus was further confirmed *in planta*. The expression of either AtPAN N-terminus deletion mutant, AtPAN with mutated N-terminal Cys68 and Cys87, or of AtPAN with substituted of all N-terminal cysteines to serines could not complement the *pan* knockout plant phenotype (Gutsche and Zachgo, 2016). Nevertheless, the mechanisms of redox-dependent sensitivity based on N-terminal cysteines remain to be elucidated.

On the other hand, the TGA N-terminus effects on interactions with other proteins should also be considered. The AtTGA2 N-terminus can be bound by the copper chaperone induced by pathogens (CPP) and the AtTGA2-CPP interaction enhances AtTGA2 binding to the *PR-1* promoter (Chai et al., 2020). The N-terminal half of AtTGA3, including the bZIP domain, interacts with the WRKY53 transcription factor (Sarkar et al., 2018), while it also enhances AtTGA3 interaction with NPR1 in yeast (Zhou et al., 2000). Yeast two-hybrid

assays of AtPAN and AtTGA3 deletion mutants indicated that their N-termini strengthen interactions with ROXY1 CC-type glutaredoxin (Li et al., 2011), however, these results have not been confirmed with quantitative analyses.

C-terminus: Glutamine rich regions and the DOG1 domain

TGAs share a relatively conserved C-terminus of about 250 aa in length. Overall, it has a lower intrinsic disorder-probability in all clades (**Figure 3** and **Supplementary Figure 1**) and contains two 20–30 aa long regions rich in glutamine residues, designated Q1 and Q2 (**Figure 2B**; Katagiri et al., 1989; Gatz, 2013). Glutamine-rich regions occur in transcription activation domains and can modulate transcription activation through unknown mechanisms (Arnold et al., 2018). Activation and/or repression function of individual TGAs is therefore likely the result of both N- and C-terminal contributions of the same protein. As described above, the TGA N-terminus is important for modulation of protein-protein interactions. However, the C-terminus has been identified as the main protein-protein interaction region in several studies. It is sufficient for interaction of AtTGA2 and AtTGA3 with NPR1 (Fan and Dong, 2002; Johnson et al., 2008; Kumar et al., 2022). AtPAN and AtTGA3 interact with ROXY1 primarily through the Q2 and the intervening region, which represents the first third of their C-termini (Li et al., 2011).

Consistent with the TGA AF model (**Figure 2A**), the cryo-EM and crystallographic data presented by Kumar et al. (2022) show that the 3D structure of AtTGA3 C-terminus (aa 166–377) is predominantly alpha-helical, containing five longer and three shorter alpha-helices that are connected with flexible linkers. The alpha-helices envelop a single molecule of palmitic acid, the role of which has yet to be elucidated. The C-terminus is involved in AtTGA3 dimerization as well as interaction with the ankyrin repeat region of NPR1, leading the authors to refer to it as NPR1-interacting domain (NID). The NID forms contacts with NPR1 through four residues near the center of the AtTGA3 C-terminus sequence (Glu263, Pro264, Thr266, and Asp267) and four residues close to its C-terminal end (Thr351, Thr352, Arg353, and Arg357). Additionally, by using a series of chimeric AtTGA1/AtTGA2 proteins, Després et al. (2003) show the importance of AtTGA2 C-terminal aa 236–266 in establishing the interaction with NPR1. Furthermore, the NPR1 Broad-Complex, Tramtrack, and Bric-a-brac/Pox virus and Zinc finger domain was shown to interact with AtTGA2 N-terminus (Boyle et al., 2009). While the cryo-EM structures of AtTGA3 bZIP and N-terminus could not be determined likely due to flexible linker connecting them to the NID (Kumar et al., 2022), it would be interesting to compare their involvement in the TGA-NPR complex.

The TGA C-terminus contains also the DOG1 domain, which spans most of the region according to the ExPASy Prosite domain prediction tool (Sigrist et al., 2013). The domain name originates from the Arabidopsis DOG1 protein, a plant-specific protein involved in seed dormancy control (Bentsink et al., 2006). DOG1 was also identified as a microprotein, a transcription factor-like protein of low molecular weight without the DNA-binding ability that could be involved in modulation of TGA activity (Magnani et al., 2014). Circular dichroism spectra of a recombinant DOG1 revealed it to be an alpha-helical protein as well, containing a heme-binding site important for DOG1 function (Nishimura et al., 2018). Despite low sequence identity between AtTGAs and DOG1, some DOG1 domain residues remain conserved and may contribute to the final protein fold or any key structural and functional characteristics (Sall et al., 2019), indicating the possibility of heme-binding activity also in TGAs. Phylogenetic analyses have shown that TGAs form a monophyletic group outside of DOG1 family members, which include DOG1 and five DOG1-like (DOGL) proteins (Nishiyama et al., 2021). The presence of conserved amino acid residues related to Calmodulin (CaM) binding in the C-terminus in both protein groups, indicates that DOG1 could act as a CaM-binding domain in AtTGA transcription factors (Sall et al., 2019). CaM is an important calcium (Ca^{2+}) sensor and affects a number of cellular processes in response to increased concentrations of free Ca^{2+} (Bergey et al., 2014). Several TGAs have been identified as CaM interactors (Popescu et al., 2007). The CaM/Ca^{2+} complex enhanced AtTGA3 binding to TGACG elements *in vivo* and *in vitro* by direct protein-protein interaction (Szymanski et al., 1996; Fang et al., 2017), signifying a close connection of TGA transcription regulation with Ca^{2+} influx, a primary occurrence following stress-related events in plant cells (Tian et al., 2020).

Additionally, interactions with a variety of structurally distinct proteins have been shown to affect TGA activity, but the protein part important for the interaction has not yet been identified. For instance, clade II AtTGAs interact with WRKY50 transcription factor to cooperatively activate *PR-1* gene expression (Hussain et al., 2018), or with NPR1 paralogues NPR3 and NPR4 to repress the expression of *SAR DEFICIENT 1* (*SARD1)* and *WRKY70* (Ding et al., 2018). Recently it has been shown that AtTGA5 and AtTGA7 interact with CYCLIN-DEPENDENT KINASE 8 (CDK8), involved in the recruitment of RNA polymerase II (Chen et al., 2019), and AtTGA2 with HIGH OSMOTIC STRESS GENE EXPRESSION 15 (HOS15) corepressor (Shen et al., 2020). All Arabidopsis TGAs interact with ROXY glutaredoxins, albeit with different binding affinities (Li et al., 2009; Murmu et al., 2010; Zander et al., 2012). Being able to interact with various cofactors suggests a certain flexibility in the binding region itself or the presence of multiple binding sites. Considering the C-terminus to be fairly structured, it could contain more than one protein-binding region. To further facilitate interaction specificity and modulate

binding affinity of different TGA-interactor combinations, additional contact sites are likely mediated by the variable N-terminus.

What is the role of post-translational modifications?

Post-translational modifications (PTMs) are known to be important in regulation of protein function and can affect protein activity on many levels, including dimerization, DNA-binding or protein interactions. TGAs have been found to be subjected to phosphorylation (Kang and Klessig, 2005; Kim et al., 2022), S-nitrosylation, S-glutathionylation (Lindermayr et al., 2010; Gutsche and Zachgo, 2016) and most notably redox-dependent regulation through disulphide bond formation (Després et al., 2003; Lindermayr et al., 2010; Gutsche and Zachgo, 2016). Phosphorylation was among the first PTMs studied in TGAs. Clade II AtTGAs and to a lesser extent AtTGA3, but not clade I AtTGAs, can be phosphorylated by casein kinase II (CK2) and experiments with AtTGA2 deletion mutants revealed the phosphorylation site to be within the first 20 aa of its N-terminus. The CK2-mediated phosphorylation of AtTGA2 reduced its DNA-binding activity (Kang and Klessig, 2005). On the other hand, the clade I AtTGAs have been shown to be phosphorylated by BR-INSENSITIVE2 (BIN2) at their C-terminus. This phosphorylation destabilized AtTGA4 and inhibited its interaction with NPR1 *in vivo* (Kim et al., 2022).

TGAs seem tightly connected to redox-dependent regulation and several studies focused on examining the importance of TGA cysteine residues. Clade II AtTGAs contain only one cysteine in their C-terminus, but its function remains unknown (Huang et al., 2016; Findling et al., 2018). AtPAN retains five cysteines in the N-terminus alone, which are involved in redox-dependent modulation of AtPAN DNA-binding, while the S-glutathionylation of the Cys340, localized in a putative transcription activation domain in AtPAN C-terminus, indicates that additional mechanisms could modify its activity post-translationally (Gutsche and Zachgo, 2016). Clade I AtTGAs contain four cysteines, two of which are unique and were found to facilitate redox-dependent interaction with NPR1 in the presence of salicylic acid (Després et al., 2003), with the redox regulation proposedly mediated by nitric oxide (Lindermayr et al., 2010). Furthermore, substitutions of the same cysteines prevent clade I AtTGA interactions with the NPR-family Blade-on-Petiole (BOP) proteins (Wang et al., 2019). However, Budimir et al. (2021) showed that the reduction of cysteines in AtTGA1 may not affect its function in salicylic acid-dependent gene expression. The conflicting results regarding clade I redox regulation were discussed recently (Li and Loake, 2020), and although a tight connection between TGAs and the intracellular redox state is clearly important for their activity, the role of TGA cysteine residues is still a subject of debate.

Discussion

Although TGAs represent a relatively small group of regulatory proteins, they are able to affect a wide range of cellular processes. The three main TGA protein parts have individual yet overlapping roles, jointly contributing to the functional variability of each paralogue (**Figure 4**). Thus far, the link between TGA structural characteristics and plant phenotype is not well studied. Extensive work has been dedicated to understanding how TGAs bind to DNA, as well as to which genes they regulate (Gatz, 2013). However, while one area of studies focuses on the biochemistry of binding, the other is concentrated mainly on the function of TGAs in interaction with other proteins, using knockout plants. The use of high-throughput DNA-binding methodologies in TGA studies can be challenging due to their low abundance in plant tissue and the question of how different TGAs with similar DNA-binding preferences can regulate so many different functions remains to be elucidated.

Specific studies should be designed to understand how the ability to form homo/heterodimers, tetramers or higher order complexes with other proteins at promoter regions affects target transcription and the physiological response of the plant. Most of the available molecular information is based on the studies of the *PR-1* promoter, which may not be representative. Additional gene models should be developed to study the TGA mechanism of action *in vivo*. Analyzing local DNA structural features *in silico* (Li et al., 2017; Samee et al., 2019) should also be considered for identification and analysis of TGA binding specificities. Suboptimal binding sites can contribute to TGA function as well, depending on local transcription factor molecule concentration, determined by their spatio-temporal gene expression and non-redundant subnuclear distribution (Kribelbauer et al., 2019). The mechanisms regulating TGA abundance, however, are not well understood. TGAs are differentially regulated at the expression level subsequent to pathogen infection, abiotic stress and show tissue specific expression (Chen et al., 2018; Wang et al., 2019; Seo et al., 2020). In addition they are subjected to complex post-transcriptional (Pontier et al., 2002) and post-translational processes (Kang and Klessig, 2005; Lindermayr et al., 2010; Gutsche and Zachgo, 2016).

In order to thoroughly understand the role and interplay of different TGA clade members, it is imperative to recognize key structural differences between them, individually and in higher order complexes. Resolving the complete TGA factor 3D structure in complex with their various protein interactors at target promoters would provide the basis for further experiments in studying TGA activity. Development of cryo-EM methods, which are already reaching atomic resolution (Yip et al., 2020), proved valuable in structural analysis of complexes and will continue playing an important role in the future of structure determination. Alternatively, computational modeling algorithms, such as AF (Jumper et al., 2021), can

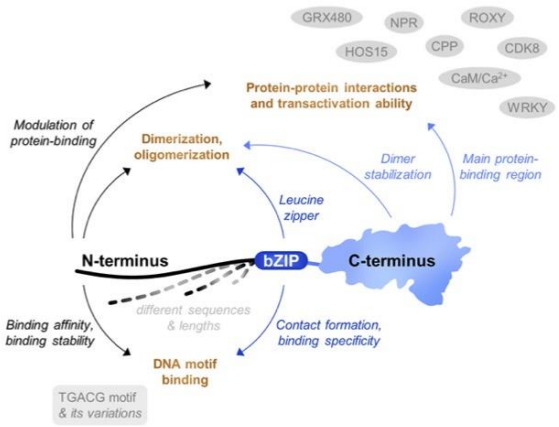


FIGURE 4
Schematic representation of TGA protein parts contribution to TGA function. All three protein parts of TGAs are multifunctional, each involved in several tasks connected to their interaction with target DNA motifs, dimerization and/or oligomerization and protein-protein interactions with transcription factors, cofactors or other proteins, resulting in a specific shift in gene expression activity.

provide a useful solution to understand the structure-function relationship better, when obtaining structures experimentally proves difficult (Greener et al., 2019; Aderinwale et al., 2020).

Author contributions

ŠT, KG, and AC conceptualized the idea. ŠT performed *in silico* analyses and wrote the initial manuscript draft. All authors contributed to the writing and revision of the manuscript.

Funding

This work was supported by the Slovenian Research Agency through the program P4-0165, the projects J4-1777 and J1-2467, the program 1000-22-0105 for young researchers, in accordance with agreement on (co) financing research activity in 2022, as well as by the European Union's Horizon 2020 research and innovation program under grant agreement No. 862858.

Acknowledgments

We would like to thank Jan Zrimec for critical revision of the manuscript. Structural model visualization and analyses were performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.935819/full#supplementary-material

References

Abe, N., Dror, I., Yang, L., Slattery, M., Zhou, T., Bussemaker, H. J., et al. (2015). Deconvolving the recognition of DNA shape from sequence. *Cell* 161, 307–318. doi: 10.1016/j.cell.2015.02.008

Aderinwale, T., Christoffer, C. W., Sarkar, D., Alnabati, E., and Kihara, D. (2020). Computational structure modeling for diverse categories of macromolecular interactions. *Curr. Opin. Struct. Biol.* 64, 1–8. doi: 10.1016/j.sbi.2020.05.017

Alvarez, J. M., Riveras, E., Vidal, E. A., Gras, D. E., Contreras-López, O., Tamayo, K. P., et al. (2014). Systems approach identifies TGA1 and TGA4 transcription factors as important regulatory components of the nitrate response of *Arabidopsis thaliana* roots. *Plant J.* 80, 1–13. doi: 10.1111/tpj.12618

Arnold, C. D., Nemčko, F., Woodfin, A. R., Wienerroither, S., Vlasova, A., Schleiffer, A., et al. (2018). A high-throughput method to identify trans-activation domains within transcription factor sequences. *EMBO J.* 37:e98896. doi: 10.15252/embj.201798896

Backer, R., Naidoo, S., and van den Berg, N. (2019). The NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) and related family: mechanistic insights in plant disease resistance. *Front. Plant Sci.* 10:102. doi: 10.3389/fpls.2019.00102

Bentsink, L., Jowett, J., Hanhart, C. J., and Koornneef, M. (2006). Cloning of *DOG1*, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17042–17047. doi: 10.1073/pnas.0607877103

Bergey, D. R., Kandel, R., Tyree, B. K., Dutt, M., and Dhekney, S. A. (2014). The role of calmodulin and related proteins in plant cell function: an ever-thickening plot. *Springer Sci. Rev.* 2, 145–159. doi: 10.1007/s40362-014-0025-z

Boyle, P., Le Su, E., Rochon, A., Shearer, H. L., Murmu, J., Chu, J. Y., et al. (2009). The BTB/POZ domain of the *Arabidopsis* disease resistance protein NPR1 interacts with the repression domain of TGA2 to negate its function. *Plant Cell* 21, 3700–3713. doi: 10.1105/tpc.109.069971

Budimir, J., Treffon, K., Nair, A., Thurow, C., and Gatz, C. (2021). Redox-active cysteines in TGACG-BINDING FACTOR 1 (TGA1) do not play a role in salicylic acid- or pathogen-induced expression of TGA1-regulated target genes in *Arabidopsis thaliana*. *New Phytol.* 230, 2420–2432. doi: 10.1111/nph.16614

Canales, J., Contreras-López, O., Álvarez, J. M., and Gutiérrez, R. A. (2017). Nitrate induction of root hair density is mediated by TGA1/TGA4 and CPC transcription factors in *Arabidopsis thaliana*. *Plant J.* 92, 305–316. doi: 10.1111/tpj.13656

Chai, L.-X., Dong, K., Liu, S.-Y., Zhang, Z., Zhang, X.-P., Tong, X., et al. (2020). A putative nuclear copper chaperone promotes plant immunity in Arabidopsis. *J. Exp. Bot.* 71, 6684–6696. doi: 10.1093/jxb/eraa401

Chen, J., Mohan, R., Zhang, Y., Li, M., Chen, H., Palmer, I. A., et al. (2019). NPR1 promotes its own and target gene expression in plant defense by recruiting CDK8. *Plant Physiol.* 181, 289–304. doi: 10.1104/pp.19.00124

Chen, L.-J., Zou, W.-S., Fei, C.-Y., Wu, G., Li, X.-Y., Lin, H.-H., et al. (2018). α-Expansin *EXPA4* positively regulates abiotic stress tolerance but negatively regulates pathogen resistance in *Nicotiana tabacum*. *Plant Cell Physiol.* 59, 2317–2330. doi: 10.1093/pcp/pcy155

Deppmann, C. D., Acharya, A., Rishi, V., Wobbes, B., Smeekens, S., Taparowsky, E. J., et al. (2004). Dimerization specificity of all 67 B-ZIP motifs in *Arabidopsis thaliana*: a comparison to *Homo sapiens* B-ZIP motifs. *Nucleic Acids Res.* 32, 3435–3445. doi: 10.1093/nar/gkh653

Deppmann, C. D., Alvania, R. S., and Taparowsky, E. J. (2006). Cross-species annotation of basic leucine zipper factor interactions: insight into the evolution of closed interaction networks. *Mol. Biol. Evol.* 23, 1480–1492. doi: 10.1093/molbev/msl022

Després, C., Chubak, C., Rochon, A., Clark, R., Bethune, T., Desveaux, D., et al. (2003). The Arabidopsis NPR1 disease resistance protein is a novel cofactor that confers redox regulation of DNA binding activity to the basic domain/leucine zipper transcription factor TGA1. *Plant Cell* 15, 2181–2191. doi: 10.1105/tpc.012849

Ding, Y., Sun, T., Ao, K., Peng, Y., Zhang, Y., Li, X., et al. (2018). Opposite roles of salicylic acid receptors NPR1 and NPR3/NPR4 in transcriptional regulation of plant immunity. *Cell* 173, 1454–1467. doi: 10.1016/j.cell.2018.03.044

Dröge-Laser, W., Snoek, B. L., Snel, B., and Weiste, C. (2018). The *Arabidopsis* bZIP transcription factor family — an update. *Curr. Opin. Plant Biol.* 45, 36–49. doi: 10.1016/j.pbi.2018.05.001

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi: 10.1186/1471-2105-5-113

Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α Helices: crystal structure of the protein-DNA complex. *Cell* 71, 1223–1237. doi: 10.1016/S0092-8674(05)80070-4

Fan, W., and Dong, X. (2002). In vivo interaction between NPR1 and transcription factor TGA2 leads to salicylic acid-mediated gene activation in Arabidopsis. *Plant Cell* 14, 1377–1389. doi: 10.1105/tpc.001628

Fang, H., Liu, Z., Long, Y., Liang, Y., Jin, Z., Zhang, L., et al. (2017). The Ca2+/calmodulin2-binding transcription factor TGA3 elevates *LCD* expression and H2S production to bolster Cr6+ tolerance in Arabidopsis. *Plant J.* 91, 1038–1050. doi: 10.1111/tpj.13627

Findling, S., Stotz, H. U., Zoeller, M., Krischke, M., Zander, M., Gatz, C., et al. (2018). TGA2 signaling in response to reactive electrophile species is not dependent on cysteine modification of TGA2. *PLoS One* 13:e0195398. doi: 10.1371/journal.pone.0195398

Fode, B., Siemsen, T., Thurow, C., Weigel, R., and Gatz, C. (2008). The *Arabidopsis* GRAS protein SCL14 interacts with class II TGA transcription factors and is essential for the activation of stress-inducible promoters. *Plant Cell* 20, 3122–3135. doi: 10.1105/tpc.108.058974

Gatz, C. (2013). From pioneers to team players: TGA transcription factors provide a molecular link between different stress pathways. *Mol. Plant Microbe Interact.* 26, 151–159. doi: 10.1094/MPMI-04-12-0078-IA

Greener, J. G., Kandathil, S. M., and Jones, D. T. (2019). Deep learning extends de novo protein modelling coverage of genomes using iteratively predicted structural constraints. *Nat. Commun.* 10:3977. doi: 10.1038/s41467-019-11994-0

Gutsche, N., and Zachgo, S. (2016). The N-terminus of the floral Arabidopsis TGA transcription factor PERIANTHIA mediates redox-sensitive DNA-binding. *PLoS One* 11:e0153810. doi: 10.1371/journal.pone.0153810

Gutsche, N., Holtmannspötter, M., Maß, L., O'Donoghue, M., Busch, A., Lauri, A., et al. (2017). Conserved redox-dependent DNA binding of ROXY glutaredoxins with TGA transcription factors. *Plant Direct* 1, 1–16. doi: 10.1002/pld3.30

Hanson, J., Yang, Y., Paliwal, K., and Zhou, Y. (2017). Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks. *Bioinformatics* 33, 685–692. doi: 10.1093/bioinformatics/btw678

Herrera-Vásquez, A., Fonseca, A., Ugalde, J. M., Lamig, L., Seguel, A., Moyano, T. C., et al. (2021). TGA class II transcription factors are essential to restrict oxidative stress in response to UV-B stress in Arabidopsis. *J. Exp. Bot.* 72, 1891–1905. doi: 10.1093/jxb/eraa534

Huang, L.-J., Li, N., Thurow, C., Wirtz, M., Hell, R., and Gatz, C. (2016). Ectopically expressed glutaredoxin ROXY19 negatively regulates the detoxification pathway in *Arabidopsis thaliana*. *BMC Plant Biol.* 16:200. doi: 10.1186/s12870-016-0886-1

Hussain, R. M. F., Sheikh, A. H., Haider, I., Quareshy, M., and Linthorst, H. J. M. (2018). Arabidopsis WRKY50 and TGA transcription factors synergistically activate expression of *PR1*. *Front. Plant Sci.* 9:930. doi: 10.3389/fpls.2018.00930

Ishida, T., and Kinoshita, K. (2007). PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res.* 35, W460–W464. doi: 10.1093/nar/gkm363

Izawa, T., Foster, R., and Chua, N.-H. (1993). Plant bZIP protein DNA binding specificity. *J. Mol. Biol.* 230, 1131–1144. doi: 10.1006/jmbi.1993.1230

Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., et al. (2002). bZIP transcription factors in *Arabidopsis*. *Trends Plant Sci.* 7, 106–111. doi: 10.1016/S1360-1385(01)02223-3

Jindrich, K., and Degnan, B. M. (2016). The diversification of the basic leucine zipper family in eukaryotes correlates with the evolution of multicellularity. *BMC Evol. Biol.* 16:28. doi: 10.1186/s12862-016-0598-z

Johnson, C., Mhatre, A., and Arias, J. (2008). NPR1 preferentially binds to the DNA-inactive form of *Arabidopsis* TGA2. *Biochim. Biophys. Acta Gene Regul. Mech.* 1779, 583–589. doi: 10.1016/j.bbagrm.2008.05.007

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

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

Kang, H., and Klessig, D. F. (2005). Salicylic acid-inducible *Arabidopsis* CK2-like activity phosphorylates TGA2. *Plant Mol. Biol.* 57, 541–557. doi: 10.1007/s11103-005-0409-1

Katagiri, F., Lam, E., and Chua, N.-H. (1989). Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. *Nature* 340, 727–730. doi: 10.1038/340727a0

Katagiri, F., Seipel, K., and Chua, N.-H. (1992). Identification of a novel dimer stabilization region in a plant bZIP transcription activator. *Mol. Cell. Biol.* 12, 4809–4816. doi: 10.1128/mcb.12.11.4809-4816.1992

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kesarwani, M., Yoo, J., and Dong, X. (2007). Genetic interactions of TGA transcription factors in the regulation of pathogenesis-related genes and disease resistance in Arabidopsis. *Plant Physiol.* 144, 336–346. doi: 10.1104/pp.106.095299

Kim, Y.-W., Youn, J.-H., Roh, J., Kim, J.-M., Kim, S.-K., and Kim, T.-W. (2022). Brassinosteroids enhance salicylic acid-mediated immune responses by inhibiting BIN2 phosphorylation of clade I TGA transcription factors in *Arabidopsis*. *Mol. Plant* 15, 991–1007. doi: 10.1016/j.molp.2022.05.002

Krawczyk, S., Thurow, C., Niggeweg, R., and Gatz, C. (2002). Analysis of the spacing between the two palindromes of *activation sequence-1* with respect to binding to different TGA factors and transcriptional activation potential. *Nucleic Acids Res.* 30, 775–781. doi: 10.1093/nar/30.3.775

Kribelbauer, J. F., Rastogi, C., Bussemaker, H. J., and Mann, R. S. (2019). Low-affinity binding sites and the transcription factor specificity paradox in eukaryotes. *Annu. Rev. Cell Dev. Biol.* 35, 357–379. doi: 10.1146/annurev-cellbio-100617-062719

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Kumar, S., Zavaliev, R., Wu, Q., Zhou, Y., Cheng, J., and Dillard, L. (2022). Structural basis of NPR1 in activating plant immunity. *Nature* 605, 561–566. doi: 10.1038/s41586-022-04699-w

Lam, E., and Lam, Y. K.-P. (1995). Binding site requirements and differential representation of TGA factors in nuclear ASF-1 activity. *Nucleic Acids Res.* 23, 3778–3785. doi: 10.1093/nar/23.18.3778

Landschulz, W. H., Johnson, P. F., and Mcknight, S. L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* 240, 1759–1764. doi: 10.1126/science.3289117

Li, J., Sagendorf, J. M., Chiu, T. P., Pasi, M., Perez, A., and Rohs, R. (2017). Expanding the repertoire of DNA shape features for genome-scale studies of transcription factor binding. *Nucleic Acids Res.* 45, 12877–12887. doi: 10.1093/nar/gkx1145

Li, N., Muthreich, M., Huang, L.-J., Thurow, C., Sun, T., Zhang, Y., et al. (2019). TGACG-binding factors (TGAs) and TGA-interacting CC-type glutaredoxins modulate hyponastic growth in *Arabidopsis thaliana*. *New Phytol.* 221, 1906–1918. doi: 10.1111/nph.15496

Li, S., Gutsche, N., and Zachgo, S. (2011). The ROXY1 C-terminal L**LL motif is essential for the interaction with TGA transcription factors. *Plant Physiol.* 157, 2056–2068. doi: 10.1104/pp.111.185199

Li, S., Lauri, A., Ziemann, M., Busch, A., Bhave, M., and Zachgo, S. (2009). Nuclear activity of ROXY1, a glutaredoxin interacting with TGA factors, is required for petal development in *Arabidopsis thaliana*. *Plant Cell* 21, 429–441. doi: 10.1105/tpc.108.064477

Li, Y., and Loake, G. J. (2020). The immune-related, TGA1 redox-switch: to be or not to be? *New Phytol.* 1–3. doi: 10.1111/nph.16785

Lindermayr, C., Sell, S., Müller, B., Leister, D., and Durner, J. (2010). Redox regulation of the NPR1-TGA1 system of *Arabidopsis thaliana* by nitric oxide. *Plant Cell* 22, 2894–2907. doi: 10.1105/tpc.109.066464

Magnani, E., de Klein, N., Nam, H.-I., Kim, J.-G., Pham, K., Fiume, E., et al. (2014). A comprehensive analysis of microProteins reveals their potentially widespread mechanism of transcriptional regulation. *Plant Physiol.* 165, 149–159. doi: 10.1104/pp.114.235903

Maier, A. T., Stehling-Sun, S., Offenburger, S.-L., and Lohmann, J. U. (2011). The bZIP transcription factor PERIANTHIA: a multifunctional hub for meristem control. *Front. Plant Sci.* 2:79. doi: 10.3389/fpls.2011.00079

Mészáros, B., Erdős, G., and Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* 46, W329–W337. doi: 10.1093/nar/gky384

Mueller, S., Hilbert, B., Dueckershoff, K., Roitsch, T., Krischke, M., Mueller, M. J., et al. (2008). General detoxification and stress responses are mediated by oxidized lipids through TGA transcription factors in *Arabidopsis*. *Plant Cell* 20, 768–785. doi: 10.1105/tpc.107.054809

Murmu, J., Bush, M. J., DeLong, C., Li, S., Xu, M., Khan, M., et al. (2010). Arabidopsis basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. *Plant Physiol.* 154, 1492–1504. doi: 10.1104/pp.110.159111

Neuhaus, G., Neuhaus-Uri, G., Katagiri, F., Seipel, K., and Chua, N.-H. (1994). Tissue-specific expression of *as-1* in transgenic tobacco. *Plant Cell* 6, 827–834. doi: 10.1105/tpc.6.6.827

Niggeweg, R., Thurow, C., Weigel, R., Pfitzner, U., and Gatz, C. (2000). Tobacco TGA factors differ with respect to interaction with NPR1, activation potential and DNA-binding properties. *Plant Mol. Biol.* 42, 775–788. doi: 10.1023/A:1006319113205

Nishimura, N., Tsuchiya, W., Moresco, J. J., Hayashi, Y., Satoh, K., Kaiwa, N., et al. (2018). Control of seed dormancy and germination by DOG1-AHG1 PP2C phosphatase complex via binding to heme. *Nat. Commun.* 9:2132. doi: 10.1038/s41467-018-04437-9

Nishiyama, E., Nonogaki, M., Yamazaki, S., Nonogaki, H., and Ohshima, K. (2021). Ancient and recent gene duplications as evolutionary drivers of the seed maturation regulators DELAY OF GERMINATION1 family genes. *New Phytol.* 230, 889–901. doi: 10.1111/nph.17201

Noshi, M., Mori, D., Tanabe, N., Maruta, T., and Shigeoka, S. (2016). *Arabidopsis* clade IV TGA transcription factors, TGA10 and TGA9, are involved in ROS-mediated responses to bacterial PAMP flg22. *Plant Sci.* 252, 12–21. doi: 10.1016/j.plantsci.2016.06.019

O'Malley, R. C., Huang, S.-S. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., et al. (2016). Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165, 1280–1292. doi: 10.1016/j.cell.2016.04.038

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., et al. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. doi: 10.1002/pro.3943

Pontier, D., Privat, I., Trifa, Y., Zhou, J.-M., Klessig, D. F., and Lam, E. (2002). Differential regulation of TGA transcription factors by post-transcriptional control. *Plant J.* 32, 641–653. doi: 10.1046/j.1365-313x.2002.01461.x

Popescu, S. C., Popescu, G. V., Bachan, S., Zhang, Z., Seay, M., Gerstein, M., et al. (2007). Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4730–4735. doi: 10.1073/pnas.0611615104

Rochon, A., Boyle, P., Wignes, T., Fobert, P. R., and Després, C. (2006). The coactivator function of *Arabidopsis* NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. *Plant Cell* 18, 3670–3685. doi: 10.1105/tpc.106.046953

Rodríguez-Martínez, J. A., Reinke, A. W., Bhimsaria, D., Keating, A. E., and Ansari, A. Z. (2017). Combinatorial bZIP dimers display complex DNA-binding specificity landscapes. *Elife* 6:e19272. doi: 10.7554/eLife.19272

Sall, K., Dekkers, B. J. W., Nonogaki, M., Katsuragawa, Y., Koyari, R., Hendrix, D., et al. (2019). DELAY OF GERMINATION 1-LIKE 4 acts as an inducer of seed reserve accumulation. *Plant J.* 100, 7–19. doi: 10.1111/tpj.14485

Samee, M. A. H., Bruneau, B. G., and Pollard, K. S. (2019). A *de novo* shape motif discovery algorithm reveals preferences of transcription factors for DNA shape beyond sequence motifs. *Cell Syst.* 8, 27–42. doi: 10.1016/j.cels.2018.12.001

Sarkar, S., Das, A., Khandagale, P., Maiti, I. B., Chattopadhyay, S., and Dey, N. (2018). Interaction of *Arabidopsis* TGA3 and WRKY53 transcription factors on *Cestrum yellow leaf curling virus* (CmYLCV) promoter mediates salicylic acid-dependent gene expression in planta. *Planta* 247, 181–199. doi: 10.1007/s00425-017-2769-6

Schiermeyer, A., Thurow, C., and Gatz, C. (2003). Tobacco bZIP factor TGA10 is a novel member of the TGA family of transcription factors. *Plant Mol. Biol.* 51, 817–829. doi: 10.1023/A:1023093101976

Schindler, U., Beckmann, H., and Cashmore, A. R. (1992). TGA1 and G-box binding factors: two distinct classes of Arabidopsis leucine zipper proteins compete for the G-box-like element TGACGTGG. *Plant Cell* 4, 1309–1319. doi: 10.1105/tpc.4.10.1309

Seo, S. Y., Wi, S. J., and Park, K. Y. (2020). Functional switching of NPR1 between chloroplast and nucleus for adaptive response to salt stress. *Sci. Rep.* 10:4339. doi: 10.1038/s41598-020-61379-3

Shen, M., Lim, C. J., Park, J., Kim, J. E., Baek, D., Nam, J., et al. (2020). HOS15 is a transcriptional corepressor of NPR1-mediated gene activation of plant immunity. *Proc. Natl. Acad. Sci. U.S.A.* 117, 30805–30815. doi: 10.1073/pnas.2016049117

Sigrist, C. J. A., de Castro, E., Cerutti, L., Cuche, B. A., Hulo, N., Bridge, A., et al. (2013). New and continuing developments at PROSITE. *Nucleic Acids Res.* 41, D344–D347. doi: 10.1093/nar/gks1067

Song, Y. H., Song, N. Y., Shin, S. Y., Kim, H. J., Yun, D.-J., Lim, C. O., et al. (2008). Isolation of CONSTANS as a TGA4/OBF4 interacting protein. *Mol. Cells* 25, 559–565.

Staller, M. V., Holehouse, A. S., Swain-Lenz, D., Das, R. K., Pappu, R. V., and Cohen, B. A. (2018). A high-throughput mutational scan of an intrinsically disordered acidic transcriptional activation domain. *Cell Syst.* 6, 444–455. doi: 10.1016/j.cels.2018.01.015

Sun, T., Busta, L., Zhang, Q., Ding, P., Jetter, R., and Zhang, Y. (2018). TGACG-BINDING FACTOR 1 (TGA1) and TGA4 regulate salicylic acid and pipecolic acid biosynthesis by modulating the expression of SYSTEMIC ACQUIRED RESISTANCE DEFICIENT 1 (SARD1) and CALMODULIN-BINDING PROTEIN 60g (CBP60g). *New Phytol.* 217, 344–354. doi: 10.1111/nph.14780

Swift, J., and Coruzzi, G. (2017). A matter of time – How transient transcription factor interactions create dynamic gene regulatory networks. *Biochim. Biophys. Acta* 1860, 75–83. doi: 10.1016/j.bbagrm.2016.08.007

Szymanski, D. B., Liao, B., and Zielinski, R. E. (1996). Calmodulin isoforms differentially enhance the binding of cauliflower nuclear proteins and recombinant TGA3 to a region derived from the Arabidopsis *Cam-3* promoter. *Plant Cell* 8, 1069–1077. doi: 10.1105/tpc.8.6.1069

Thibaud-Nissen, F., Wu, H., Richmond, T., Redman, J. C., Johnson, C., Green, R., et al. (2006). Development of Arabidopsis whole-genome microarrays and their application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. *Plant J.* 47, 152–162. doi: 10.1111/j.1365-313X.2006.02770.x

Thurow, C., Schiermeyer, A., Krawczyk, S., Butterbrodt, T., Nickolov, K., and Gatz, C. (2005). Tobacco bZIP transcription factor TGA2.2 and related factor TGA2.1 have distinct roles in plant defense responses and plant development. *Plant J.* 44, 100–113. doi: 10.1111/j.1365-313X.2005.02513.x

Tian, W., Wang, C., Gao, Q., Li, L., and Luan, S. (2020). Calcium spikes, waves and oscillations in plant development and biotic interactions. *Nat. Plants* 6, 750–759. doi: 10.1038/s41477-020-0667-6

van der Krol, A. R., and Chua, N.-H. (1991). The basic domain of plant B-ZIP proteins facilitates import of a reporter protein into plant nuclei. *Plant Cell* 3, 667–675. doi: 10.1105/tpc.3.7.667

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al. (2022). AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50, D439–D444. doi: 10.1093/nar/gkab1061

Venturuzzi, A. L., Rodriguez, M. C., Conti, G., Leone, M., Caro, M. D. P., Montecchia, J. F., et al. (2021). Negative modulation of SA signaling components by the capsid protein of tobacco mosaic virus is required for viral long-distance movement. *Plant J.* 106, 896–912. doi: 10.1111/tpj.15268

Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science* 246, 911–916. doi: 10.1126/science.2683088

Wang, P., Nolan, T. M., Yin, Y., and Bassham, D. C. (2020). Identification of transcription factors that regulate *ATG8* expression and autophagy in *Arabidopsis*. *Autophagy* 16, 123–139. doi: 10.1080/15548627.2019.1598753

Wang, W., Withers, J., Li, H., Zwack, P. J., Rusnac, D.-V., Shi, H., et al. (2020). Structural basis of salicylic acid perception by *Arabidopsis* NPR proteins. *Nature* 586, 311–316. doi: 10.1038/s41586-020-2596-y

Wang, Y., Salasini, B. C., Khan, M., Devi, B., Bush, M., Subramaniam, R., et al. (2019). Clade I TGACG-motif binding basic leucine zipper transcription factors mediate BLADE-ON-PETIOLE-dependent regulation of development. *Plant Physiol.* 180, 937–951. doi: 10.1104/pp.18.00805

Wu, Y., Zhang, D., Chu, J. Y., Boyle, P., Wang, Y., Brindle, I. D., et al. (2012). The *Arabidopsis* NPR1 protein is a receptor for the plant defense hormone salicylic acid. *Cell Rep.* 1, 639–647. doi: 10.1016/j.celrep.2012.05.008

Xu, X., Xu, J., Yuan, C., Hu, Y., Liu, Q., Chen, Q., et al. (2021). Characterization of genes associated with TGA7 during the floral transition. *BMC Plant Biol.* 21:367. doi: 10.1186/s12870-021-03144-w

Yin, Z., Machius, M., Nestler, E. J., and Rudenko, G. (2017). Activator protein-1: Redox switch controlling structure and DNA-binding. *Nucleic Acids Res.* 45, 11425–11436. doi: 10.1093/nar/gkx795

Yip, K. M., Fischer, N., Paknia, E., Chari, A., and Stark, H. (2020). Atomic-resolution protein structure determination by cryo-EM. *Nature* 587, 157–161. doi: 10.1038/s41586-020-2833-4

Zander, M., Chen, S., Imkampe, J., Thurow, C., and Gatz, C. (2012). Repression of the *Arabidopsis thaliana* jasmonic acid/ethylene-induced defense pathway by TGA-interacting glutaredoxins depends on their C-terminal ALWL motif. *Mol. Plant* 5, 831–840. doi: 10.1093/mp/ssr113

Zander, M., La Camera, S., Lamotte, O., Métraux, J. P., and Gatz, C. (2010). *Arabidopsis thaliana* class-II TGA transcription factors are essential activators of jasmonic acid/ethylene-induced defense responses. *Plant J.* 61, 200–210. doi: 10.1111/j.1365-313X.2009.04044.x

Zhang, Y., Fan, W., Kinkema, M., Li, X., and Dong, X. (1999). Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the *PR-1* gene. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6523–6528. doi: 10.1073/pnas.96.11.6523

Zhang, Y., Tessaro, M. J., Lassner, M., and Li, X. (2003). Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell* 15, 2647–2653. doi: 10.1105/tpc.014894

Zhong, L., Chen, D., Min, D., Li, W., Xu, Z., Zhou, Y., et al. (2015). AtTGA4, a bZIP transcription factor, confers drought resistance by enhancing nitrate transport and assimilation in *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 457, 433–439. doi: 10.1016/j.bbrc.2015.01.009

Zhou, J.-M., Trifa, Y., Silva, H., Pontier, D., Lam, E., Shah, J., et al. (2000). NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the *PR-1* gene required for induction by salicylic acid. *Mol. Plant Microbe Interact.* 13, 191–202. doi: 10.1094/MPMI.2000.13.2.191

Zhou, M., Wang, W., Karapetyan, S., Mwimba, M., Marqués, J., Buchler, N. E., et al. (2015). Redox rhythm reinforces the circadian clock to gate immune response. *Nature* 523, 472–476. doi: 10.1038/nature14449

2.4 A Mini-TGA Protein Modulates Gene Expression through Heterogeneous Association with Transcription Factors

Špela Tomaž, Marko Petek, Tjaša Lukan, Karmen Pogačar, Katja Stare, Erica Teixeira Prates, Daniel A. Jacobson, Jan Zrimec, Gregor Bajc, Matej Butala, Maruša Pompe Novak, Quentin Dudley, Nicola Patron, Ajda Taler-Verčič, Aleksandra Usenik, Dušan Turk, Salomé Prat, Anna Coll and Kristina Gruden

Plant Physiology, 2022. Accepted under minor revisions.

This publication represents the first report focusing on the TGA transcription factors in potato and their role in plant immunity. Using phylogenetic and sequence analysis, we identified and characterized for the first time a potato TGA protein (StTGA2.1) with an unusually compact molecular architecture due to an extremely short N-terminus and a truncated bZIP domain. We referred to this type of TGAs as mini-TGAs, which are not found in Arabidopsis. Using functional and *in silico* structural analyses, we showed that StTGA2.1 modulates gene expression through association with full-length TGAs. We prepared SA-deficient potato plants overexpressing *StTGA2.1* and discovered that its overexpression is sufficient to attenuate PVY replication in these immunocompromised plants. Furthermore, we show that this mini-TGA is directly involved in the regulation of class III peroxidase gene expression. Finally, we employed computational modeling and simulations to understand better the structural aspects of StTGA2.1 contribution to transcriptional regulation.

The PhD candidate was involved in the initial conceptualization of the study and is the first author of this publication. She planned and performed *in silico* protein sequence and phylogenetic analyses, analyses of promoter sequences and the determination of DNA-binding motifs, cloning procedures, protein-protein interaction experiments in yeast and *in planta*, cell-free protein production, and plant immunity studies using transgenic plants overexpressing *StTGA2.1*. She planned and was actively involved in protein localization experiments, protein-DNA interaction studies, transactivation assays, targeted genomic sequencing, and gene expression analyses. She wrote the first draft of the manuscript and prepared most of the figures and tables.

A mini-TGA protein modulates gene expression through heterogeneous association with transcription factors

Running title: The role of mini-TGAs in plant immunity

Authors: Špela Tomaž[1,2,*], Marko Petek[1], Tjaša Lukan[1], Karmen Pogačar[1], Katja Stare[1], Erica Teixeira Prates[3], Daniel A. Jacobson[3], Jan Zrimec[1], Gregor Bajc[4], Matej Butala[4], Maruša Pompe Novak[1,5], Quentin Dudley[6], Nicola Patron[6], Ajda Taler-Verčič[7,8], Aleksandra Usenik[7,9], Dušan Turk[7,9], Salomé Prat[10], Anna Coll[1,**] & Kristina Gruden[1,**]

[1]National Institute of Biology, 1000 Ljubljana, Slovenia

[2]Jožef Stefan International Postgraduate School, 1000 Ljubljana, Slovenia

[3]Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

[4]Biotechnical Faculty, University of Ljubljana, 1000 Ljubljana, Slovenia

[5]University of Nova Gorica, School for Viticulture and Enology, 5271 Vipava, Slovenia

[6]Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ, United Kingdom

[7]Jožef Stefan Institute, 1000 Ljubljana, Slovenia

[8]Institute of Biochemistry and Molecular genetics, Faculty of Medicine, University of Ljubljana, 1000 Ljubljana, Slovenia

[9]Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, 1000 Ljubljana, Slovenia

[10]Centre for Research in Agricultural Genomics, 08193 Cerdanyola Barcelona, Spain

*Corresponding author: Š. Tomaž (spela.tomaz@nib.si), +386 (0)59 232 833

**Authors contributed equally and share the last authorship.

ABSTRACT

TGA transcription factors, which bind their target DNA through a conserved basic region leucine zipper (bZIP) domain, are vital regulators of gene expression in salicylic acid (SA)-mediated plant immunity. Here, we investigate the role of StTGA2.1, a potato TGA lacking the full bZIP, which we name a mini-TGA. Such truncated proteins have been widely assigned as loss-of-function mutants. We, however, confirm that *StTGA2.1* overexpression compensates for SA-deficiency, indicating a distinct mechanism of action compared with model plant species. To understand the underlying mechanisms, we show that StTGA2.1 can physically interact with StTGA2.2 and StTGA2.3, while its interaction with DNA was not detected. We investigate the changes in transcriptional regulation due to *StTGA2.1* overexpression, identifying direct and indirect target genes. Using *in planta* transactivation assays, we confirm that StTGA2.1 interacts with StTGA2.3 to activate *StPRX07*, a member of class III peroxidases, which are known to play role in immune response. Finally, via structural modelling and molecular dynamics simulations, we hypothesise that the compact molecular architecture of StTGA2.1 distorts DNA conformation upon heterodimer binding to enable transcriptional activation. This study demonstrates how protein truncation can lead to novel functions and that such events should be studied carefully in other protein families.

INTRODUCTION

Plants have developed efficient strategies to withstand the invasion of surrounding microbes. Pathogen recognition is mediated by plant cell-surface and intracellular receptors, triggering a cascade of intracellular reactions, orchestrated by phytohormones, ultimately leading to a finely modulated transcriptional reprogramming (Zhou and Zhang, 2020). Regulation of defense-related gene expression is among the most fundamental aspects of the immune response, involving multiple transcription factors and cofactor proteins. Since their initial discovery in tobacco over 30 years ago (Katagiri et al., 1989), the importance of TGACG-binding (TGA) transcription factors in plant immunity, as well as modulation of other cellular processes, has been widely studied (Gatz, 2013).

TGAs are a group of transcription factors belonging to the basic region leucine zipper (bZIP) protein family. Their mechanism of action has been thoroughly studied in *Arabidopsis thaliana*, where the ten Arabidopsis TGAs (AtTGAs) group into five clades (Jakoby et al., 2002). Clade II members, AtTGA2, AtTGA5, and AtTGA6, are essential regulators of the salicylic acid (SA)-mediated defense response, where they play a redundant, yet vital role in establishing resistance following infection (Zhang et al., 2003; Zhou and Zhang, 2020). They co-regulate the expression of key defense-related genes and genes involved in SA synthesis through interaction with NON-EXPRESSOR OF PR (NPR) cofactors (Zhang et al., 1999; Ding et al., 2018), while also participating in jasmonic acid and ethylene-mediated signaling (Zander et al., 2010). Structurally, TGAs consist of an intrinsically disordered N-terminus of varying length, a conserved bZIP domain, which entails a basic region and a leucine zipper, and a C-terminal region that contains a putative Delay of Germination 1 (DOG1) domain (Tomaž et al., 2022). TGAs bind their target DNA through the bZIP basic region, while the leucine zipper is important for protein dimerization (Thurow et al., 2005) and oligomerization (Boyle et al., 2009). The TGACG core sequence is sufficient for TGA binding, although high-

throughput DNA-binding studies revealed the TGACGTCA palindrome as the representative binding motif (Thibaud-Nissen et al., 2006; O'Malley et al., 2016).

The molecular mechanisms of TGA-mediated regulation involve complex interactions between TGAs and other proteins (Gatz, 2013). For example, the SA-receptor NPR1 interacts with AtTGA2 to activate the expression of the *Pathogenesis Related-1* (*PR-1*) gene expression (Zhang et al., 1999; Backer et al., 2019), but the mechanistic aspect of this cooperation is not yet entirely clear. Several studies suggest that AtTGA2 acts as a constitutive repressor of *PR-1* in absence of biotic stress (Zhang et al., 2003; Rochon et al., 2006; Kesarwani et al., 2007). Its repressive activity is then alleviated through NPR1 interaction with AtTGA2 N-terminus, affecting the binding stoichiometry and leading to the formation of a transcriptional activation complex (Rochon et al., 2006; Boyle et al., 2009). Additionally, other reports propose NPR1 interacts with TGAs not yet bound to DNA and indicate it could facilitate TGA binding to target promoter (Johnson et al., 2008). Furthermore, regulatory proteins, such as WRKY50 (Hussain et al., 2018) and histone acyltransferase (HAC) transcription factors (Jin et al., 2018), have also been shown to contribute to AtTGA2 transcriptional function.

Although the results obtained in Arabidopsis provide a molecular framework for understanding the role of TGAs in plant immunity, we know much less about their function in crops. The involvement of TGAs in biotic stress response has been reported in several species, including rice (Moon et al., 2018), soybean (Lawaju et al., 2018), strawberry (Feng et al., 2020), tobacco (Thurow et al., 2005), and tomato (Ekengren et al., 2003). Potato (*Solanum tuberosum* L.) is one of the most widely grown crops (FAO, 2020) and tuber production is severely threatened by pathogen infections. Several transcription factor families have been associated with the regulation of potato defense response (Chacón-Cerdas et al., 2020), but the mechanisms underlying potato TGA (StTGA) activity remain largely unexplored.

Here we identify the mini-TGA StTGA2.1, a potato clade II TGA, which lacks most of the bZIP DNA-binding domain and has a shorter N-terminus. We hypothesize that StTGA2.1 cannot bind DNA by itself because of the truncated bZIP and therefore modulates gene expression through its interaction with additional DNA-binding StTGAs. By combining *in vivo* and *in vitro* functional studies, we confirm the role of StTGA2.1 in potato immunity. Furthermore, using *in silico* structural analysis and molecular dynamics simulations, we provide insights into the molecular basis for a different mechanism of action of StTGA2.1 compared to other StTGAs.

100 **RESULTS**

101 **Potato encodes clade II TGAs with truncated bZIP domain**

102 In previous work, we investigated gene expression in response to viral infection in non-
103 transgenic resistant potato (NT) and its transgenic derivative (NahG), which is impaired in SA
104 accumulation and thus sensitive to infection (Baebler et al., 2014). To identify the TGA
105 transcription factors involved in potato immunity, we examined the expression patterns of the
106 fourteen StTGA genes, orthologues of AtTGAs (Supplementary Table 1 and Supplementary
107 Table 2). Notably, *Sotub10g022560* was up-regulated in infected NahG transgenic plants, but
108 not in the parental lines, suggesting that it may be an important component of SA signaling.

109 To classify the StTGAs, we conducted a phylogenetic analysis of all candidate potato
110 proteins, along with the ten AtTGAs and thirteen TGAs from tomato (SlTGAs) (Hou et al.,
111 2019; Lemaire-Chamley et al., 2022). Interestingly, the three clade II AtTGAs are
112 orthologous to five StTGAs and four SlTGAs (Fig. 1a). Three closely related members of this
113 clade, including *Sotub10g022560*, named StTGA2.1, StTGA2.4 (*Sotub10g022570*) and
114 SlTGA2.3 (*Solyc10g080780*) (Hou et al., 2019; Lemaire-Chamley et al., 2022) have shorter
115 protein sequences than other TGAs (Fig. 1b). Domain prediction studies showed that they
116 retain the putative C-terminal DOG1 domain, however, the bZIP domain is almost completely
117 lost, retaining only a partial zipper region. In addition, their N-terminus is very short and
118 dissimilar to the N-termini of other clade II TGAs. We named these three proteins mini-
119 TGAs.

120 By targeted sequencing of a ~36.5 kbp region on chromosome 10, where *StTGA2.1*, *StTGA2.2*
121 (*Sotub10g022550*) and *StTGA2.4* loci are co-located, we confirmed the reduced length of
122 *StTGA2.1* and *StTGA2.4* in a tetraploid cultivar that was used for further analyses
123 (Supplementary Fig. 1).

StTGA2.1 improves immune response in salicylic acid-deficient potato

SA signaling has proven vital for the establishment of an efficient defense response against potato virus Y (PVY) infection in resistant potato cultivars (Baebler et al., 2014; Lukan et al., 2020). We thus investigated the role of StTGA2.1 in plant immunity using the potato-PVY pathosystem. We generated SA-deficient NahG transgenic potato plants inducibly overexpressing *StTGA2.1* (TGA2.1-NahG) using the glucocorticoid-system (Aoyama and Chua, 1997), in which target gene expression is controlled by external application of dexamethasone (DEX). Three transgenic TGA2.1-NahG lines, showing more than 6-fold induction in *StTGA2.1* expression after DEX treatment (Supplementary Fig. 2a-c), were selected for further analysis. We observed that viral replication was significantly reduced in TGA2.1-NahG compared to NahG at 10 days post infection (dpi) (Fig. 2, Supplementary Fig. 2d-f). As expected, little to no PVY was detected in NT plants exhibiting a typical resistant phenotype (Baebler et al., 2014). This shows that overexpression of *StTGA2.1* can compensate for the lack of SA in potato immune response to PVY.

StTGA2.1 retains its dimerization ability and shows a distinct localization pattern

Protein interaction studies using the yeast two-hybrid assay showed that StTGA2.1 can form both homodimers and heterodimers with StTGA2.2 and StTGA2.3 (*Sotub01g009430*) (Fig. 3a), further confirmed by *in planta* co-immunoprecipitation assay (Fig. 3b, Supplementary Fig. 3a−c). Additionally, the size-exclusion chromatography elution volume of a recombinant His_6-tagged StTGA2.1 corresponded to the size of a dimer (Supplementary Fig. 3d,e), while chemical cross-linking of a non-tagged protein yielded monomers, dimers, and higher order complexes (Supplementary Fig. 3f). Overall, these results demonstrate that StTGA2.1 retains protein-protein interaction ability. In addition, we examined whether StTGA2.1 can interact with two potato NPR cofactors, an orthologue of AtNPR1, StNPR1 (*Sotub07g011600*), and

148 an orthologue of AtNPR3 and AtNPR4, StNPR3/4 (*Sotub02g015550*). Our results showed
149 that StTGA2.1 as well as StTGA2.2 and StTGA2.3 interact with both StNPRs in yeast and
150 that the addition of SA to the media promotes these interactions (Supplementary Fig. 4).
151 Thus, the ability to interact with NPR proteins is not perturbed in mini-TGA StTGA2.1.

152 Subcellular localization of GFP-tagged StTGA2.1 in the *N. benthamiana* leaf epidermis and
153 mesophyll showed that it can localize to cell nuclei (Fig. 3c). StTGA2.1 also showed a
154 distinct localization pattern with intense fluorescence in the cytoplasm, which was enhanced
155 around the chloroplasts (Supplementary Fig. 5a). We also detected its fluorescence in the ER
156 and in granular formations of about 0.5-1.0 μm in size (Supplementary Fig. 5). In contrast,
157 StTGA2.2 and StTGA2.3 showed predominantly nuclear localization and were organized into
158 subnuclear formations of different sizes within the nuclei (Fig. 3c, Supplementary Fig. 6).

159 **Identification of potential StTGA2.1 targets with spatial transcriptomic profiling**
160 To gain insight into the mini-TGA mechanism of action in plant immunity, we examined the
161 expression profile of NahG plants overexpressing *StTGA2.1*. By sampling tissue sections
162 containing lesions and their immediate surrounding area after PVY infection (Supplementary
163 Fig. 7a), we were able to follow transcriptomic changes in PVY-responding cells (Lukan et
164 al., 2020). RNA sequencing results showed a regulation of 217 genes due to *StTGA2.1*
165 overexpression in the NahG background (TGA2.1-NahG vs. NahG plants comparison,
166 Supplementary Table 3). However, over 1,800 genes were differentially expressed
167 exclusively in TGA2.1-NahG, when plants were exposed to pathogen infection
168 (Supplementary Fig. 7b). Technical validation of the RNA sequencing data by qPCR is shown
169 in Supplementary Table 4.

170 Gene Set Enrichment Analysis (Supplementary Table 5) enabled us to extract key differences
171 in gene expression at the level of processes or functionally related gene groups (BINs), as

172 they are defined by the MapMan ontology (Ramšak et al., 2014). Important immunity-related
173 or regulatory BINs, enriched in up- or down-regulated genes of PVY- vs. mock-inoculated
174 plants for all three genotypes, are listed in Table 1. Up-regulated genes enriched uniquely in
175 TGA2.1-NahG included isoprenoid metabolism-related genes, PHD finger and PHOR1
176 transcription factors, and peroxidases (Table 1). On the other hand, the C2C2(ZN) DOF
177 transcription factors were enriched in down-regulated genes in TGA2.1-NahG plants. PVY-
178 regulation of cytokinin and jasmonate metabolism was lost in TGA2.1-NahG compared with
179 the other two genotypes, as was the up-regulated expression of WRKY transcription factors
180 (Table 1).

181 **StTGA2.1 and StTGA2.3 activate the class III peroxidase StPRX07**

182 As the expression of several peroxidases was up-regulated after *StTGA2.1* overexpression in
183 PVY-infected NahG plants (Table 1, Supplementary Table 5), we recognized them as
184 potential direct targets of StTGA2.1. To test our hypothesis, we selected three class III
185 peroxidases (StPRX, Supplementary Table 6), *StPRX07* (*Sotub09g020950*), *StPRX15*
186 (*Sotub02g035680*), and *StPRX46* (*Sotub03g007840*), which were up-regulated in TGA2.1-
187 NahG compared with NahG (Supplementary Table 3 and Supplementary Table 7), for further
188 analysis. Analysis of their promoter regions revealed predicted TGA-binding motifs between
189 450 and 750 bp upstream of the transcription start site (Fig. 4a).

190 To investigate the ability of StTGAs to bind these motifs, we first tested whether the
191 StTGA2.3 and StTGA2.1 proteins could bind to four candidate DNA fragments from *StPRX*
192 promoter regions, PRX07p_1, PRX07p_2, PRX15p_1 and PRX46p (Fig. 4a), using surface
193 plasmon resonance. Titration of a recombinant StTGA2.3 over chip-immobilized PRX07p_1
194 and PRX07p_2 fragments, carrying the predicted TGA-binding motifs of the *StPRX07*
195 promoter, resulted in a dose-dependent increase in response, compared to reference (Fig. 4b).
196 Interaction with PRX15p_1 and PRX46p fragments was negligible (Fig. 4b). As predicted by

197 the absence of the basic region, we did not measure any interaction between the His_6-tagged
198 StTGA2.1 and the tested DNA (Supplementary Fig. 8a). These results suggest that StTGA2.3,
199 but not StTGA2.1, binds specifically to the TGA-binding motifs in the *StPRX07* promoter.
200 Furthermore, titration of StTGA2.3 premixed with StTGA2.1 over PRX07p_1 and PRX07p_2
201 fragments resulted in higher responses compared with StTGA2.3 alone, whereas this was not
202 the case for PRX15p_1 (Supplementary Fig. 8b). These results support the formation of a
203 StTGA2.1-StTGA2.3 complex at the *StPRX07* regulatory region.

204 Finally, we tested the ability of StTGA2.3 and StTGA2.1 to activate the *StPRX07* promoter *in*
205 *planta*, using a transient transactivation assay (Lasierra and Prat, 2018). For this purpose, the
206 2.95 kbp long promoter region upstream of the *StPRX07* start codon, containing both
207 predicted TGA-binding motifs, was fused to a luciferase (*LucF*) coding sequence. GFP-
208 tagged StTGA2.1 and BFP-tagged StTGA2.3 were then co-expressed with the reporter
209 construct, confirmed by confocal microscopy. Co-expression of the reporter construct with
210 StTGA2.3 induced the expression of *StPRX07*::*LucF* by approximately 20% compared with
211 basal promoter activity, whereas co-expression with StTGA2.1 resulted in only minor
212 induction (Fig. 4c, Supplementary Fig. 8c). In contrast, more than two-fold induction in
213 promoter activity was observed when co-expressed with both StTGA2.1 and StTGA2.3,
214 compared to control, meaning the induction was about three to four-times stronger when both
215 StTGAs were overexpressed (Fig. 4c, Supplementary Fig. 8c). These results indicate that
216 strong activation of *StPRX07* promoter is achieved only when both StTGA2.3 and StTGA2.1
217 are present.

StTGA2.1 N-terminus likely contributes to protein interactions and alters TGA binding to DNA

Comparative structural analysis using AlphaFold (AF) (Jumper et al., 2021) revealed important singularities in the molecular architecture of StTGA2.1, mostly contained in its N-terminus. In the AF models of StTGA2.2 and StTGA2.3, the intrinsically disordered N-terminus is followed by an α-helical bZIP domain, which includes several basic residues and three heptads that comprise the leucine zipper (Fig. 5a,b, Supplementary Fig. 9). In contrast, StTGA2.1 has a short N-terminus with low helical propensity due to the Pro25 α-helix breaker, harboring only the basic residues Arg20, Arg30, and Arg24, and lacking the first and most of the second leucine zipper heptad (Fig. 5a,b). The conserved hydrophobic residues Leu34, Val31, and Phe38 may contribute to protein dimerization by forming a partial zipper that could be stabilized by the hydrophobic Val28 and Val29. Persistent contacts identified in molecular dynamics (MD) simulations of StTGA2.2 and StTGA2.3 homo- and heterodimers are preserved in StTGA2.1 and involve the said hydrophobic residues (Fig. 5b, Supplementary Fig. 10).

We then inquired if the StTGA2.1 N-terminus binds to DNA. Based on prior knowledge of AtTGA2 and its cognate TGACG motif (Boyle et al., 2009), we modelled the DNA-bound StTGA2.2-StTGA2.2 and StTGA2.2-StTGA2.1 dimers and, via MD simulations, identified the key DNA-binding residues (Fig. 5c,d, Supplementary Table 8 and Supplementary Table 9). Persistent interactions in the StTGA2.2 homodimer mostly correspond to salt bridges, formed between the bZIP basic residues and the DNA phosphate groups, and do not explain the StTGA2.2 motif-binding specificity. In contrast, sequence specificity is provided by hydrophobic contacts between the StTGA2.2 Ala172 residue and the DNA T2 methyl group or between Ala171 and T-4 or T-5. Indeed, most of the predicted TGA-binding motifs (Fig. 4a) have a thymine or adenine in these positions. While the contacts involving StTGA2.2 in

the StTGA2.1-StTGA2.2 heterodimer are highly preserved, the StTGA2.1-DNA interactions are dramatically reduced, with only Arg11, Arg20, and Arg24 forming persistent contacts with the DNA phosphate groups (Fig. 5d, Supplementary Table 9). Moreover, the StTGA2.1 partial zipper positions the last basic residue (Arg24) more distant to the DNA compared to StTGA2.2 (Lys180), breaking the protein-DNA complex symmetry.

Another important singularity of StTGA2.1 is that its partial zipper connects directly to the putative DOG1 domain, while the two domains are connected by a 13 aa peptide linker in StTGA2.2 and StTGA2.3. This may greatly influence its interdomain conformational flexibility. Pro51 in StTGA2.1 (Ala227 and Ala110, in StTGA2.2 and StTGA2.3, respectively), disrupts the DOG1 α-helix 1 (α1) and forms a shorter, disordered linker that could somehow compensate for this absence. These results indicate that the compact molecular architecture of StTGA2.1, which causes an asymmetric distribution of basic residues in the StTGA2.1-StTGA2.2 heterodimer, significantly distorts the DNA conformation near the binding site, supporting strong promoter activation upon binding of the heterodimer compared with binding of the homodimer.

DISCUSSION

TGAs are involved in modulation of various cellular processes, acting as positive or negative regulators of gene expression (Gatz, 2013). Their structural features provide the basis for their functional variability, defining their subcellular localization, target recognition, DNA-binding, as well as their ability to form dimers, oligomers or interact with other proteins (Tomaž et al., 2022). In Arabidopsis, all ten AtTGAs share a highly conserved bZIP domain, essential for establishing specific interactions with DNA. Here, we report on the structural-functional relationship of a mini-TGA from potato, StTGA2.1, which lacks a full bZIP but still acts at target gene activation. Mini-TGAs were identified in potato (this study), tomato (Hou et al., 2019; Lemaire-Chamley et al., 2022), and strawberry (Feng et al., 2020), but not in Arabidopsis.

Homo- and heterodimerization of StTGA2.1 (Fig. 3a,b, Supplementary Fig. 3), which contains only a part of an already short and presumably unstable TGA zipper region (Deppmann et al., 2004), corroborates the findings of Boyle *et al.* (2009), who established that the leucine zipper is not crucial for dimerization of AtTGA2 (Boyle et al., 2009). Instead, StTGA2.1 dimerization is likely mediated by interactions involving its C-terminal region, previously reported to contain a dimer stabilization region (Katagiri et al., 1992). In line with this, the recent cryo-EM structure of an NPR1-AtTGA3 complex revealed the formation of stable homodimer contacts between two AtTGA3 C-terminal regions (Kumar et al., 2022). StTGA2.1 nuclear localization allows its role in gene regulation, however its unusually broad localization pattern (Fig. 3c, Supplementary Fig. 5) suggests that it might perform different tasks, as has been shown for other plant transcription factors with intrinsically disordered regions (Powers et al., 2019). TGAs also interact with different proteins present in both nuclei and cytoplasm, including NPR cofactors (Boyle et al., 2009; Ding et al., 2018), which was already confirmed for StTGA2.1 (Supplementary Fig. 4), glutaredoxins (Li et al., 2009), and

283 calmodulins (Popescu et al., 2007). Thus, StTGA2.1 could modulate the function of these
284 partners by sequestering them into inactive complexes, as was proposed for AtTGA2 and
285 NPR1 (Fan and Dong, 2002).

286 Multiple functions of StTGA2.1 are supported also by the diversity of detected transcriptional
287 changes following exposure to pathogen infection (Table 1, Supplementary Table 3 and
288 Supplementary Table 5). StTGA2.1 may directly affect the activity of several transcriptional
289 regulators and its effects on gene expression are likely amplified by further secondary
290 regulation. Thus, we hypothesize StTGA2.1 is an important player in shaping the
291 transcriptional landscape during infection. Interestingly, our results show StTGA2.1 improves
292 potato immunity even in the absence of SA, even though clade II TGAs are known for their
293 regulatory role in the SA pathway (Zhang et al., 2003; Zhou and Zhang, 2020). Our results
294 thus indicate that the TGA mechanism of action in potato may differ substantially from that in
295 Arabidopsis. It is important to note that *StTGA2.1* overexpression in SA-deficient plants may
296 not reflect its function in non-transgenic potato completely, as several parts of immune
297 signaling are under the control of SA and are thus not fully functional in our transgenic plants.
298 We have, however, focused on overexpression in a SA-deficient background to check whether
299 resistance towards PVY can be improved. Overexpression of *StTGA2.1* in a non-transgenic
300 background would not result in observable phenotypic differences, because the plants are
301 already fully resistant to PVY. Furthermore, as we did not find *StTGA2.1* transcriptionally
302 regulated during PVY infection in non-transgenic plants (Supplementary Table 2,
303 Supplementary Table 4 and Supplementary Table 7), the preparation of knockout non-
304 transgenic plants is unlikely to affect plant immunity.

305 Several studies have evaluated the influence of clade II TGA dominant-negative bZIP mutants
306 on plant immunity during bacterial infection, leading to contradictory results (Niggeweg et
307 al., 2000; Pontier et al., 2001; Fan and Dong, 2002). Reactive oxygen species (ROS) act as

signaling molecules in biotic stress (Bleau and Spoel, 2021). Early ROS production is central to plant defense and TGAs have previously been associated with cellular redox control, physically interacting with or regulating the expression of CC-type glutaredoxins (Ndamukong et al., 2007; Li et al., 2009; Hou et al., 2019). Furthermore, clade II TGAs modulate the expression of glutathione-S-transferases in ROS-processing responses to UV-B stress (Herrera-Vásquez et al., 2021), while clade IV AtTGAs are regulated by flg22-induced ROS production (Noshi et al., 2016). Here we show that the synergistic activity of TGAs regulates the expression of yet another group of enzymes involved in ROS-metabolism, the class III peroxidases (Fig. 4b,c, Supplementary Fig. 8). Class III peroxidases are haem-containing glycoproteins, secreted to the apoplast or localized in vacuoles (Almagro et al., 2009). Among them, AtPRX33 and AtPRX34 proved vital for apoplastic ROS production in response to flg22 and elf26 (Daudi et al., 2012). Most of the StPRX protein sequences from the peroxidase functional group (Ramšak et al., 2014) contain predicted secretory signal peptides (Supplementary Table 6), indicating StTGA2.1 could affect apoplastic ROS production in plant defense.

Transcription factor cooperativity is essential in eukaryotic transcription regulation and can arise through various mechanisms, involving protein-protein and/or protein-DNA interactions (Morgunova and Taipale, 2017). For example, the interaction between two NPR1 proteins bridges two AtTGA3 homodimers bound to separate DNA-binding motifs, creating a complex with an $AtTGA3_2$-$NPR1_2$-$AtTGA3_2$ stoichiometry (Kumar et al., 2022). Previous studies have shown that TGA mutants, impaired in DNA-binding through diverse modifications of the bZIP domain, prevent DNA-binding of wild type homologues (Rieping, 1994; Niggeweg et al., 2000; Pontier et al., 2001), which somewhat opposes the cooperative activation via an StTGA2.1-StTGA2.3 complex. Compared to homodimers of its paralogues, our molecular dynamics simulations suggest that the asymmetrical distribution of basic residues in the bZIP-

like domain in the StTGA2.1-StTGA2.2 heterodimer significantly distorts the DNA conformation near its binding site (Fig. 5). We hypothesize that StTGA2.1 dramatically affects the overall conformation of the regulatory complex due to its compact molecular architecture.

In conclusion, we show that, although mini-TGAs are not able to bind DNA on their own, their unusual structure supports diverse functionalities, such as allowing the induction of class III peroxidases in immune signaling. We thus provide evidence that truncation in evolution of genes does not necessary lead to a loss-of-function phenotype. Instead, additional functions can be attained. Through this, we shed additional perspective on immune signaling in non-model species, as Arabidopsis does not encode such proteins.

343 **MATERIALS AND METHODS**

344 ***In silico* sequence and structural analysis**

345 TGA transcription factor orthologues from potato were identified based on orthologue
346 information included in the GoMapMan database (Ramšak et al., 2014). The initial list was
347 further pruned based on protein sequence alignments created with Geneious Alignment in
348 Geneious Prime 2020.1.1 (https://www.geneious.com/) and BLAST results to exclude
349 technical errors of orthologue detection and sequencing. Identified StTGAs are listed in
350 Supplementary Table 1. Basic protein information was calculated using the ExPASy
351 ProtParam tool (Gasteiger et al., 2005). Protein domain prediction was performed with
352 ExPASy Prosite (de Castro et al., 2006). Protein sequences of SlTGAs (Hou et al., 2019;
353 Lemaire-Chamley et al., 2022) were retrieved from the Sol Genomics Network (Fernandez-
354 Pozo et al., 2015) and sequences of AtTGAs from The Arabidopsis Information Resource
355 (Berardini et al., 2015).

356 For the phylogenetic analysis, the sequences were aligned with MAFFT (Katoh and Standley,
357 2013), using the L-INS-I iterative refinement method, and the alignment used for a
358 maximum-likelihood phylogenetic tree construction in MEGA-X (Kumar et al., 2018), using
359 the Jones-Taylor-Thorton matrix-based model (Jones et al., 1992) and 1,000 bootstrap
360 repetitions. The rice OsTGA2.1 (Chern et al., 2001) protein sequence (Q7X993) was retrieved
361 from UniProtKB (https://www.uniprot.org/) and used as tree root. For sequence similarity
362 visualization, the protein sequences were aligned with Geneious Alignment in Geneious
363 Prime 2020.1.1 (https://www.geneious.com/). Potato peroxidases were identified with protein
364 sequence BLAST against the RedoxiBase database (Savelli et al., 2019) and the secretory
365 signal peptides were predicted with SignalP 5.0 (Almagro Armenteros et al., 2019)
366 (Supplementary Table 6). Predictions of transcription factor binding motifs in promoter

sequences were performed with TRANSFAC (Matys et al., 2006) and predictions of transcription start sites with TSSFinder (de Medeiros Oliveira et al., 2021).

Structural models of StTGA2.1, StTGA2.2, and StTGA2.3 were generated with AlphaFold (Jumper et al., 2021). The top-ranked models were selected. The VMD (Visual Molecular Dynamics, version 1.9.4a48) molecular visualization program was used for visual analysis and structural alignment of protein models.

Molecular dynamics simulations

The initial homo- and heterodimeric configurations of StTGA2.1-StTGA2.2, StTGA2.2-StTGA2.2 and StTGA2.2-StTGA2.3 N-terminal fragments were defined using the crystal structure of CREB bZIP-CRE (PDB id: 1DH3) as template (Schumacher et al., 2000). Corresponding amino acid changes to the template were done using the VMD psfgen plugin, preserving the coordinates of the backbone and C_β atoms. StTGA2.1, StTGA2.2 and StTGA2.3, are truncated, keeping the amino acids 1-43, 159-206, and 42-89, respectively. The N-termini of StTGA2.2 and StTGA2.3 are capped with N-methylamide and the C-termini of the three proteins with acetyl. For simulations of DNA-bound StTGA2.1-StTGA2.2 and StTGA2.2-StTGA2.2, the DNA fragment from the template crystal structure was kept and the nucleotides were modified using psfgen. The final DNA sequence corresponds to the TGACGT motif, complementary to the linker scan 5 element and its adjacent regions of Arabidopsis *PR-1* promoter *as-1*-like sequence (Lebel et al., 1998) (Fig. 5c,d). The crystal Mg^{2+} cation and the six coordinated water molecules were kept.

GROMACS-2020 (Abraham et al., 2015) was used to prepare inputs and run molecular dynamics (MD) simulations. The simulation boxes were generated as an octahedron, defining a solvation layer of 10 Å minimum thickness around the molecular complex. 0.15 M NaCl was used to establish electroneutrality. Protonation states were defined for pH 7.0. Amber

ff99SB (Lindorff-Larsen et al., 2010) and ff14SB9 (Maier et al., 2015) were used to describe the protein in the free and DNA-bound TGA dimers, respectively, and PARMBSC1 (Ivani et al., 2015) was used to describe the DNA. TIP4P-D (Piana et al., 2015) or SPC (Berendsen et al., 1981) was used to describe water molecules in the simulations of the free and DNA-bound TGA dimers, respectively. CHARMM-formatted topology and parameter files were converted to GROMACS input files using the VMD plugin TopoGromacs (Vermaas et al., 2016).

The MD simulations were performed on the Summit supercomputer at the Oak Ridge Leadership Computing Facility. Energy minimization was performed for all systems with steepest descent. Periodic electrostatic interactions were treated with the particle mesh Ewald method (Darden et al., 1993). LINCS (Hess et al., 1997) was used to constrain bonds involving hydrogen atoms.

Similar protocols of simulation were applied for the free and DNA-bound TGA dimers. Preceding the classical simulations, we performed long equilibration runs of 315 ns as part of our protocol adapted from the MD simulation-based method of structural refinement described by Heo *et al.* (2021). In this protocol, potential sampling is accelerated with hydrogen mass repartitioning and by applying fairly high temperatures. Weak position restraint potentials were applied for minimum bias and to compensate for the high thermal energy. Velocity Langevin dynamics was performed using a friction constant of 1 ps^{-1}. During the equilibration phase, position restraints applied to C_α atoms in the leucine heptads were gradually released and the temperature gradually increased, reaching the maximum of 360 K (Supplementary Table 10). After long sampling at 340 K and 320 K, a final phase of equilibration is conducted at 298.15 K, the temperature of the following production runs. During the final equilibration phase, flat-bottom harmonic restraint potentials were applied, using a force constant of 0.25 kcal/mol/$Å^2$ and a flat-bottom width of 4 Å. To adjust box size, part of the equilibration phase was conducted in the *NpT* ensemble, using the Berendsen

416 barostat (Berendsen et al., 1984) applying a compressibility of 4.5 x 10^{-5} bar^{-1} and a time
417 constant of 1.0 ps. In the final phase of equilibration, the atomic velocities were assigned from
418 a Maxwell-Boltzmann distribution using random numbers of seed. The production runs of
419 free and DNA-bound dimers consisted of five unbiased independent simulations of 128 ns
420 and 200 ns, respectively. The position restraint potential applied to Mg^{2+} and its coordinated
421 water molecules was kept during these simulations.

422 For simulation analysis, the VMD plugin Hbonds was used to compute hydrogen bond
423 statistics. The geometric criteria adopted are a cut-off of 3.0 Å for donor-acceptor distance
424 and 20 ° for acceptor-donor-H angle. In Fig. 5c,d, salt bridges and hydrogen bonds between
425 protein DNA occurring during more than 10% of the simulation time are shown. Persistent
426 contacts were identified using the VMD plugin Timeline. In Fig. 5c,d, amino acid residues
427 involved interactions or hydrophobic contacts persisting for more than 30% of the simulation
428 time are shown. Grace was used for plots (https://plasma-gate.weizmann.ac.il/Grace/).

429 **Plant material**

430 Potato (*Solanum tuberosum* L.) non-transgenic cultivar Rywal (NT) and Rywal-NahG
431 (NahG), a transgenic line impaired in SA accumulation due to salicylate hydroxylase
432 expression (Baebler et al., 2014), were used in this study. Plants were propagated from stem
433 node tissue cultures and transferred to soil two weeks after node segmentation, where they
434 were kept in growth chambers under controlled environmental conditions at 22/20 °C with a
435 long-day (16 h) photoperiod of light (light intensity 4000 lm/m^2) and 60-70% relative
436 humidity. Tobacco *Nicotiana benthamiana* plants were grown from seeds and kept in growth
437 chambers under the same conditions.

438 **DNA constructs**

439 Full-length coding sequences (cds) of StTGA2.1 (Genbank accession number OM569617),
440 StTGA2.2 (Genbank accession number OM569618), StTGA2.3 (Genbank accession number
441 OM569619), StNPR1 (Genbank accession number OM569620) and StNPR3/4 (Genbank
442 accession number OM569621) were amplified from potato cultivar Rywal cDNA and inserted
443 into the pJET1.2/blunt cloning vector using the CloneJET PCR Cloning Kit (Thermo
444 Scientific, USA), following the manufacturer's instructions.

445 The selected genes were subsequently cloned into pENTR D-TOPO vector using pENTR™
446 Directional TOPO® Cloning Kit (Invitrogen, USA) and recombined through LR reaction
447 using the Gateway® LR Clonase TM II Enzyme Mix (Invitrogen, USA) into several Gateway
448 destination vectors (VIB, Belgium). For co-immunoprecipitation experiments, localization
449 studies and transactivation assays, the StTGA2.1, StTGA2.2 and StTGA2.3 cds were inserted
450 into pH7FWG2 and pJCV52 expression vectors (Karimi et al., 2002) to produce proteins with
451 C-terminal enhanced green fluorescent protein (GFP) and hemagglutinin A (HA) fusions,
452 respectively. For transactivation assays, StTGA2.3 was fused with a C-terminal mTagBFP2
453 (from Addgene plasmid # 102638) (Stark et al., 2018) blue fluorescent protein (BFP) prior to
454 cloning into pENTR D-TOPO vector (Invitrogen, USA) and subsequently recombined into
455 the pK7WG2 vector (Karimi et al., 2002). A short linker of six Gly residues was introduced
456 between the StTGA2.3 and BFP sequence. BFP fused with an N7 nuclear localization signal
457 (Ghareeb et al., 2016) (N7-BFP) was recombined into pK7WG2 (Karimi et al., 2002) as
458 control.

459 For overexpression experiments, the StTGA2.1 cds was amplified with primers harboring
460 *XhoI* and *Spe*I restriction enzyme cleavage sites and inserted into the pTA7002 vector
461 (Aoyama and Chua, 1997), enabling glucocorticoid-inducible gene expression *in planta*,
462 through restriction-ligation cloning.

463 For the yeast two-hybrid assays, the cds of StTGA2.1, StTGA2.2, StTGA2.3, StNPR1 and
464 StNPR374 were amplified and inserted into the pGBKT7 (bait) yeast expression vector
465 through *in vivo* cloning with Matchmaker Gold Yeast Two-Hybrid System (Clontech, USA),
466 to produce proteins with an N-terminal Gal4 DNA-binding domain. StTGA2.1, StNPR1 and
467 StNPR3/4 were inserted also into the pGADT7 (prey) vector (Clontech, USA), to produce
468 proteins with an N-terminal Gal4 activation domain, using the same cloning system.

469 Promoter sequences of *StPRX07* (Genbank accession number OM569622), *StPRX15*
470 (Genbank accession number OM569623) and *StPRX46* (Genbank accession number
471 OM569624) were amplified from potato cultivar Rywal genomic DNA and inserted into the
472 pENTR D-TOPO vector (Invitrogen, USA). The *StPRX07* promoter sequence was
473 subsequently recombined through LR reaction into the pGWB435 Gateway vector (Nakagawa
474 et al., 2007), as described above, inserting the promoter upstream of a luciferase reporter
475 (*LucF*).

476 For recombinant protein production, the StTGA2.1 cds was inserted into the pMCSG7
477 bacterial expression vector (Eschenfeldt et al., 2009) by ligation-independent cloning
478 (Aslanidis and Jong, 1990) to produce a protein with an N-terminal hexahistidine (His_6) tag.
479 The cds of StTGA2.3 was amplified using primers, enabling the digestion-ligation reaction
480 with the *Bsa*I restriction enzyme. Three silent mutations were introduced into its sequence, to
481 remove two native *Bsa*I restriction sites. The amplified fragment was subsequently ligated
482 into the pEPQD0KN0025 acceptor backbone (Addgene plasmid #162283) (Dudley et al.,
483 2021), together with pEPQD0CM0030 (Addgene plasmid #162312) (Dudley et al., 2021),
484 which adds an additional GS peptide to the protein C-terminus.

485 All primer pairs used in the cloning procedure are listed in Supplementary Table 11.
486 Sequence verification was performed with Sanger sequencing (Eurofins Genomics,
487 Germany).

488 **Transient expression assays**

489 Homemade electrocompetent *Agrobacterium tumefaciens* GV3101 cells were transformed
490 with prepared constructs by electroporation. Transformants were used for agroinfiltration of
491 the bottom three fully developed leaves of 3-4 week old *N. benthamiana* plants, as described
492 previously (Lazar et al., 2014). In cases of co-transformation with agrobacteria carrying
493 different constructs, the 1:1 ratio was applied. An equal volume of agrobacteria carrying *p19*
494 silencing suppressor (kindly provided by prof. Jacek Hennig, PAS, Poland) was added to the
495 mixture. Agrobacteria carrying *p19* only were used as controls.

496 **Confocal microscopy**

497 Protein fluorescence was visualized three to five days after transient *N. benthamiana*
498 transformation. For protein localization, the Leica TCS SP5 laser scanning confocal
499 microscope mounted on a Leica DMI 6000 CS inverted microscope with an HC PL
500 FLUOTAR 10x objective or HCX PL APO lambda blue 63.0x1.40 oil-immersion objective
501 (Leica Microsystems, Germany) was used, using the settings described previously (Lukan et
502 al., 2018b). The red Histone 2B-mRFP1 (H2B-RFP) nuclear marker (Federici et al., 2012)
503 was used to visualize cell nuclei. For co-immunoprecipitation and transactivation assays, the
504 protein fluorescence was confirmed with the Leica TCS LSI macroscope with Plan APO 5x
505 and 20x objectives (Leica Microsystems, Germany), using the settings described previously
506 (Lukan et al., 2018a). The green, blue or red fluorescent protein fluorescence was excited
507 using 488 nm, 405 nm and 543 nm laser lines, respectively. The emission was measured in
508 the window of 505-520 nm for GFP, 450-465 nm for BFP, 570-630 nm for H2B-RFP and

690-750 nm for autofluorescence. The Leica LAS AF Lite software (Leica Microsystems, Germany) was used for image processing.

Yeast two-hybrid assay

Bait (containing *StTGA2.1*, *StTGA2.2*, *StTGA2.3*, *StNPR1* or *StNPR3/4* cds), and prey (containing *StTGA2.1*, *StNPR1* or *StNPR3/4* cds) construct combinations were transformed into the Y2H Gold strain using the Matchmaker Gold Yeast Two-Hybrid System (Clontech, USA) and the transformants selected on control SD media without Leu and Trp (-L-W). Interactions were analyzed on selection SD media without Leu, Trp, His and adenine, with added X-α-Gal and Aureobasidin A (-L-W-H-A+Xgal+Aur). The proteins were tested for autoactivation through co-transformation of bait constructs with an empty prey vector. To evaluate the strength of interaction, saturated yeast culture dilutions (10^{-1}, 10^{-2} and 10^{-3}) were spotted onto selection media. To evaluate the effect of SA on the strength of interaction, the dilutions were spotted onto selection media containing 0.1 mM or 1.0 mM SA.

Co-immunoprecipitation assay

HA or GFP-tagged StTGA2.1, StTGA2.2 and StTGA2.3 were transiently expressed in *N. benthamiana* leaves in different combinations. The empty pB7WGF2 vector (Karimi et al., 2002), expressing the GFP protein, was used as control. The fluorescence of GFP and GFP-tagged proteins was confirmed with confocal microscopy after 4 days. Total proteins were extracted from ~500 mg leaf material with immunoprecipitation (IP) buffer, containing 10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1 mM dithiothreitol and 1x EDTA-free Protease Inhibitor Cocktail (Roche, Switzerland), followed by 1 h incubation with GFP-Trap® Magnetic Agarose beads (ChromoTek, Germany) at 4 °C. The beads were washed three times with IP buffer and eluted into SDS-PAGE loading buffer, containing 100 mM Tris-HCl, pH 6.8, 4% (w/v) SDS, 0.2% bromophenol blue, 20% (v/v) glycerol and 200 mM

533 dithiothreitol. The immunoprecipitated proteins and protein extracts were analyzed by SDS-
534 PAGE and Western blot, using anti-GFP (diluted 1:3.000 or 1: 5.000, Invitrogen, USA) and
535 anti-HA (diluted 1:1.000, ChromoTek, Germany) antibodies.

536 **Generation of StTGA2.1 overexpression plants**

537 Transgenic NahG-TGA2.1 plants were obtained by stable transformation of the Rywal-NahG
538 potato genotype (Baebler et al., 2014). Electrocompetent *A. tumefaciens* strain LBA4404 was
539 electroporated with the pTA7002 vector (Aoyama and Chua, 1997) carrying the *StTGA2.1*
540 cds, as described above. Agrobacteria were used for stable transformation of sterile plantlet
541 stem internodes from tissue culture, as described previously (Lukan et al., 2022). Plantlets
542 grown on regeneration media plates with hygromycin selection were sub-cultured in order to
543 generate independent transgenic lines. Transgenic lines were confirmed with PCR
544 (Supplementary Table 11). Lines 7, 12 and 13 were selected for further analysis.

545 **Virus inoculation and plant treatments**

546 Three to four weeks old potato plants were inoculated with GFP-tagged infectious PVY
547 clones PVY^{N605}-GFP (Rupar et al., 2015) or PVY^{N605}(123)-GFP (Lukan et al., 2022) or mock
548 inoculum, as described previously (Baebler et al., 2009). To induce gene overexpression,
549 plants were treated with dexamethasone (DEX) foliar spray solution containing 30 μM DEX
550 and 0.01 % (v/v) Tween-20 or control spray solution without DEX (control), 3 h prior to virus
551 inoculation, 3 h after virus inoculation and every day post inoculation until sampling.

552 **Gene expression analysis with qPCR**

553 For gene expression analysis, total RNA isolation, reverse transcription and qPCR were
554 performed as described previously (Baebler et al., 2014). DEX-induced *StTGA2.1*
555 overexpression in fully developed leaves of TGA2.1-NahG transgenic lines was confirmed 3
556 h after DEX treatment using a qPCR assay targeting *StTGA2.1* cds. The leaves of three DEX-

557 treated plants and two or three non-treated plants were sampled, one leaf per plant. For PVY
558 abundance analysis, PVY-infected leaves of DEX-treated TGA2.1-NahG, NahG and NT
559 genotypes were sampled at 3, 5 and 7 dpi or 3, 7 and 10 dpi. For each genotype and treatment,
560 three plants were analyzed, sampling one leaf per plant per dpi. PVY abundance and
561 *StTGA2.1* expression were quantified using two sample dilutions and a relative standard curve
562 method by normalization to the endogenous control *StCOX1* with quantGenius
563 (http://quantgenius.nib.si) (Baebler et al., 2017). A two-tailed *t*-test was used to compare
564 treatments, when applicable. The qPCR analysis was performed for both TGA2.1-NahG
565 transgenic lines.

566 RNA sequencing results were validated technically and biologically with qPCR, as described
567 above. For technical validation, the expression of *StACX3*, *StCS*, *StPti5, StPRX28* and
568 *StTGA2.1* was followed. Biological validation was performed in an independent experiment
569 repetition with both TGA2.1-NahG transgenic lines and following gene expression of
570 *StPRX07*, *StPRX15*, *StPRX46*, *StTGA2.1* and *PVY*. *StCOX1* and *StEF-1* were used for
571 normalization in both cases, as described above. A two-tailed *t*-test was used to compare
572 treatments, when applicable.

573 All primers and probes used for qPCR analysis together with the target gene IDs are listed in
574 Supplementary Table 12. New qPCR assays, targeting *StPRX07*, *StPRX15*, *StPRX46*,
575 *StTGA2.1* and *StPti5* were designed with Primer Express v2.0 (Applied Biosystems, USA),
576 using the sequences from the potato reference genome (Xu et al., 2011), cultivar Rywal cds,
577 and cultivar Rywal and cultivar Désirée reference transcriptomes (Petek et al., 2020).

578 **RNA sequencing analysis**

579 For RNA sequencing, 2-25 early visible lesions and their immediate surroundings were
580 sampled from PVY-inoculated leaves of DEX-treated TGA2.1-NahG, NahG and NT plants

581 and control-treated TGA2.1-NahG plants at 4 dpi, as described previously (Lukan et al.,
582 2020). About 20-30 sections of comparable size were sampled from mock-inoculated leaves
583 as controls. Three plants per genotype per treatment were analyzed, pooling together all
584 lesions or mock-sections from one leaf per plant. Total RNA was isolated as described
585 previously (Lukan et al., 2020). Strand-specific library preparation and sequencing were
586 performed by Novogene (HongKong), using the NovaSeq platform (Illumina) to generate
587 150-bp paired-end reads. Read quality control was performed using FastQC (Andrews, 2010).
588 The presence of contaminant organism reads was determined using Centrifuge (Kim et al.,
589 2016). Reads were mapped to the reference group Phureja DM1-3 potato genome v4.04 (Xu
590 et al., 2011) using the merged PGSC and ITAG genome annotation (Petek et al., 2020) and
591 counted using STAR (Dobin et al., 2013) with default parameters. Differential expression
592 analysis was performed in R using the limma package (Smyth et al., 2018). Raw and
593 normalized read counts as well as a processed data table were deposited at GEO under
594 accession number GSE196078. Genes with Benjamini-Hochberg FDR adjusted p-values <
595 0.05 and $|\log_2 FC| \leq -1$ were considered statistically significantly differentially expressed. The
596 Venn diagram was drawn, according to results obtained with the Gene List Venn Diagram
597 tool (http://genevenn.sourceforge.net/).

598 Gene Set Enrichment Analysis (Subramanian et al., 2005) was performed using non-filtered
599 normalized counts to search for regulated processes and functionally related gene groups,
600 altered significantly by virus inoculation in different genotypes (FDR corrected q-value <
601 0.05) using MapMan ontology (Ramšak et al., 2014) as the source of gene groups.

602 **Targeted genomic sequencing**

603 Genomic DNA was isolated from potato cultivar Rywal leaves using the DNeasy Plant Mini
604 Kit (Qiagen, Germany). Two sets of primers were designed to target the region of interest
605 (Supplementary Fig. 1a, Supplementary Table 13). Droplet-based PCR-free target region

enrichment, library preparation using the SQK-LSK109 kit (Oxford Nanopore Technologies, United Kingdom) and long-read sequencing on the MinION platform using the R9.4.1-type flow cell was performed by Samplix (Denmark). Nanopore read basecalling was performed using Guppy 4.2.2. The reads were error corrected with NECAT (Chen et al., 2021) setting GENOME_SIZE=100000000 and PREP_OUTPUT_COVERAGE=20000. Chimeric reads were split using Pacasus (Warris et al., 2018) and all reads designated as "passed" were mapped to the group Phureja DM1-3 potato genome v6.1 (Pham et al., 2020) using Minimap2 (Li, 2018). The obtained BAM file was indexed and sorted using SAMtools (Danecek et al., 2021). Raw Nanopore reads were deposited at SRA under accession number PRJNA803339.

Transactivation assay

GFP-tagged StTGA2.1, BFP-tagged StTGA2.3 and their combination were transiently expressed in *N. benthamiana* leaves, with N7-BFP and either a GFP-tagged SNF-related serine/threonine-protein kinase (StSAPK8) or an empty pH7FWG2 vector (Karimi et al., 2002) as controls. Protein fluorescence was confirmed with confocal microscopy after 3-5 days. The transactivation assays were performed as described previously (Lasierra and Prat, 2018). In brief, 0.5-cm-diameter leaf discs were sampled at 4 dpi and pre-incubated in MS liquid media with 35 μM D-luciferin substrate for 4 hours before analysis. Luminescence was measured in 10 min intervals with Centro LB963 Luminometer (Berthold Technologies, Germany). Seventeen to 18 leaf discs per construct combination were analyzed. The experiment was repeated twice.

Protein production, purification, characterization and antibody preparation

For recombinant production of His_6-tagged StTGA2.1, *Escherichia coli* BL21(DE3) cells were transformed with the pMCSG7 vector (Eschenfeldt et al., 2009) carrying the *StTGA2.1* cds, grown overnight and subsequently transferred to the liquid auto-induction media

(Studier, 2005), where they were incubated for 4 h at 37 °C and further 20 h at 20 °C to produce the protein. Cells were harvested by centrifugation, lysed and the protein purified by nickel affinity chromatography using the His-Trap HP column coupled with size-exclusion chromatography (SEC) using the HiPrep 26/60 Sephacryl S-200 column (GE Healthcare Life Sciences, UK). The protein was eluted into a buffer containing 30 mM Tris, pH 7.5, and 400 mM NaCl, and used for rabbit polyclonal anti-StTGA2.1 antibody preparation, provided by GenScript (USA).

The protein oligomeric state was determined based on SEC elution volume and Gel Filtration LWM Calibration Kit (standard sizes: conalbumin 75 kDa, ovalbumin 44 kDa, carbonic anhydrase 29 kDa, ribonuclease A 13.7 kDa and aprotinin 6.5 kDa, GE Healthcare, USA). Additionally, the His_6-tag was removed by His_6-tagged TEV protease cleavage and a secondary nickel affinity chromatography followed by an additional SEC, as well as an anion-exchange chromatography purification step. Chemical crosslinking was performed after His_6-tag removal, for which the protein buffer was exchanged to 30 mM Hepes, pH 7.5, 400 mM NaCl using ultrafiltration with Amicon Ultra centrifugal filter units (Merck, Germany). The reaction was performed using the BS^3 crosslinker according to the manufacturer's instructions (Thermo Scientific, USA) and the protein oligomeric state evaluated by SDS-PAGE.

The *E. coli* cell-free protein synthesis (CFPS) was used for the production of StTGA2.3. All CFPS reactions (total volume 30 or 75 µL) were performed as described previously (Dudley et al., 2021), with 20-24h incubation at 16, 20 or 25 °C. Either the empty pEPQD0KN0025 vector (Addgene plasmid #162283) (Dudley et al., 2021) or water was added to the reagent mixture to prepare a CFPS components reference. Proteins were detected by SDS-PAGE and Western blot, using anti-StTGA2.1 antibodies (diluted 1:4.000, GenScript, USA). Additionally, the protein identity was confirmed with mass spectrometry, performed at the

654 Department of Biochemistry and Molecular and Structural Biology at the Jožef Stefan
655 Institute (Slovenia).

656 **Surface plasmon resonance**

657 Surface plasmon resonance measurements were performed on Biacore T200 (GE Healthcare,
658 USA) at 25 °C at the Infrastructural Centre for Analysis of Molecular Interactions, University
659 of Ljubljana (Slovenia). To prepare the DNA, the PRX07p_1, PRX07p_2, PRX15p_1 and
660 PRX46p complementary primers (Integrated DNA Technologies, Belgium, Supplementary
661 Table 14) were mixed in a 2:3 molar ratio (long:short primers) and annealed by cooling the
662 mixtures from 95 °C to 4 °C. The resulting DNA fragments carried the selected 20 bp
663 promoter regions with a 15-nucleotide overhang that allowed hybridization with the
664 complementary biotinylated S1 primer (Caveney et al., 2019), immobilized on the
665 streptavidin sensor chip (GE Healthcare, USA). StTGA2.3, either alone or premixed with
666 StTGA2.1, protein-DNA binding experiments were performed in a running buffer containing
667 25 mM Tris, pH 7.4, 140 mM NaCl, 1 mM $MgCl_2$ and 0.005% P20. For StTGA2.1 DNA
668 binding assays, the running buffer contained 180 mM instead of 140 mM NaCl. Flow cell 1
669 was used as a reference and the DNA fragments were injected across the flow cell 2 at a flow
670 rate of 5 µL/min to immobilize 42-105 response units.

671 A kinetic titration approach was used to study the interactions between the CFPS-produced
672 StTGA2.3 protein, the CFPS components reference that lacked StTGA2.3 or the His_6-tagged
673 StTGA2.1 (18.75, 37.5, 75, 150, or 300 nM) and the DNA fragments. The highest
674 concentration of total protein (264 µg/mL) and four sequential 1.5-fold dilutions were used
675 for the CFPS-produced StTGA2.3 and the CFPS components reference. The proteins were
676 injected across DNA at five concentrations, with no dissociation time between protein
677 injections, at a flow rate of 30 µL/min. We used the multi cycle kinetics approach to study the
678 interaction between StTGA2.1 (300 nM) premixed with CFPS-produced StTGA2.3 (total

679 protein concentration of 130 µg/mL) with DNA fragments PRX07p_1, PRX07p_2 or
680 PRX15p_1. The proteins were injected over DNA at a flow rate of 30 µL/min with an
681 association time of 120 s and followed by dissociation for 300 s.

682 Regeneration of the sensor surface was performed with 50 mM NaOH solution for 10 s and
683 300 mM NaCl for 10 s at a flow rate of 30 µL/min. The sensorgrams for the StTGA2.3 and/or
684 the StTGA2.1 proteins were double subtracted for the response of the reference flow cell 1
685 and for the response of the CFPS components reference or of the running buffer, respectively.

686 **ACKNOWLEDGEMENTS**

687 We thank our colleagues Barbara Jaklič, Valentina Levak, Tjaša Mahkovec Povalej, Lidija
688 Matičič, Nastja Marondini and Rebecca Vollmeier for technical support and laboratory
689 assistance. We thank prof. Dr. Jim Haseloff and Dr. Fernán Federici (University of
690 Cambridge, UK) for providing the plasmid containing H2B-RFP. Template to amplify N7
691 was a kind gift from Prof. Volker Lipka (Georg-August-Universität, Göttingen). This work
692 was supported by the Slovenian Research Agency through the research core funding no. P4-
693 0165, the projects no. J4-1777 and J1-2467, the contract no. 1000-22-0105, in accordance
694 with the agreement on (co) financing research activity in 2022, as well as by the European
695 Union's Horizon 2020 research and innovation program under grant agreement no. 862858.
696 Funding was also provided by the Plant-Microbe Interfaces (PMI) Scientific Focus Area in
697 the Genomic Sciences Program of from the U.S. Department of Energy, Office of Science,
698 Office of Biological and Environmental Research. This manuscript has been co-authored by
699 UT-Battelle, LLC under contract no. DE-AC05-00OR22725 with the U.S. Department of
700 Energy. The United States Government retains and the publisher, by accepting the article for
701 publication, acknowledges that the United States Government retains a nonexclusive, paid-up,
702 irrevocable, world-wide license to publish or reproduce the published form of this manuscript,
703 or allow others to do so, for United States Government purposes. The Department of Energy
704 will provide public access to these results of federally sponsored research in accordance with
705 the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan, last
706 accessed September 16, 2020).

AUTHOR CONTRIBUTIONS

Š.T., A.C. and K.G. designed the research. Š.T., A.C., T.L., M.P., K.P., K.S., M.P.N., J.Z., Q.D., G.B. and E.T.P. performed the experiments and analyzed the data. Š.T. wrote the initial manuscript draft. All authors contributed to writing or revision of the manuscript.

711 **TABLES**

712 **Table 1. Selected functional groups (BINs) enriched in up- or down-regulated genes in**

713 **TGA2.1-NahG, NahG and NT plants after PVY infection.** FDR corrected q-value < 0.05.

714 (+), enriched in up-regulated genes; (−), enriched in down regulated genes.

BIN	Functional group	Size	TGA2.1-NahG	NahG	NT
	Secondary metabolism				
16.1.2	mevalonate pathway	47	+		
16.1.5	terpenoids	155	+		
16.10	simple phenols	38			+
	Hormone metabolism				
17.1.3	abscisic acid-regulated	32	−	−	
17.2	auxin	351	−	−	
17.2.3	auxin-regulated	287	−	−	
17.4.1	cytokinin	85		−	−
17.7	jasmonate	64		+	+
17.8	salicylic acid	22			+
	Stress – biotic				
20.1.7	PR proteins	208	+	+	+
20.1.7.1	PR-1	33	+		
20.1.7.3	PR-3/PR-4/PR-8/PR-11	44	+	+	+
	Miscellaneous				
26.12	peroxidases	139	+		
26.21	protease inhibitor/seed storage/lipid transfer proteins	117		−	
	Transcription regulation				
27.3.32	WRKY transcription factors	97		+	+
27.3.63	PHD finger transcription factors	49	+		
27.3.64	PHOR1 transcription factors	21	+		
27.3.8	C2C2(Zn) DOF transcription factors	41	−		
	DNA synthesis – chromatin structure				
28.1.3	histone	83	+	+	+
28.1.3.2	histone core	76	+	+	+
28.1.3.2.1	histone core H2A proteins	26	+	+	
28.1.3.2.3	histone core H3 proteins	20	+	+	+
	Protein degradation				
29.5.11.20	ubiquitin-proteasome	73	+	+	
	Signaling – receptor kinases				
30.2.8.1	leucine-rich repeat VIII (type 1)	20			−
30.2.16	*Catharanthus roseus*-like RLK1	79			+
30.2.17	DUF26	94			+
30.2.19	legume-lectin	38	+		+
30.2.99	miscellaneous	206	+		+

715

716 **FIGURE LEGENDS**

717 **Fig. 1. Phylogenetic analysis and domain characterization of StTGAs. a,** A rooted
718 phylogenetic tree of potato, tomato and Arabidopsis TGAs. The mini-TGA branch is shaded
719 in orange and StTGA2.1 is marked (arrow). The branch length scale represents the number of
720 amino acid substitutions per site. The rice OsTGA2.1 (Chern et al., 2001) serves as tree root.
721 **b,** Protein sequence alignment of clade II TGAs, showing the position of the bZIP domain
722 (orange box) and the shorter sequences of mini-TGA members, StTGA2.1, StTGA2.4 and
723 SlTGA2.3. The alignment is colored with the Geneious Prime (https://www.geneious.com/)
724 sequence similarity color scheme, based on the identity score matrix. Sequence numbering
725 (aa) is shown above the alignment.

726 **Fig. 2. StTGA2.1 attenuates PVY replication in salicylic acid-deficient plants.** Relative
727 expression levels of **a,** PVY and **b,** *StTGA2.1* in PVY-infected leaves of dexamethasone
728 (DEX)-treated TGA2.1-NahG line 12 (white), NahG (light grey) and NT (dark grey) plants at
729 3, 7 and 10 days post infection (dpi). Average values ± standard error from three biological
730 replicates are shown. Significance was determined using a two-tailed *t*-test comparing
731 TGA2.1-NahG with NahG (p_1) and TGA2.1-NahG with NT (p_2). **c,** Phenotypic differences in
732 PVY-infected leaves of DEX-treated TGA2.1-NahG line 12, NahG and NT plants at 10 dpi.

733 **Fig. 3. StTGA2.1 can form homodimers, heterodimers and localizes to diverse cellular**
734 **compartments. a**, StTGA2.1 interactions with itself, StTGA2.2, or StTGA2.3 in the yeast
735 two-hybrid assay. Yeast were co-transformed with bait (BD) and prey (AD) construct
736 combinations and selected on control media without Leu and Trp (-L-W). Positive
737 interactions were determined by yeast growth on selection media without Leu, Trp, His and
738 adenine, with added X-α-galactosidase and Aureobasidin A (-L-W-H-A+Xgal+Aur). **b**,
739 StTGA2.1 interactions with itself, StTGA2.2, or StTGA2.3 in the co-immunoprecipitation
740 assay. The combination of GFP and HA-tagged proteins expressed in *N. benthamiana* is

741 indicated for each sample (+/−). Positive interactions were determined by detection of
742 immunoprecipitated (GFP-IP) complexes with anti-HA antibodies. Arrows indicate expected
743 bands. In case of HA-tagged StTGA2.3 two bands were detected, likely due to partial protein
744 degradation. Controls are shown in Supplementary Fig. 3. **c**, Subcellular localization of GFP-
745 tagged StTGA2.1, StTGA2.2 and StTGA2.3 (yellow) with H2B-RFP nuclear marker (red) in
746 *N. benthamiana* leaves. The *p19* silencing suppressor was expressed as control. Protein
747 fluorescence is represented as the z-stack maximum projection. Arrows indicate examples of
748 StTGA2.1 nuclear localization. Scale bars, 100 µm.

749 **Fig. 4**. **StTGA2.1, together with its interactor StTGA2.3, activates expression of**
750 **StPRX07**. **a**, TGA-binding motifs in selected *StPRX* promoters. The predicted motifs are
751 boxed and the nucleotides, differing from the core TGACG(T) sequence, its reverse
752 complement or the TGACGTCA palindrome, are underlined. Numbers indicate position
753 upstream of transcription start site. **b**, Surface plasmon resonance results, showing the
754 interaction between StTGA2.3 and chip-immobilized PRX15p_1, PRX46p, PRX07p_1 or
755 PRX07p_2 DNA fragments, bound to the chip at ~38, 41, 65, or 53 response units (RU),
756 respectively. Representative sensorgrams are shown. **c**, Transactivation assay results, showing
757 *in planta StPRX07* promoter activation by GFP-tagged StTGA2.1 (light orange), BFP-tagged
758 StTGA2.3 (dark orange) or a combination of both (black). BFP or GFP-tagged controls and
759 their combination (control) were used to detect the basal promoter activity (grey). Average
760 values ± standard error of 18 biological replicates in the first 10 h of measurement are shown.
761 Significance was determined using a two-tailed *t*-test and is shown below the response curve
762 for GFP-tagged StTGA2.1 (light orange), BFP-tagged StTGA2.3 (dark orange) or a
763 combination of both (black) compared with control. The experiment was repeated twice with
764 similar results (Supplementary Fig. 8c).

765 **Fig. 5. Comparative structural analysis and simulations of StTGA2.1 N-terminus**
766 **interactions with StTGA2.2 and StTGA2.3 bZIP domains. a**, Protein sequence alignment
767 of StTGA2.1 N-terminus with AtTGA2, StTGA2.2 and StTGA2.3 bZIP domains. The basic
768 region (orange box) and the leucine zipper heptads (grey dashed boxes) are indicated.
769 Conserved amino acids, in respect to the consensus sequence, are marked with dots.
770 StTGA2.1 contains hydrophobic residues (Val31 and Phe38) in two out of three Leu positions
771 in the heptads (grey) and has a completely conserved third heptad. **b**, Molecular architectures
772 of StTGA2.1 (orange) and StTGA2.3 (green). The StTGA2.3 bZIP domain (aa 40-95), and
773 StTGA2.1 N-terminus (aa 1-45) are shown. The C-terminal region is highly conserved
774 between StTGA2.1 (aa 47-240), StTGA2.2 (aa 222-446), and StTGA2.3 (aa 105-327). Amino
775 acid residues that are discussed in this study are represented as liquorice and labelled. Those
776 forming persistent contacts in the leucine zipper, according to molecular dynamics (MD)
777 simulations, are shown in bold. Basic amino acid residues that may contribute to DNA-
778 binding are depicted and labelled. Fully non-conservative substitution sites in the putative
779 DOG1 domains are also represented as liquorice. **c**, Representative snapshot of the MD
780 simulations of the DNA-bound StTGA2.2 homodimer and **d**, StTGA2.2-StTGA2.1
781 heterodimer. The DNA double helix is represented in violet. The DNA sequence is shown at
782 the bottom and the binding site core is underlined. A dot is used as a reference at the center of
783 the sequence for numbering the nucleotide residues. StTGA2.1 (orange) and StTGA2.2 (cyan)
784 are represented as cartoon. Salt bridges and hydrogen bonds between protein and DNA are
785 indicated with black dashed lines. Hydrophobic contacts are indicated as dotted lines. Amino
786 acid residues forming persistent interactions are labelled in bold. The presence of a
787 hexacoordinated Mg^{2+} was assumed (yellow sphere), based on its importance for CREB-bZIP
788 (Moll et al., 2002).

789 **REFERENCES**

790 **Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindah E** (2015) Gromacs:
791 High performance molecular simulations through multi-level parallelism from laptops to
792 supercomputers. SoftwareX **1–2**: 19–25
793 **Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak**
794 **S, von Heijne G, Nielsen H** (2019) SignalP 5.0 improves signal peptide predictions
795 using deep neural networks. Nat Biotechnol **37**: 420–423
796 **Almagro L, Gómez Ros L V., Belchi-Navarro S, Bru R, Ros Barceló A, Pedreño MA**
797 (2009) Class III peroxidases in plant defence reactions. J Exp Bot **60**: 377–390
798 **Andrews S** (2010) FastQC: a quality control tool for high throughput sequence data.
799 **Aoyama T, Chua NH** (1997) A glucocorticoid-mediated transcriptional induction system in
800 transgenic plants. Plant J **11**: 605–612
801 **Aslanidis C, Jong PJ de** (1990) Ligation-independent cloning of PCR products (LIC-PCR).
802 Nucleic Acids Res **18**: 6069–6074
803 **Backer R, Naidoo S, Berg N Van Den** (2019) The NONEXPRESSOR OF
804 PATHOGENESIS-RELATED GENES 1 (NPR1) and Related Family: Mechanistic
805 Insights in Plant Disease Resistance. Front Plant Sci **10**: 102
806 **Baebler Š, Krečič-Stres H, Rotter A, Kogovšek P, Cankar K, Kok EJ, Gruden K, Kovač**
807 **M, Žel J, Pompe-Novak M, et al** (2009) PVYNTN elicits a diverse gene expression
808 response in different potato genotypes in the first 12 h after inoculation. Mol Plant Pathol
809 **10**: 263–275
810 **Baebler Š, Svalina M, Petek M, Stare K, Rotter A, Pompe-Novak M, Gruden K** (2017)
811 QuantGenius: Implementation of a decision support system for qPCR-based gene
812 quantification. BMC Bioinformatics **18**: 276
813 **Baebler Š, Witek K, Petek M, Stare K, Tušek-Žnidarič M, Pompe-Novak M, Renaut J,**

814 **Szajko K, Strzelczyk-Zyta D, Marczewski W, et al** (2014) Salicylic acid is an
815 indispensable component of the Ny-1 resistance-gene-mediated response against Potato
816 virus Y infection in potato. J Exp Bot **65**: 1095–1109
817 **Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E** (2015) The
818 Arabidopsis Information Resource: Making and mining the "gold standard" annotated
819 reference plant genome. Genesis **53**: 474–485
820 **Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR** (1984) Molecular
821 dynamics with coupling to an external bath. J Chem Phys **81**: 3684–3690
822 **Berendsen HJC, Postma JPM, Van Gunsteren WF, Hermans J** (1981) Interaction Models
823 for Water in Relation to Protein Hydration. *In* B Pullman, ed, Intermol. Forces. Reidel,
824 Dordrecht, pp 331–342
825 **Bleau JR, Spoel SH** (2021) Selective redox signaling shapes plant-pathogen interactions.
826 Plant Physiol **186**: 53–65
827 **Boyle P, Le Su E, Rochon A, Shearer HL, Murmu J, Chu JY, Fobert PR, Despres C**
828 (2009) The BTB/POZ Domain of the Arabidopsis Disease Resistance Protein NPR1
829 Interacts with the Repression Domain of TGA2 to Negate Its Function. Plant Cell **21**:
830 3700–3713
831 **de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger**
832 **E, Bairoch A, Hulo N** (2006) ScanProsite: Detection of PROSITE signature matches
833 and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res
834 **34**: 362–365
835 **Caveney NA, Pavlin A, Caballero G, Bahun M, Hodnik V, de Castro L, Fornelos N,**
836 **Butala M, Strynadka NCJ** (2019) Structural Insights into Bacteriophage GIL01 gp7
837 Inhibition of Host LexA Repressor. Structure **27**: 1094–1102
838 **Chacón-Cerdas R, Barboza-Barquero L, Albertazzi FJ, Rivera-Méndez W** (2020)

839 Transcription factors controlling biotic stress response in potato plants. Physiol Mol
840 Plant Pathol **112**: 101527
841 **Chen Y, Nie F, Xie SQ, Zheng YF, Dai Q, Bray T, Wang YX, Xing JF, Huang ZJ, Wang**
842 **DP, et al** (2021) Efficient assembly of nanopore reads via highly accurate and intact
843 error correction. Nat Commun **12**: 60
844 **Chern MS, Fitzgerald HA, Yadav RC, Canlas PE, Dong X, Ronald PC** (2001) Evidence
845 for a disease-resistance pathway in rice similar to the NPR1-mediated signaling pathway
846 in Arabidopsis. Plant J **27**: 101–113
847 **Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A,**
848 **Keane T, McCarthy SA, Davies RM, et al** (2021) Twelve years of SAMtools and
849 BCFtools. Gigascience **10**: giab008
850 **Darden T, York D, Pedersen L** (1993) Particle mesh Ewald: An Nlog(N) method for Ewald
851 sums in large systems. J Chem Phys **98**: 10089–10092
852 **Daudi A, Cheng Z, O'Brien JA, Mammarella N, Khan S, Ausubel FM, Paul Bolwell G**
853 (2012) The apoplastic oxidative burst peroxidase in Arabidopsis is a major component of
854 pattern-triggered immunity. Plant Cell **24**: 275–287
855 **Deppmann CD, Acharya A, Rishi V, Wobbes B, Smeekens S, Taparowsky EJ, Vinson C**
856 (2004) Dimerization specificity of all 67 B-ZIP motifs in Arabidopsis thaliana: A
857 comparison to Homo sapiens B-ZIP motifs. Nucleic Acids Res **32**: 3435–3445
858 **Ding Y, Sun T, Ao K, Peng Y, Zhang Y, Li X, Zhang Y** (2018) Opposite roles of salicylic
859 acid receptors NPR1 and NPR3/NPR4 in transcriptional regulation of plant immunity.
860 Cell **173**: 1454–1467
861 **Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,**
862 **Gingeras TR** (2013) STAR: Ultrafast universal RNA-seq aligner. Bioinformatics **29**:
863 15–21

Dudley QM, Cai YM, Kallam K, Debreyne H, Carrasco Lopez JA, Patron NJ (2021) Biofoundry-assisted expression and characterization of plant proteins. Synth Biol **6**: ysab029

Ekengren SK, Liu Y, Schiff M, Dinesh-Kumar SP, Martin GB (2003) Two MAPK cascades, NPR1, and TGA transcription factors play a role in Pto-mediated disease resistance in tomato. Plant J **36**: 905–917

Eschenfeldt WH, Stols L, Sanville Millard C, Joachimiak A, Donnelly MI (2009) A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. Methods Mol Biol **498**: 105–115

Fan W, Dong X (2002) In Vivo Interaction between NPR1 and Transcription Factor TGA2 Leads to Salicylic Acid-Mediated Gene Activation in Arabidopsis. Plant Cell **14**: 1377–1389

FAO (2020) World Food and Agriculture - Statistical Pocketbook 2020. Rome. doi: 10.4060/cb1521en

Federici F, Dupuy L, Laplaze L, Heisler M, Haseloff J (2012) Integrated genetic and computation methods for in planta cytometry. Nat Methods **9**: 483–485

Feng J, Cheng Y, Zheng C (2020) Expression patterns of octoploid strawberry TGA genes reveal a potential role in response to Podosphaera aphanis infection. Plant Biotechnol Rep **14**: 55–67

Fernandez-Pozo N, Menda N, Edwards JD, Saha S, Tecle IY, Strickler SR, Bombarely A, Fisher-York T, Pujar A, Foerster H, et al (2015) The Sol Genomics Network (SGN)-from genotype to phenotype to breeding. Nucleic Acids Res **43**: D1036–D1041

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. *In* JM Walker, ed, Proteomics Protoc. Handb., 1st ed. Humana Press, pp 571–607

889 **Gatz C** (2013) From Pioneers to Team Players: TGA Transcription Factors Provide a
890 Molecular Link Between Different Stress Pathways. Mol Plant-Microbe Interact **26**:
891 151–159
892 **Ghareeb H, Laukamm S, Lipka V** (2016) COLORFUL-circuit: A platform for rapid
893 multigene assembly, delivery, and expression in plants. Front Plant Sci **7**: 246
894 **Heo L, Arbour CF, Janson G, Feig M** (2021) Improved Sampling Strategies for Protein
895 Model Refinement Based on Molecular Dynamics Simulation. J Chem Theory Comput
896 **17**: 1931–1943
897 **Herrera-Vásquez A, Fonseca A, Ugalde JM, Lamig L, Seguel A, Moyano TC, Gutiérrez**
898 **RA, Salinas P, Vidal EA, Holuigue L** (2021) TGA class II transcription factors are
899 essential to restrict oxidative stress in response to UV-B stress in Arabidopsis. J Exp Bot
900 **72**: 1891–1905
901 **Hess B, Bekker H, Berendsen HJC, Fraaije JGEM** (1997) LINCS: A Linear Constraint
902 Solver for molecular simulations. J Comput Chem **18**: 1463–1472
903 **Hou J, Sun Q, Li J, Ahammed GJ, Yu J, Fang H, Xia X** (2019) Glutaredoxin S25 and its
904 interacting TGACG motif-binding factor TGA2 mediate brassinosteroid-induced
905 chlorothalonil metabolism in tomato plants. Environ Pollut **255**: 113256
906 **Hussain RMF, Sheikh AH, Haider I, Quareshy M, Linthorst HJM** (2018) Arabidopsis
907 WRKY50 and TGA transcription factors synergistically activate expression of PR1.
908 Front Plant Sci **9**: 930
909 **Ivani I, Dans PD, Noy A, Pérez A, Faustino I, Hospital A, Walther J, Andrio P, Goñi R,**
910 **Balaceanu A, et al** (2015) Parmbsc1: A refined force field for DNA simulations. Nat
911 Methods **13**: 55–58
912 **Jakoby M, Weisshaar B, Dröge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T,**
913 **Parcy F** (2002) bZIP transcription factors in Arabidopsis. Trends Plant Sci **7**: 106–111

914 **Jin H, Choi SM, Kang MJ, Yun SH, Kwon DJ, Noh YS, Noh B** (2018) Salicylic acid-
915 induced transcriptional reprogramming by the HAC-NPR1-TGA histone
916 acetyltransferase complex in Arabidopsis. Nucleic Acids Res **46**: 11712–11725
917 **Johnson C, Mhatre A, Arias J** (2008) NPR1 preferentially binds to the DNA-inactive form
918 of Arabidopsis TGA2. Biochim Biophys Acta - Gene Regul Mech **1779**: 583–589
919 **Jones DT, Taylor WR, Thornton JM** (1992) The rapid generation of mutation data matrices
920 from protein sequences. Bioinformatics **8**: 275–282
921 **Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool**
922 **K, Bates R, Žídek A, Potapenko A, et al** (2021) Highly accurate protein structure
923 prediction with AlphaFold. Nature **596**: 583–589
924 **Karimi M, Inzé D, Depicker A** (2002) GATEWAY^{TM} vectors for Agrobacterium-mediated
925 plant transformation. Trends Plant Sci **7**: 193–195
926 **Katagiri F, Lam E, Chua NH** (1989) Two tobacco DNA-binding proteins with homology to
927 the nuclear factor CREB. Nature **340**: 727–730
928 **Katagiri F, Seipel K, Chua N-H** (1992) Identification of a Novel Dimer Stabilization Region
929 in a Plant bZIP Transcription Activator. Mol Cell Biol **12**: 4809–4816
930 **Katoh K, Standley DM** (2013) MAFFT multiple sequence alignment software version 7:
931 Improvements in performance and usability. Mol Biol Evol **30**: 772–780
932 **Kesarwani M, Yoo J, Dong X** (2007) Genetic interactions of TGA transcription factors in
933 the regulation of Pathogenesis-Related genes and disease resistance in Arabidopsis. Plant
934 Physiol **144**: 336–346
935 **Kim D, Song L, Breitwieser FP, Salzberg SL** (2016) Centrifuge: rapid and accurate
936 classificaton of metagenomic sequences. Genome Res **26**: 1721–1729
937 **Kumar S, Stecher G, Li M, Knyaz C, Tamura K** (2018) MEGA X: Molecular evolutionary
938 genetics analysis across computing platforms. Mol Biol Evol **35**: 1547–1549

939 **Kumar S, Zavaliev R, Wu Q, Zhou Y, Cheng J, Dillard L** (2022) Structural basis of NPR1
940 in activating plant immunity. Nature. doi: 10.1038/s41586-022-04699-w
941 **Lasierra P, Prat S** (2018) Transient transactivation studies in Nicotiana benthamiana leaves.
942 *In* L Oñate-Sánchez, ed, Methods Mol. Biol. Springer Science+Business Media, pp 311–
943 322
944 **Lawaju BR, Lawrence KS, Lawrence GW, Klink VP** (2018) Harpin-inducible defense
945 signaling components impair infection by the ascomycete Macrophomina phaseolina.
946 Plant Physiol Biochem **129**: 331–348
947 **Lazar A, Coll A, Dobnik D, Baebler Š, Bedina-Zavec A, Žel J, Gruden K** (2014)
948 Involvement of potato (Solanum tuberosum L.) MKK6 in response to potato virus Y.
949 PLoS One **9**: e104553
950 **Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E** (1998) Functional analysis of
951 regulatory sequences controlling PR-1 gene expression in Arabidopsis. Plant J **16**: 223–
952 233
953 **Lemaire-Chamley M, Koutouan C, Jorly J, Assali J, Yoshida T, Nogueira M, Tohge T,**
954 **Ferrand C, Peres LEP, Asamizu E, et al** (2022) A chimeric TGA repressor slows
955 down fruit maturation and ripening in tomato. Plant Cell Physiol **63**: 120–134
956 **Li H** (2018) Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics **34**:
957 3094–3100
958 **Li S, Lauri A, Ziemann M, Busch A, Bhave M, Zachgo S** (2009) Nuclear activity of
959 ROXY1, a glutaredoxin interacting with TGA factors, is required for petal development
960 in Arabidopsis thaliana. Plant Cell **21**: 429–441
961 **Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE**
962 (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field.
963 Proteins **78**: 1950–1958

964 **Lukan T, Baebler Š, Pompe-Novak M, Guček K, Zagorščak M, Coll A, Gruden K**
965 (2018a) Cell death is not sufficient for the restriction of potato virus Y spread in
966 hypersensitive response-conferred resistance in potato. Front Plant Sci **9**: 168
967 **Lukan T, Machens F, Coll A, Baebler Š, Messerschmidt K, Gruden K** (2018b) Plant X-
968 tender: An extension of the AssemblX system for the assembly and expression of
969 multigene constructs in plants. PLoS One **13**: e0190526
970 **Lukan T, Pompe-Novak M, Baebler Š, Tušek-Žnidarič M, Kladnik A, Križnik M, Blejec**
971 **A, Zagorščak M, Stare K, Dušak B, et al** (2020) Precision transcriptomics of viral foci
972 reveals the spatial regulation of immune-signaling genes and identifies RBOHD as an
973 important player in the incompatible interaction between potato virus Y and potato. Plant
974 J **104**: 645–661
975 **Lukan T, Županič A, Mahkovec Povalej T, Brunkard JO, Kmetič M, Juteršek M,**
976 **Baebler Š, Gruden K** (2022) Chloroplast redox state changes mark cell-to-cell
977 signalling in the hypersensitive response. New Phytol Accepted Manuscript, doi:
978 10.1111/nph.18425
979 **Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C** (2015)
980 ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from
981 ff99SB. J Chem Theory Comput **11**: 3696–3713
982 **Matys V, Kel-Margoulis O V., Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I,**
983 **Chekmenev D, Krull M, Hornischer K, et al** (2006) TRANSFAC(R) and its module
984 TRANSCompel(R): transcriptional gene regulation in eukaryotes. Nucleic Acids Res **34**:
985 D108–D110
986 **de Medeiros Oliveira M, Bonadio I, Lie de Melo A, Mendes Souza G, Durham AM**
987 (2021) TSSFinder—fast and accurate ab initio prediction of the core promoter in
988 eukaryotic genomes. Brief Bioinform **22**: bbab198

989 **Moll JR, Acharya A, Gal J, Mir AA, Vinson C** (2002) Magnesium is required for specific
990 DNA binding of the CREB B-ZIP domain. Nucleic Acids Res **30**: 1240–1246
991 **Moon SJ, Park HJ, Kim TH, Kang JW, Lee JY, Cho JH, Lee JH, Park DS, Byun MO,**
992 **Kim BG, et al** (2018) OsTGA2 confers disease resistance to rice against leaf blight by
993 regulating expression levels of disease related genes via interaction with NH1. PLoS One
994 **13**: e0206910
995 **Morgunova E, Taipale J** (2017) Structural perspective of cooperative transcription factor
996 binding. Curr Opin Struct Biol **47**: 1–8
997 **Nakagawa T, Suzuki T, Murata S, Nakamura S, Hino T, Maeo K, Tabata R, Kawai T,**
998 **Tanaka K, Niwa Y, et al** (2007) Improved Gateway Binary Vectors: High-Performance
999 Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. Biosci
1000 Biotechnol Biochem **71**: 2095–2100
1001 **Ndamukong I, Abdallat A Al, Thurow C, Fode B, Zander M, Weigel R, Gatz C** (2007)
1002 SA-inducible Arabidopsis glutaredoxin interacts with TGA factors and suppresses JA-
1003 responsive PDF1.2 transcription. Plant J **50**: 128–139
1004 **Niggeweg R, Thurow C, Kegler C, Gatz C** (2000) Tobacco transcription factor TGA2.2 is
1005 the main component of as-1- binding factor ASF-1 and is involved in salicylic acid- and
1006 auxin-inducible expression of as-1-containing target promoters. J Biol Chem **275**:
1007 19897–19905
1008 **Noshi M, Mori D, Tanabe N, Maruta T, Shigeoka S** (2016) Arabidopsis clade IV TGA
1009 transcription factors, TGA10 and TGA9, are involved in ROS-mediated responses to
1010 bacterial PAMP flg22. Plant Sci **252**: 12–21
1011 **O'Malley RC, Huang SSC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M,**
1012 **Gallavotti A, Ecker JR** (2016) Cistrome and epicistrome features shape the regulatory
1013 DNA landscape. Cell **165**: 1280–1292

1014 **Petek M, Zagorščak M, Ramšak Ž, Sanders S, Tomaž Š, Tseng E, Zouine M, Coll A,**
1015 **Gruden K** (2020) Cultivar-specific transcriptome and pan-transcriptome reconstruction
1016 of tetraploid potato. Sci Data **7**: 249
1017 **Pham GM, Hamilton JP, Wood JC, Burke JT, Zhao H, Vaillancourt B, Ou S, Jiang J,**
1018 **Robin Buell C** (2020) Construction of a chromosome-scale long-read reference genome
1019 assembly for potato. Gigascience **9**: giaa100
1020 **Piana S, Donchev AG, Robustelli P, Shaw DE** (2015) Water dispersion interactions
1021 strongly influence simulated structural properties of disordered protein states. J Phys
1022 Chem B **119**: 5113–5123
1023 **Pontier D, Miao ZH, Lam E** (2001) Trans-dominant suppression of plant TGA factors
1024 reveals their negative and positive roles in plant defense responses. Plant J **27**: 529–538
1025 **Popescu SC, Popescu G V., Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-**
1026 **Kumar SP** (2007) Differential binding of calmodulin-related proteins to their targets
1027 revealed through high-density Arabidopsis protein microarrays. Proc Natl Acad Sci U S
1028 A **104**: 4730–4735
1029 **Powers SK, Holehouse AS, Korasick DA, Schreiber KH, Clark N, Jing H, Emenecker R,**
1030 **Han S, Tycksen E, Sozzani R, et al** (2019) Nucleo-cytoplasmic partitioning of ARF
1031 proteins controls auxin responses in Arabidopsis thaliana. Mol Cell **76**: 177–190
1032 **Ramšak Ž, Baebler Š, Rotter A, Korbar M, Mozetič I, Usadel B, Gruden K** (2014)
1033 GoMapMan: integration, consolidation and visualization of plant gene annotations
1034 within the MapMan ontology. Nucleic Acids Res **42**: 1167–1175
1035 **Rieping M** (1994) A Dominant Negative Mutant of PG13 Suppresses Transcription from a
1036 Cauliflower Mosaic Virus 35S Truncated Promoter in Transgenic Tobacco Plants. Plant
1037 Cell Online **6**: 1087–1098
1038 **Rochon A, Boyle P, Wignes T, Fobert PR, Despres C** (2006) The coactivator function of

1039 Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-
1040 terminal cysteines. Plant Cell **18**: 3670–3685
1041 **Rupar M, Faurez F, Tribodet M, Gutiérrez Aguirre I, Delaunay A, Glais L, Kriznik M,**
1042 **Dobnik D, Gruden K, Jacquot E, et al** (2015) Fluorescently tagged Potato virus Y: A
1043 versatile tool for functional analysis of plant-virus interactions. Mol Plant-Microbe
1044 Interact **28**: 739–750
1045 **Savelli B, Li Q, Webber M, Jemmat AM, Robitaille A, Zamocky M, Mathé C, Dunand C**
1046 (2019) RedoxiBase: A database for ROS homeostasis regulated proteins. Redox Biol **26**:
1047 101247
1048 **Schumacher MA, Goodman RH, Brennan RG** (2000) The structure of a CREB
1049 bZIP·somatostatin CRE complex reveals the basis for selective dimerization and divalent
1050 cation-enhanced DNA binding. J Biol Chem **275**: 35242–35247
1051 **Smyth GK, Ritchie ME, Law CW, Alhamdoosh M, Su S, Dong X, Tian L** (2018) RNA-
1052 seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Research **5**: 1408
1053 **Stark JC, Huang A, Nguyen PQ, Dubner RS, Hsu KJ, Ferrante TC, Anderson M,**
1054 **Kanapskyte A, Mucha Q, Packett JS, et al** (2018) BioBits^{TM} Bright: A fluorescent
1055 synthetic biology education kit. Sci Adv **4**: eaat5107
1056 **Studier FW** (2005) Protein production by auto-induction in high density shaking cultures.
1057 Protein Expr Purif **41**: 207–234
1058 **Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA,**
1059 **Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al** (2005) Gene set enrichment
1060 analysis: A knowledge-based approach for interpreting genome-wide expression profiles.
1061 Proc Natl Acad Sci U S A **102**: 15545–15550
1062 **Thibaud-Nissen F, Wu H, Richmond T, Redman JC, Johnson C, Green R, Arias J,**
1063 **Town CD** (2006) Development of Arabidopsis whole-genome microarrays and their

application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. Plant J **47**: 152–162

Thurow C, Schiermeyer A, Krawczyk S, Butterbrodt T, Nickolov K, Gatz C (2005) Tobacco bZIP transcription factor TGA2.2 and related factor TGA2.1 have distinct roles in plant defense responses and plant development. Plant J **44**: 100–113

Tomaž Š, Gruden K, Coll A (2022) TGA transcription factors - Structural characteristics as basis for functional variability. Front Plant Sci **13**: 935819

Vermaas J V., Hardy DJ, Stone JE, Tajkhorshid E, Kohlmeyer A (2016) TopoGromacs: Automated Topology Conversion from CHARMM to GROMACS within VMD. J Chem Inf Model **56**: 1112–1116

Warris S, Schijlen E, Van De Geest H, Vegesna R, Hesselink T, Te Lintel Hekkert B, Sanchez Perez G, Medvedev P, Makova KD, De Ridder D (2018) Correcting palindromes in long reads after whole-genome amplification. BMC Genomics **19**: 798

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, et al (2011) Genome sequence and analysis of the tuber crop potato. Nature **475**: 189–195

Zander M, La Camera S, Lamotte O, Métraux JP, Gatz C (2010) Arabidopsis thaliana class-II TGA transcription factors are essential activators of jasmonic acid/ethylene-induced defense responses. Plant J **61**: 200–210

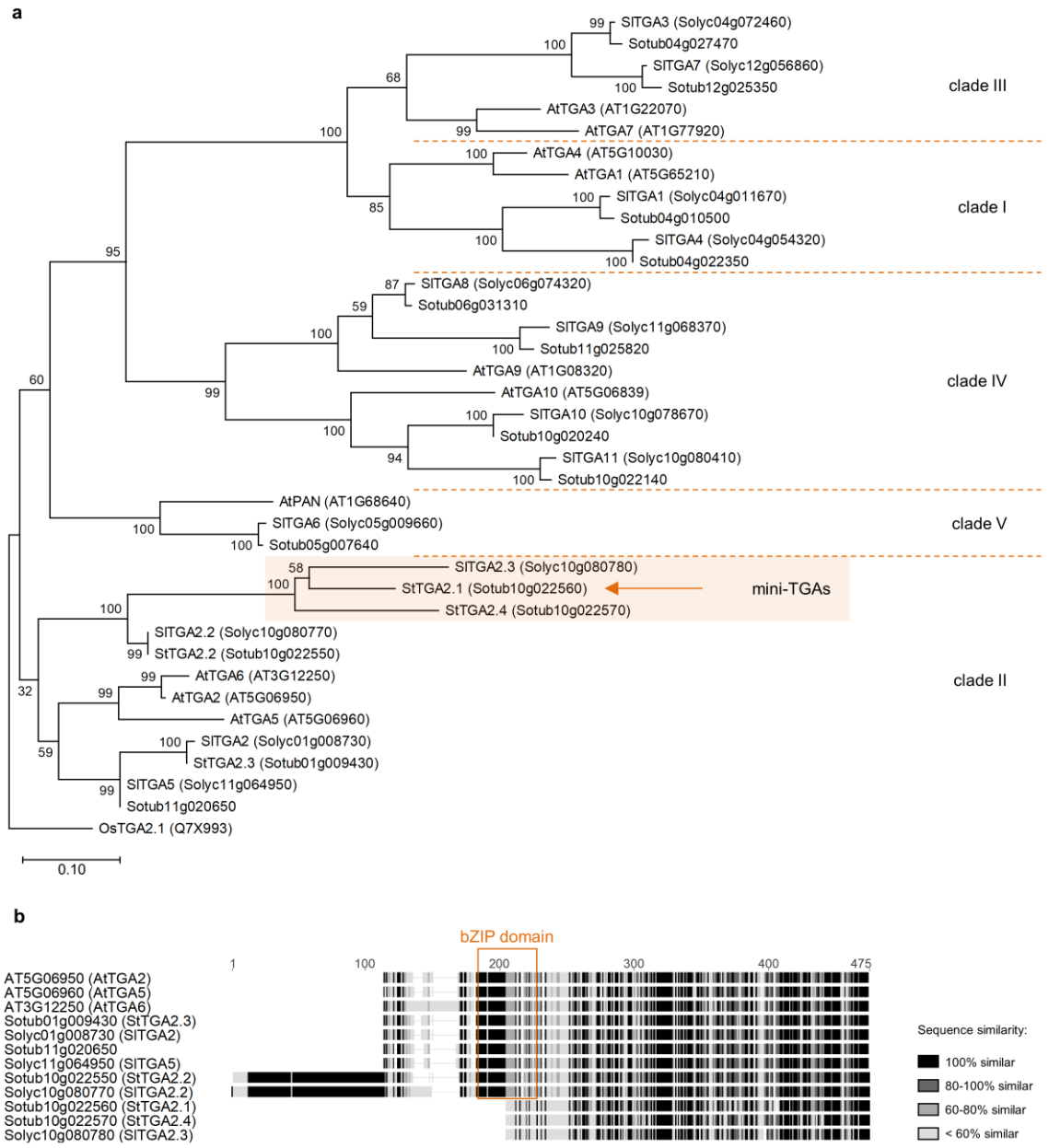
Zhang Y, Fan W, Kinkema M, Li X, Dong X (1999) Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-1 gene. Proc Natl Acad Sci U S A **96**: 6523–6528

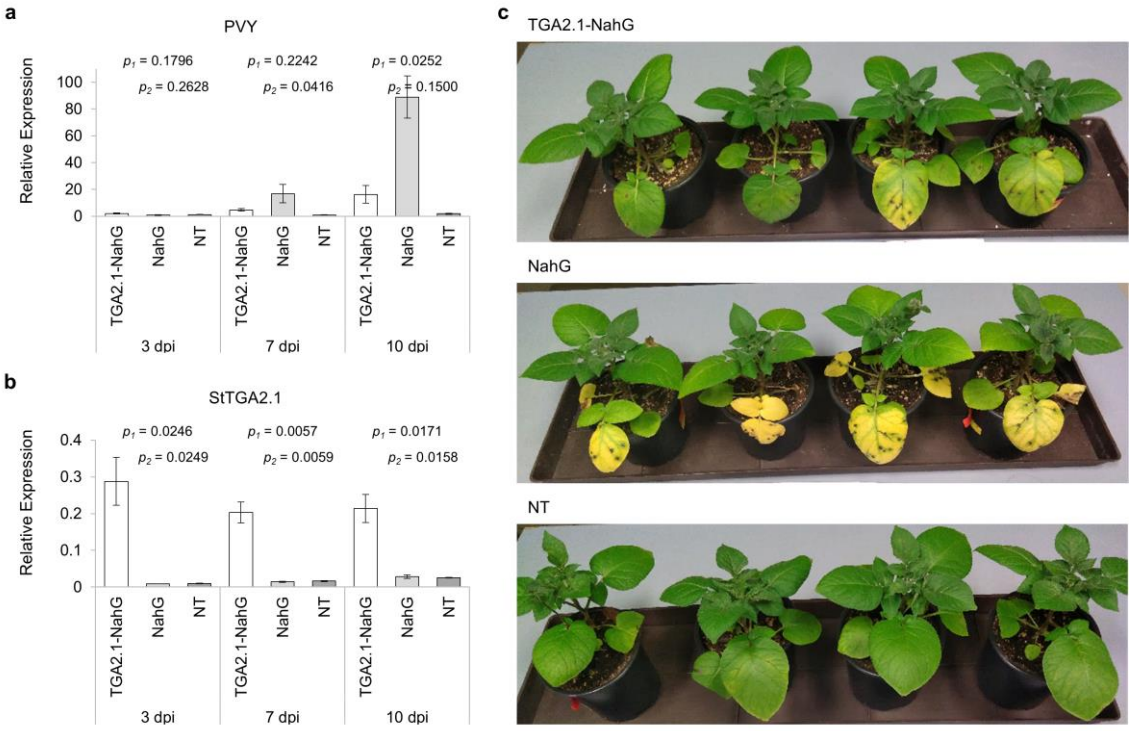
Zhang Y, Tessaro MJ, Lassner M, Li X (2003) Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their eedundant and essential roles in systemic acquired resistance. Plant Cell **15**: 2647–2653

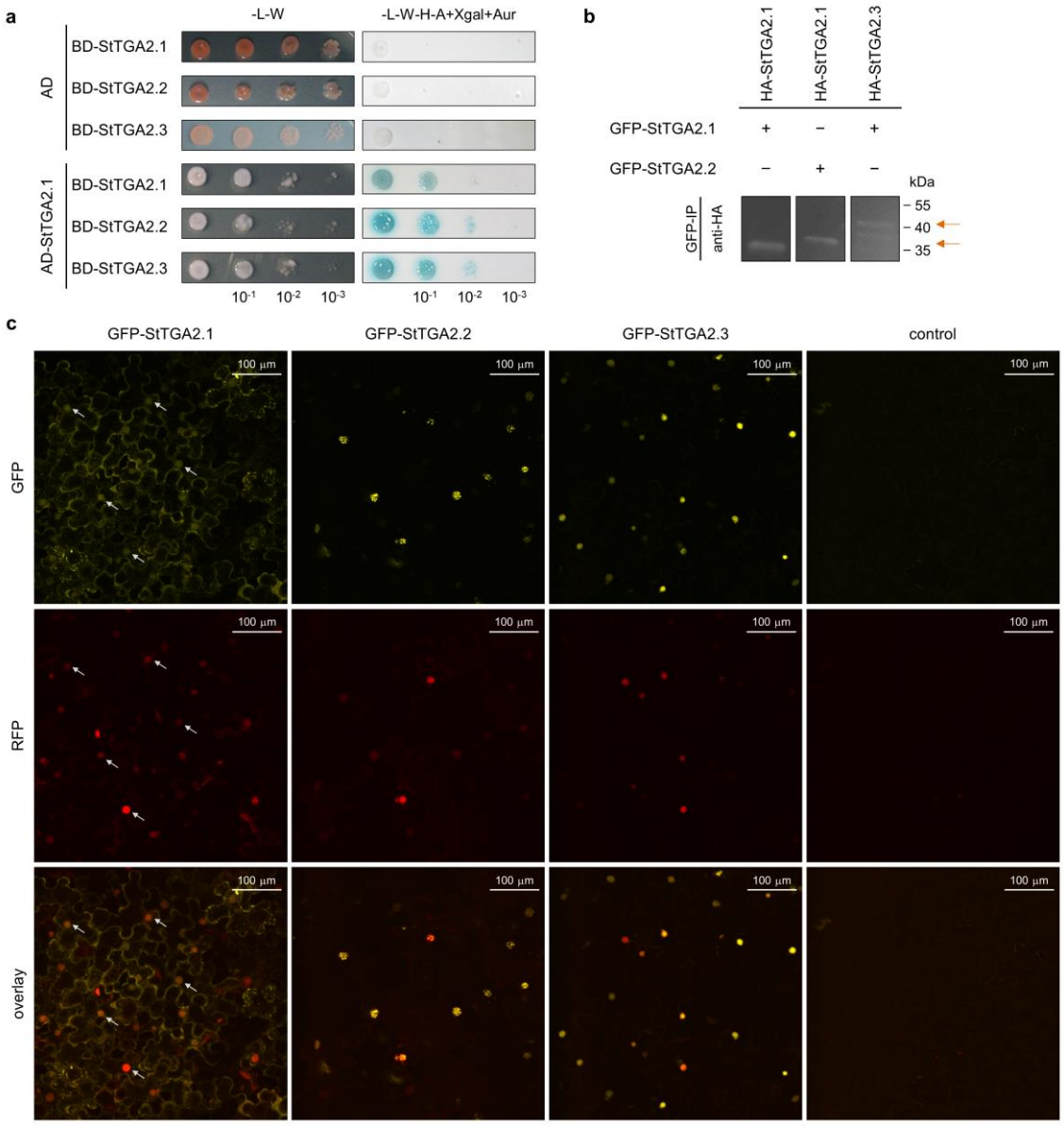
Zhou JM, Zhang Y (2020) Plant Immunity: Danger Perception and Signaling. Cell **181**:

1089 978–989

1090

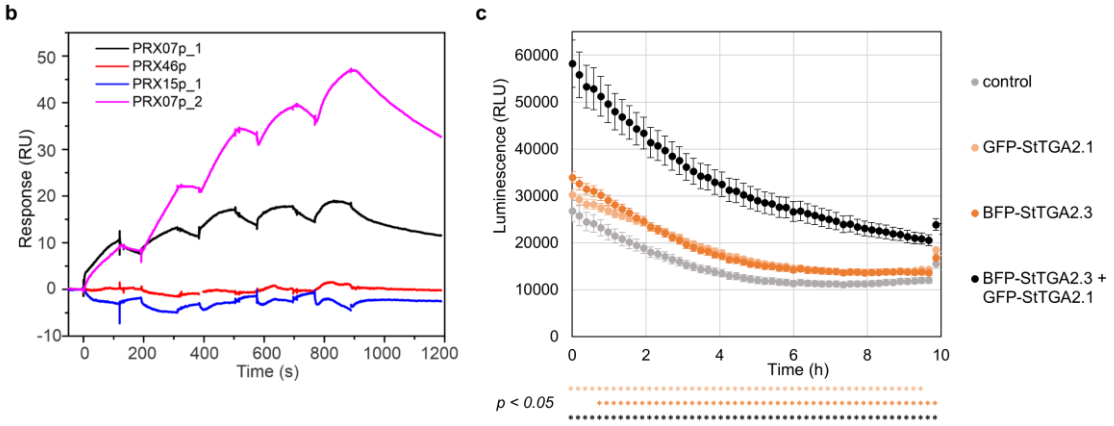

a
SlTGA3 (Solyc04g072460)
Sotub04g027470
SlTGA7 (Solyc12g056860)
Sotub12g025350
AtTGA3 (AT1G22070)
AtTGA7 (AT1G77920)
AtTGA4 (AT5G10030)
AtTGA1 (AT5G65210)
SlTGA1 (Solyc04g011670)
Sotub04g010500
SlTGA4 (Solyc04g054320)
Sotub04g022350
SlTGA8 (Solyc06g074320)
Sotub06g031310
SlTGA9 (Solyc11g068370)
Sotub11g025820
AtTGA9 (AT1G08320)
AtTGA10 (AT5G06839)
SlTGA10 (Solyc10g078670)
Sotub10g020240
SlTGA11 (Solyc10g080410)
Sotub10g022140
AtPAN (AT1G68640)
SlTGA6 (Solyc05g009660)
Sotub05g007640
SlTGA2.3 (Solyc10g080780)
StTGA2.1 (Sotub10g022560)
StTGA2.4 (Sotub10g022570)
SlTGA2.2 (Solyc10g080770)
StTGA2.2 (Sotub10g022550)
AtTGA6 (AT3G12250)
AtTGA2 (AT5G06950)
AtTGA5 (AT5G06960)
SlTGA2 (Solyc01g008730)
StTGA2.3 (Sotub01g009430)
SlTGA5 (Solyc11g064950)
Sotub11g020650
OsTGA2.1 (Q7X993)
0.10
clade III
clade I
clade IV
clade V
clade II
mini-TGAs
b
bZIP domain
1
100
200
300
400
475
AT5G06950 (AtTGA2)
AT5G06960 (AtTGA5)
AT3G12250 (AtTGA6)
Sotub01g009430 (StTGA2.3)
Solyc01g008730 (SlTGA2)
Sotub11g020650
Solyc11g064950 (SlTGA5)
Sotub10g022550 (StTGA2.2)
Solyc10g080770 (SlTGA2.2)
Sotub10g022560 (StTGA2.1)
Sotub10g022570 (StTGA2.4)
Solyc10g080780 (SlTGA2.3)
Sequence similarity:
100% similar
80-100% similar
60-80% similar
< 60% similar

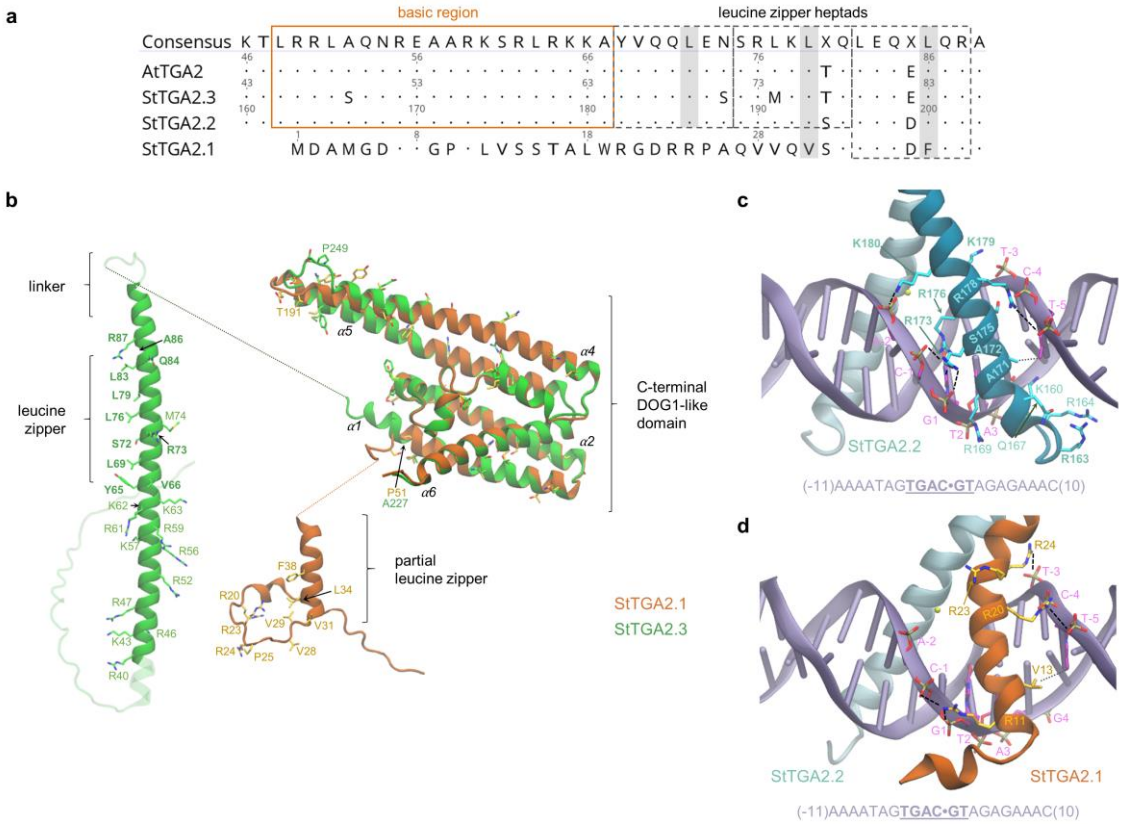

a
PVY
Relative Expression
p_1 = 0.1796
p_2 = 0.2628
p_1 = 0.2242
p_2 = 0.0416
p_1 = 0.0252
p_2 = 0.1500
100
80
60
40
20
0
TGA2.1-NahG
NahG
NT
3 dpi
7 dpi
10 dpi
b
StTGA2.1
Relative Expression
p_1 = 0.0246
p_2 = 0.0249
p_1 = 0.0057
p_2 = 0.0059
p_1 = 0.0171
p_2 = 0.0158
0.4
0.3
0.2
0.1
0
TGA2.1-NahG
NahG
NT
3 dpi
7 dpi
10 dpi
c
TGA2.1-NahG
NahG
NT


a
-L-W
-L-W-H-A+Xgal+Aur
AD
BD-StTGA2.1
BD-StTGA2.2
BD-StTGA2.3
AD-StTGA2.1
BD-StTGA2.1
BD-StTGA2.2
BD-StTGA2.3
10-1 10-2 10-3
10-1 10-2 10-3
b
HA-StTGA2.1
HA-StTGA2.1
HA-StTGA2.3
GFP-StTGA2.1 + − +
GFP-StTGA2.2 − + −
kDa
55
40
35
GFP-IP
anti-HA
c
GFP-StTGA2.1
GFP-StTGA2.2
GFP-StTGA2.3
control
GFP
RFP
overlay
100 µm

a

Gene Name	Gene ID	Promoter fragment	
StPRX07	Sotub09g020950	PRX07p_1	-574 TGTGCCCA**ACGTCA**CAATCC -555
		PRX07p_2	-464 TTAGGGG**TGACGT**TTCCAAT -445
StPRX15	Sotub02g035680	PRX15p_1	-733 TTTAATAA**TGATG**ACATTTG -714
		PRX15p_2	-470 TACGCC**TGATGTCA**AGTCCA -451
StPRX46	Sotub03g007840	PRX46p	-628 TGAACC**TGAGGTCA**ACCGTT -609




a
basic region
leucine zipper heptads
Consensus K T L R R L A Q N R E A A R K S R L R K K A Y V Q Q L E N S R L K L X Q L E Q X L Q R A
AtTGA2
StTGA2.3
StTGA2.2
StTGA2.1 M D A M G D G P L V S S T A L W R G D R R P A Q V V Q V S D F
b
linker
leucine zipper
C-terminal DOG1-like domain
partial leucine zipper
StTGA2.1
StTGA2.3
c
StTGA2.2
(-11)AAAATAGTGAC•GTAGAGAAAC(10)
d
StTGA2.2
StTGA2.1
(-11)AAAATAGTGAC•GTAGAGAAAC(10)

Chapter 3

Discussion

3.1 Approaches for Studying Grapevine-Phytoplasma Interactions

A thorough analysis of disease pathogenesis requires the use of various approaches, and the challenges of adapting the available methodology for studying a particular plant species can pose a considerable drawback. Grapevine is among the more demanding species to study. As a relatively large perennial woody plant, it requires a long juvenile period before reaching the adult phase and is difficult to grow in controlled conditions (Chaïb et al., 2010). Furthermore, its tissues are rich in secondary metabolites, such as polysaccharides and polyphenols, which can interfere with different experimental procedures (Gambino, Perrone, & Gribaudo, 2008).

To investigate the processes underlying '*Ca.* P. solani' infection of grapevine, we focused on implementing a transcriptomics approach in our first publication, which included RNA and sRNA sequencing. Careful sample preparation and isolation of high-quality RNA are essential for downstream transcriptomic analyses. As we wanted to study primarily the phytoplasma-inhabited phloem sieve tissue, our samples consisted of excised leaf veins, presenting an additionally difficult tissue in comparison to whole leaves. We encountered difficulties in isolating total RNA from grapevine leaf veins using established guanidine isothiocyanate-based methods, commercial RNA isolation kits, or their adaptations in combination with additional purification on Zymo Spin columns. Even though similar approaches are used for total RNA isolation from grapevine leaves for RNA sequencing (Guan et al., 2018; Haider et al., 2017; Hily et al., 2018), our test experiments resulted in low or inconsistent RNA concentrations, low RNA integrity and/or purity. The high content of phenolic compounds and polysaccharides, which can bind nucleic acids and hinder purification (Carra, Gambino, & Schubert, 2007; Salzman, Fujita, Zhu-Salzman, Hasegawa, & Bressan, 1999), are likely the causes of these problems. We optimized the isolation procedure by implementing a cetyltrimethylammonium bromide (CTAB)-based protocol, which was initially optimized for RNA isolation from pine trees (Chang, Puryear, & Cairney, 1993). CTAB has proven to be effective in RNA isolation from difficult tissues, including mature canes and wood, and its adaptation yielded better results than other protocols, according to Gambino et al. (2008). We combined a modified version of the beginning steps of the CTAB-based protocol by Carra *et al.* (2007), which has also been adapted for the isolation of sRNAs, with additional RNA purification on Zymo Spin columns, as described in our first publication. Using our optimized approach, we were able to isolate high-quality total RNA, including sRNAs, in high amounts with consistent concentrations and high integrity. In addition, our optimized version could be completed in three hours, instead of six or more, drastically shortening the time required for isolation.

The transcriptomic data obtained revealed important differences in transcriptional regulation and sRNA levels prior to symptom development and in the symptomatic phase of infection. Our approach allowed us to identify novel genes associated with '*Ca.* P. solani' infection, providing the basis for further experiments. We obtained new data on phytoplasma-induced hormonal signaling pathways at different time points of infection. Moreover, our method has become established in our laboratory and is already being used for other analyses.

High-throughput methodologies provide a comprehensive overview of the dynamic changes inside the cell and are also usually the first step in identifying specific targets for further analyses. In our second publication, we identified several putative '*Ca.* P. solani' effector proteins using *in silico* analyses and investigated their function. We used a proteomics approach to screen for protein targets of the candidate '*Ca.* P. solani' effector PoSTOSP28, which is sequentially identical to a previously reported stolbur antigenic membrane protein (StAMP). This 16 kDa protein was previously described as a membrane-bound antigen, presented on the phytoplasma outside membrane surface (Fabre, Danet, & Foissac, 2011); however, some AMP proteins are also cleaved from the membrane surface (Barbara, Morton, Clark, & Davies, 2002). In contrast to many other bacterial pathogens, phytoplasma are cell wall-less and reside inside host cells, so their membrane or secreted proteins can interact with intracellular plant proteins, possibly affecting their function. Using protein pull-down coupled with mass spectrometry (MS), we identified 58 protein groups from tobacco (*Nicotiana benthamiana*) as potential PoSTOSP28 interactors. Many of them encompassed proteins highly abundant in plant tissues, such as ribosomal subunits and components of the photosynthetic machinery, which we likely detected also due to their high concentrations in the sample. It is important to note that many parameters, such as plant growing conditions, leaf maturity, and extraction buffer composition, could affect the selection of proteins available for interaction. For example, the use of reducing agents in the extraction buffer may affect interactions between PoSTOSP28 and redox-sensitive proteins. Moreover, we performed the experiment by transiently expressing PoSTOSP28 in tobacco leaves, which likely have a different protein composition than the grapevine phloem tissue, where the phytoplasma reside.

Interaction confirmation was therefore important not only to rule out false positives and technical errors, but also to account for species-specific differences. We used an *in planta* co-immunoprecipitation assay to confirm the interactions between PoSTOSP28 and grapevine orthologues of selected tobacco proteins identified by MS. Our results showed positive interaction between PoSTOSP28 and two grapevine phosphoglucomutases (VvPGM). Phosphoglucomutases are involved in sugar metabolism; specifically, they catalyze the conversion of glucose 1-phosphate and glucose 6-phosphate, and reducing PGM activity leads to impaired growth and development, as well as metabolic changes in Arabidopsis (Malinova et al., 2014). Our enzyme activity assays revealed that PoSTOSP28, as well as the candidate effector PoSTOSP18, increased phosphoglucomutase activity in tobacco, indicating the mechanism through which '*Ca.* P. solani' may influence host carbohydrate metabolism. The bands for both VvPGMs detected in Western blot traveled substantially slower than expected for their calculated protein sizes of about 70 kDa and corresponded to proteins closer to 100 kDa in size. This could be due to post-translational modifications or their intrinsic protein charge, which can affect the travel speed in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). On the other hand, we could not confirm the interaction of PoSTOSP28 with two of the selected proteins identified by MS, i.e., the grapevine fructose-bisphosphate aldolase and geranylgeranyl diphosphate reductase enzymes (Appendix A, Figure A.1). Since we could detect the fructose-bisphosphate aldolase in the protein extract, we postulate that the interaction was either falsely detected in MS, this protein is not the correct grapevine orthologue of the detected

tobacco fructose-bisphosphate aldolase, or this orthologue contains species-specific structural differences, which prevent interaction. In the case of the geranylgeranyl diphosphate reductase, the absence of interaction could also be due to low protein levels in tobacco leaves, as we could not detect its presence in the extract.

Overall, we were able to optimize and apply several established approaches to study plant processes involved in grapevine-phytoplasma interactions, including RNA isolation, protein pull-down, and co-immunoprecipitation assays, thus confirming our first hypothesis. We were also able to show that the putative effector PoSTOSP28 interacts with two VvPGMs, identifying new interactors of phytoplasma effector proteins and confirming our second hypothesis.

3.2 TGA Transcription Factors in Potato Immune Response

In the second part, we continue our investigation of plant protein function by studying the TGA transcription factors and their role in potato immune response. TGAs are a relatively small group of plant-specific proteins that have been recognized for their role in SA-mediated transcriptional regulation in plant immunity (Gatz, 2013). Most known TGAs consist of three complete protein parts, the N-terminus, the bZIP DNA-binding domain, and the C-terminus with a putative Delay of Germination 1 (DOG1) domain, each of them performing different tasks. The TGA structural specifics contribute to the functional variability of different TGAs, which we discuss in our third publication, a review on TGA structural-functional features in Arabidopsis and tobacco.

Interestingly, some plant species contain TGA proteins with substantially shorter sequences, as we have shown in the case of potato TGAs, which we study in our fourth publication. Among the fourteen identified StTGAs, two of them, StTGA2.1 and StTGA2.4, had exceptionally short N-termini and lacked most of their bZIP domain. A similar TGA protein, an orthologue of StTGA2.1 and StTGA2.4, was also identified in tomato (Hou et al., 2019; Lemaire-Chamley et al., 2022), and we referred to this particular protein type as mini-TGAs. Both mini-TGA genes reside immediately downstream of the *StTGA2.2* in the potato genome, indicating they are likely a result of tandem gene duplications, events that are highly frequent in plants and common for plant transcription factors (Panchy, Lehti-Shiu, & Shiu, 2016). Mini-TGAs are not found in Arabidopsis but were reported in strawberry (Feng, Cheng, & Zheng, 2020). Yet another type of TGAs, called TGA-related proteins, which retain only the DOG1 domain with an extension to their C-termini, are found in moss *Physcomitrella patens* and lycophyte *Selaginella moellendorffii* (Nishiyama, Nonogaki, Yamazaki, Nonogaki, & Ohshima, 2021). Another TGA-related protein, with the DOG1 domain and an N-terminal extension, was found in the algae *Klebsormidium nitens* (Nishiyama et al., 2021) and might be evolutionary related to mini-TGAs. TGAs are often identified based on their bZIP domain, which could prevent the identification of mini-TGAs. It would be interesting to see how common they are in other plant species.

Unlike Arabidopsis, which has three clade II TGA (AtTGA2, AtTGA5, and AtTGA6) members, potato clade II has five TGAs (StTGA2.1, StTGA2.2, StTGA2.3, StTGA2.4, and *Sotub11g020650*) with substantial differences in sequence and length, especially in the N-terminal region. Only StTGA2.3 and *Sotub11g020650* have N-termini of similar length to clade II AtTGAs. In addition to the extremely short N-termini of mini-TGAs, StTGA2.2 has an exceptionally long N-terminus, which is characteristic for Arabidopsis clade IV and V members (Gatz, 2013). These structural specifics denote distinct functionalities of clade II StTGAs in comparison to AtTGA, which we further demonstrate through the functional characterization of StTGA2.1. Our results reveal the previously unknown role of mini-

TGAs in the regulation of gene expression and show that they are likely multifunctional and perform other tasks outside the nucleus. One of them could be the modulation of TGA-interacting proteins in the cytoplasm. We show that StTGA2.1 interacts with potato StNPR1 and StNPR3/4 proteins. Localization experiments described in Appendix B.1 show StNPR3/4 is localized predominantly in the nucleus, while StNPR1 forms irregular formations of different sizes in the cytoplasm and near cell nuclei (Figure B.1a). Our preliminary data also show that StTGA2.1 and StNPR1 can co-localize in these formations (Figure B.1b), additionally supporting StTGA2.1 multifunctionality. Altogether, these results confirm our third hypothesis that the mechanisms of potato TGA function differ from those of their orthologues in Arabidopsis.

To investigate the involvement of mini-TGAs in plant immunity, we prepared transgenic SA-deficient potato plants overexpressing *StTGA2.1* and followed whole-genome gene expression after PVY infection. Our results revealed that *StTGA2.1* overexpression attenuated PVY replication and resulted in diverse transcriptional changes following exposure to pathogen infection in comparison to control plants. That a single protein can elicit such extensive changes in the transcriptional landscape is astounding. Its effects are likely augmented by secondary regulation of other transcription factors. Using Gene Set Enrichment Analysis of RNA sequencing data, we show that four groups of transcription factors (WRKY, plant homeodomain finger, photoperiod-responsive 1, and C_2C_2(Zn) DNA-binding One Zinc Finger) are indeed differentially regulated in *StTGA2.1* overexpressing plants compared to controls. We also show that *StTGA2.1* overexpression positively affects the regulation of class III peroxidases (PRX). This group of enzymes is involved in the metabolism of reactive oxygen species, which act in biotic stress-related signaling (Bleau & Spoel, 2021) and could be one of the mechanisms through which StTGA2.1 improves potato immunity against PVY. Altogether, our results demonstrate that StTGA2.1 is involved in transcriptional regulation following PVY infection and that its overexpression greatly reduces PVY replication and symptom development. Although further experiments might be required to confirm the role of StTGA2.1 in the SA-mediated immune response unambiguously, our results strongly indicate that StTGA2.1 functions as part of the SA pathway against PVY, since we confirmed its association with the StNPR1 cofactor, the master regulator of SA-mediated signaling (Backer, Naidoo, & Berg, 2019; W. Wang et al., 2020), and showed the influence of SA on StTGA-StNPR interactions. Therefore, with these results, we for the most part confirm our fourth hypothesis.

Mini-TGAs lack a complete DNA-binding domain and are unable to bind target DNA by themselves, which we also demonstrated by the lack of interactions between StTGA2.1 and several target DNA-binding motifs using surface plasmon resonance. Their domain composition resembles that of microProteins, which are defined as truncated transcription factor-like proteins of small molecular weight that cannot bind DNA but regulate transcription factor activity (Magnani et al., 2014). Our results indeed show StTGA2.1 can form homodimers and heterodimers with StTGA2.2 and StTGA2.3, in yeast and *in planta*, and modulates transcriptional regulation of *StPRX07* through association with StTGA3, confirming our fifth and last hypothesis. Although StTGA2.1 retains a short segment of the leucine zipper, which could contribute to dimerization, contact formation likely occurs through stable interactions involving the C-terminal region, as was shown recently in AtTGA3 (Kumar et al., 2022). The exact mechanism of cooperative promoter activation by the StTGA2.1-StTGA2.3 complex remains to be determined; however, our molecular dynamics simulations suggest that StTGA2.1 dramatically affects the overall conformation of the regulatory heterodimer due to its compact structure. The presence of StTGA2.1 in the activation complex could also affect its affinity for the establishment of interactions with other regulatory proteins, such as StNPR cofactors. The transactivation assays used to determine promoter activity in the presence of StTGAs were performed in *N.*

benthamiana leaves. The regulatory proteins from tobacco could contribute to StTGA-mediated activation.

Chapter 4

Conclusions

Understanding the molecular mechanisms of plant diseases and the different roles of plant proteins is of paramount importance if we are to reduce crop losses due to pathogen infections. Our knowledge in this area continues to grow, but we still know very little about the molecular mechanisms underlying most plant-pathogen interactions in crops. Within this thesis, we demonstrate the application of different functional analysis approaches for studying molecular processes and their components in two plant pathosystems, leading to new discoveries in disease pathogenesis and plant immune response mechanisms in crops.

In studying the grapevine-phytoplasma interactions, we successfully adapted a protocol for efficient isolation of total RNA from grapevine leaf veins, which enabled extensive analyses of grapevine response to infection with '*Ca.* P. solani' at the transcriptomic and small RNA levels. Secondly, using an optimized protein pull-down assay coupled with mass spectrometry, we could identify several potential targets of the putative '*Ca.* P. solani' effector PoSTOSP28 and confirm the interaction between PoSTOSP28 and grapevine phosphoglucomutase enzymes, which are involved in plant sugar metabolism. These results present an important contribution to our understanding of phytoplasma infections and provide the basis for further experiments.

In studying the potato-PVY interactions, we focused on an essential group of transcription factors involved in plant immunity, the TGA transcription factors. We characterized a novel TGA protein type found in potato, the mini-TGAs, through functional analysis of StTGA2.1. We showed how protein truncation can lead to diversification of function and not only loss of function as generally assumed. This can be an important event in gene evolution and should perhaps be investigated in more detail for other gene families. Furthermore, we identified an important component of the potato immune response, revealing new insights into plant immune signaling in crops.

Altogether, the work presented here provides an important contribution to plant disease research at the molecular level. Such knowledge is crucial for the development of more efficient plant breeding programs and attaining new and more resilient crop species.

Appendix A

Effector PoSTOSP28 Interactions with Other Enzymes

A.1 Co-immunoprecipitation Assay

A.1.1 Experimental Procedures

Full length coding sequences of grapevine fructose-bisphosphate aldolase (VvFBA, *Vitvi03g00048*) and geranylgeranyl diphosphate reductase (VvGDR, *Vitvi06g01069*) were amplified from '*Ca.* P. solani'-infected grapevine cultivar Zweigelt complementary DNA and cloned into the pJET1.2/blunt vector using the CloneJET PCR Cloning Kit (Thermo Scientific, USA). The coding sequences were then inserted into the pENTR D-TOPO vector using pENTR™ Directional TOPO® Cloning Kit (Invitrogen, USA), verified with Sanger sequencing (Eurofins Genomics, Germany), and recombined into the pJCV52 Gateway vector (VIB, Belgium) (Karimi, Inzé, & Depicker, 2002) containing the hemagglutinin A (HA) tag, using the Gateway® LR Clonase TM II Enzyme Mix (Invitrogen, USA). Primer pairs used for cloning and sequence confirmation are listed in Table A.1.

Table A.1: Primer pairs used for cloning of grapevine VvFBA and VvGDR.

Target	**Primer Name**	**Primer sequence 5'- 3'**	**Purpose**
VvFBA *Vitvi03g00048*	F_Vitvi03g00048_fba	ATGGCATCAGCAGCTCCATCT	pJET cloning
	R_Vitvi03g00048_fba	TTAGTAGACGTAGCCCTTGACGAAC	
VvGDR *Vitvi06g01069*	F_Vitvi06g01069_ger	ATGGCAGCTGCTACTCCTCTTTA	
	R_Vitvi06g01069_ger	TCACACCCTCAAAGCCTGTACGTC	
VvFBA *Vitvi03g00048*	Vv421_pENTR_Fw	CACCATGGCATCTGCCTCTCTTCTCAA	pENTR cloning
	Vv421_pENTR_Rv_noSTOP	GTAGACATAACCCTTCACGAACATTCCT	
VvGDR *Vitvi06g01069*	Vv608_pENTR_Fw	CACCATGGCCTCCATTTCCCTCAAAACCT	
	Vv608_pENTR_Rv_noSTOP	TACGCTCATCTTATCCATCTCCCTC	
pENTR D-TOPO	M13 F	GTAAAACGACGGCCAGT	sequence confirmation
	M13 R	CAGGAAACAGCTATGACC	
pJET1.2/blunt	Forward Sequencing Primer	CGACTCACTATAGGGAGAGCGGC	
	Reverse Sequencing Primer	AAGAACATCGATTTTCCATGGCAG	

Homemade electrocompetent *Agrobacterium tumefaciens* GV3101 cells were transformed with prepared constructs by electroporation. Transformants were used for the

agroinfiltration of the bottom three fully developed leaves of 3–4 week-old *N. benthamiana* plants, as described previously (Lazar et al., 2014). In cases of co-transformation with agrobacteria carrying different constructs, the 1:1 ratio was applied. An equal volume of agrobacteria carrying the p19 silencing suppressor was added to the mixture. Agrobacteria carrying only p19 were used as controls.

For the co-immunoprecipitation assay, the HA-tagged VvFBA and VvGDR proteins were co-expressed with a yellow fluorescent protein (YFP)-tagged PoSTOSP28 in *N. benthamiana* leaves for four days. Protein fluorescence was confirmed by confocal microscopy. Total proteins were extracted from about 500 mg of leaf material with immunoprecipitation (IP) buffer (25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM DTT, 0.1 mM PMSF, 0.02% Nonidet P-40 (NP-40), 10% glycerol, and cOmplete™ ULTRA Tablets, Mini, EDTA-free Protease Inhibitor Cocktail (Roche Switzerland)). The protein extract was then diluted in IP buffer without NP-40 at a ratio of 1:2.5, followed by incubation with GFP-Trap® Magnetic Agarose beads (ChromoTek, Germany) at 4 °C for 1 h. The beads were washed three times with IP buffer without NP-40, after which the immunoprecipitated proteins were eluted in SDS-PAGE loading buffer (100 mM Tris-HCl, pH 6.8, 4% SDS, 0.2% bromophenol blue, 20% glycerol, and 200 mM DTT). The immunoprecipitated proteins and protein extracts were analyzed by SDS-PAGE and Western blot using antibodies against green fluorescent protein (GFP) (1:3,000 dilution, Invitrogen, USA) and against HA (1:1,000 dilution, ChromoTek, Germany).

A.1.2 Results

Interactions between the candidate effector protein PoSTOSP28 and grapevine VvFBA and VvGDR enzymes were tested using an *in planta* co-immunoprecipitation assay. As shown in Figure A.1, we did not detect any interaction between the tested proteins.

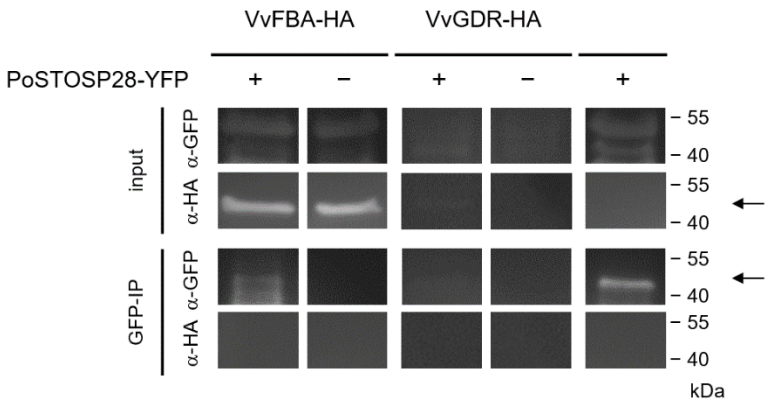


Figure A.1: Co-immunoprecipitation assay results. No interactions were detected between PoSTOSP28 and VvFBA or VvGDR. The combination of YFP- and HA-tagged proteins expressed in *N. benthamiana* is indicated for each sample, whereas each of the interactors was also expressed alone as control (+/−). Interactions were determined by the detection of immunoprecipitated (GFP-IP) complexes with anti-HA antibodies. Detection of YFP-tagged PoSTOSP28 with anti-GFP antibodies in GFP-IP samples and detection of proteins with anti-GFP and anti-HA antibodies in leaf protein extracts (input) are shown as controls. The arrows indicate expected bands.

Appendix B

Experiments with Potato NPR Transcription Cofactors

B.1 Subcellular Localization and Co-localization

B.1.1 Experimental Procedures

Full-length coding sequences of StNPR1 (*Sotub07g011600*), StNPR3/4 (*Sotub02g015550*), and StTGA2.1 (*Sotub10g022560*) were inserted into the pENTR D-TOPO vector using pENTR™ Directional TOPO® Cloning Kit (Invitrogen, USA) and the primer pairs listed in Table B.1. The coding sequences of StNPR1 and StNPR3/4 were then recombined through LR reaction using the Gateway® LR Clonase TM II Enzyme Mix (Invitrogen, USA) into the pK7RWG2.0 expression vector (Karimi et al., 2002), to produce proteins with a C-terminal monomeric red fluorescent protein 1 (RFP) fusion. StTGA2.1 was inserted into the pH7FWG2.0 expression vector (Karimi et al., 2002) to produce a protein with a C-terminal GFP using the same procedure.

Table B.1: Primer pairs used for cloning of potato StTGA and StNPR.

Target	**Primer Name**	**Primer sequence 5'- 3'**	**Purpose**
StNPR1 *Sotub07g011600*	NPR1_pENTR_Fw	CACCATGGATAGTAGGACTGCTTTTTCG	pENTR cloning
	StNPR1_no_stop_Rv	TTTCCTAAAAGGGAGATTATTGGGC	
StNPR3/4 *Sotub02g015550*	NPR3/4A_pENTR_Fw	CACCATGGGAAGTTCTGCTGAACCATC	
	StNPR3/4a_no_stop_Rv	TAGGTTCCTAGCTTTGACACTTGC	
StTGA2.1 *Sotub10g022560*	Sotub10g022560_pENTR_Fw	CACCATGGATGCAATGGGTGATAGGGA	
	Sotub10g022550_560_no_stop _Rv	TTGCTCTCGTGGTCTGGCA	
pENTR D-TOPO	M13 F	GTAAAACGACGGCCAGT	sequence confirmation
	M13 R	CAGGAAACAGCTATGACC	

Agrobacteria-mediated *N. benthamiana* transformation was performed, as described in subchapter A.1.1. Protein localization was visualized three to five days after agroinfiltration, using the Leica TCS SP5 laser scanning confocal microscope mounted on a Leica DMI 6000 CS inverted microscope with an HCX PL APO lambda blue 63.0×1.40 oil-immersion objective (Leica Microsystems, Germany), using previously described settings (Lukan, Machens, et al., 2018). StNPR1 and StTGA2.1 co-localization was visualized with the Leica TCS LSI macroscope with Plan APO 5× objective (Leica Microsystems, Germany), using previously described settings (Lukan, Baebler, et al., 2018).

The red Histone 2B-eYFP (H2B-YFP) nuclear marker (Federici, Dupuy, Laplaze, Heisler, & Haseloff, 2012) was used to visualize cell nuclei. Yellow and green or red fluorescent protein fluorescence was excited with laser lines of 488 nm or 543 nm, respectively. Emission was measured in the window of 505–525 nm for GFP, 525–535 nm for H2B-YFP, 570–630 nm for RFP, and 690–750 nm or 700–800 nm for autofluorescence. Leica LAS AF Lite software (Leica Microsystems, Germany) was used for image processing.

B.1.2 Results

Subcellular localization of RFP-tagged StNPR1 and StNPR3/4 proteins in *N. benthamiana* leaf epidermis and mesophyll was compared with the H2B-YFP nuclear marker (Figure B.1a). StNPR1-RFP formed compact and irregularly shaped formations in the cytoplasm, often near cell nuclei. In contrast, StNPR3/4A was predominantly localized in the nucleus and, to a lesser extent, in the cytoplasm. Co-localization of RFP-tagged StNPR1 and GFP-tagged StTGA2.1 showed that the proteins can co-localize in unevenly sized formations in the cytoplasm (Figure B.1b).

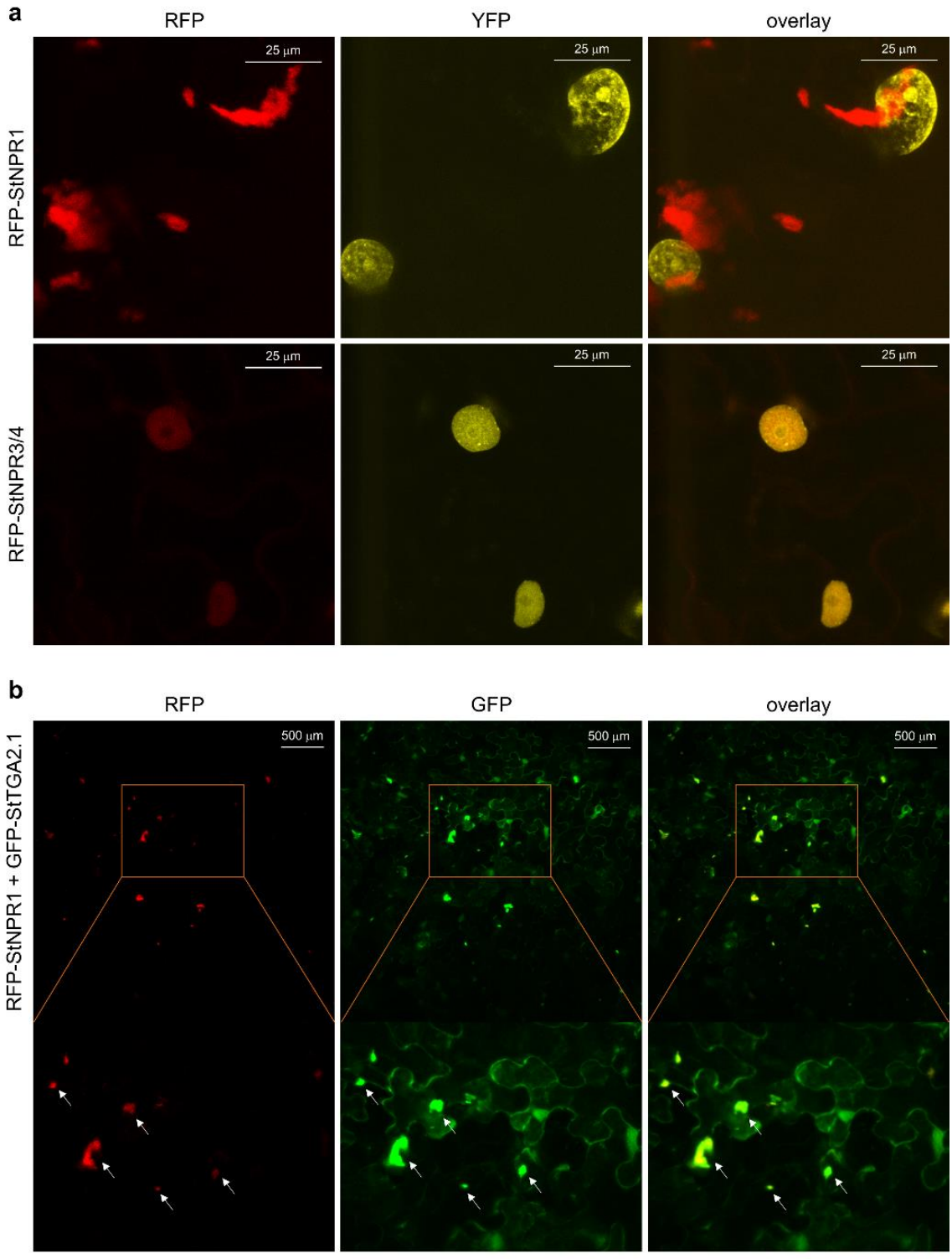


Figure B.1: Subcellular localization and co-localization. a) Localization of RFP-tagged StNPR1 and StNPR3/4 (red) with H2B-YFP nuclear marker (yellow) in *N. benthamiana* leaves. b) Co-localization of RFP-tagged StNPR1 (red) and GFP-tagged StTGA2.1 (green). Orange lines depict the image section close-up. White arrows indicate regions of co-localization, which are colored yellow in channel overlay. a,b) Protein fluorescence is represented as the z-stack maximum projection. Scale bars, a) 25 µm and b) 500 µm.

Appendix C

Supplementary Material of Included Publications

C.1 Supplementary Material for Publication 2.1

Supplementary material is available at https://www.mdpi.com/1422-0067/22/7/3531. Accessed: 31-March-2022.

M. Dermastia, B. Škrlj, R. Strah, B. Anžič, Š. Tomaž, M. Križnik, C. Schönhuber, M. Riedle-Bauer, Ž. Ramšak, M. Petek, A. Kladnik, N. Lavrač, K. Gruden, T. Roitsch, G. Brader and M. Pompe-Novak, "Differential Response of Grapevine to Infection with '*Candidatus* Phytoplasma solani' in Early and Late Growing Season Through Complex Regulation of mRNA and Small RNA Transcriptomes," *International Journal of Molecular Sciences,* vol. 22, no. 7:3531, 2021.

C.2 Supplementary Material for Publication 2.2

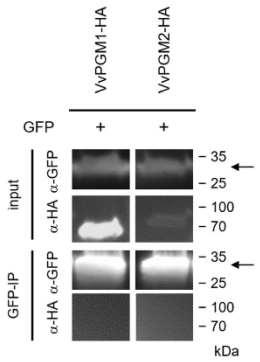
Supplemental Information

Title: Candidate effector proteins of '*Candidatus* Phytoplasma solani' are associated with modulation of plant carbohydrate metabolism toward effective glycolysis, with accelerated ascorbate-glutathione cycle, and with induction of autophagosomes

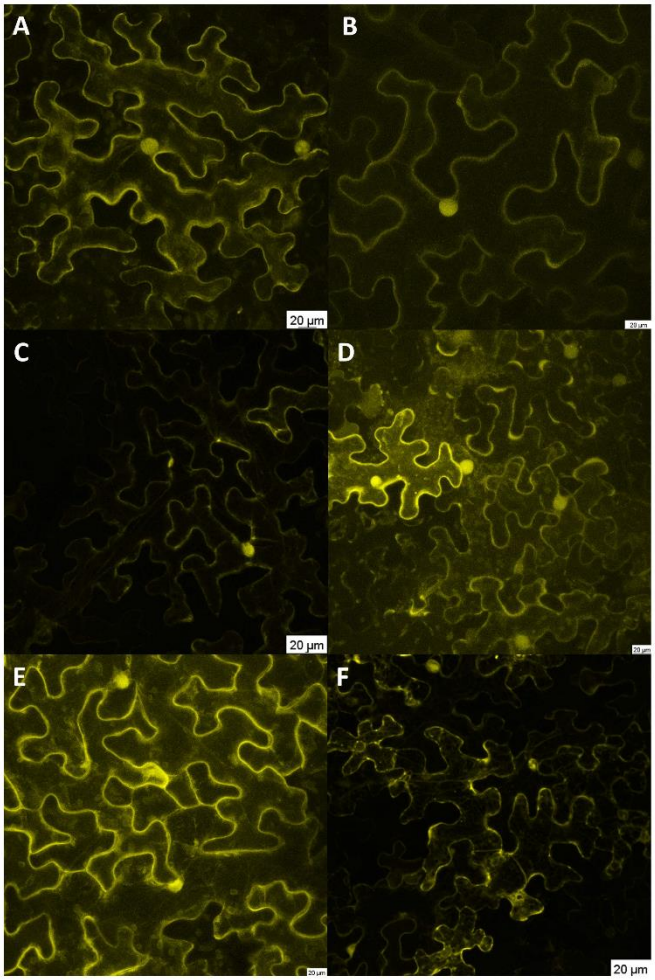
Authors: Marina Dermastia*, Špela Tomaž, Rebeka Strah, Timotej Čepin, Tjaša Lukan, Anna Coll, Barbara Dušak, Barbara Anžič, Thomas Roitsch, Stefanie Wienkoop, Wolfram Weckwert, Kristina Gruden, Maruša Pompe Novak, Günter Brader

* Correspondence: marina.dermastia@nib.si

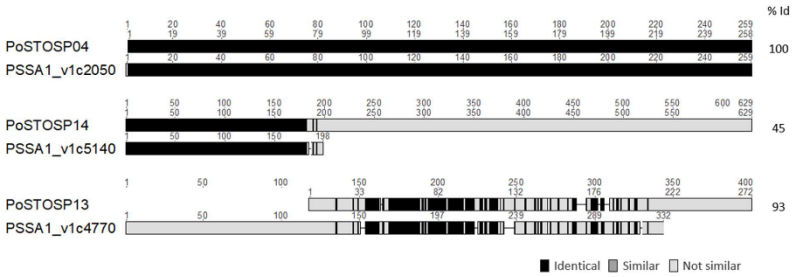
Supplemental Figures



Supplemental Fig. 1. Co-immunoprecipitation assay negative controls. The HA-labeled grapevine phosphoglucomutases (VvPGMs) were co-expressed with GFP (+), to show they do not interact with fluorescent proteins after immunoprecipitation (GFP-IP). Detection of proteins in leaf protein extracts is shown as control (input). Arrows indicate expected bands.



Supplemental Figure 2. Expression of effector proteins tagged with YFP. (A) PoSTOSP04, (B) PoSTOSP06, (C) PoSTOSP13, (D) PoSTOSP14, (E) PoSTOSP18, and (F) PoSTOSP28. Scale bars 20 µm.



Supplemental Figure 3. Pairwise sequence alignments of candidate effector proteins. PoSTOSP04, PoSTOSP13 and PoSTOSP14 align or partially align with previously identified '*Ca*. P. solani' strain SA-1 effector sequences (Music Seruga et al. 2019). The alignments are colored with the Geneious Prime (https://www.geneious.com) sequence similarity color scheme, based on the identity score matrix. Pairwise sequence identity (% Id) of the aligned regions is specified for each alignment. Sequence numbering (amino acid) is shown above each alignment.

Supplemental Tables

Supplemental Table 1. List of primers.

Target	Primer name	Primer sequence 5'-3'	Purpose
PoStoSP04	PoStoSP04_ FP	CACCATGAGAAATAACAAAATAAAAATAAAATTTAAATTG	cloning into pENTR and colony PCR
	PoStoSP04_ RP	TTATTCTATTGTGTTTACAAAAGTACAAAGTTTTAC	
PoStoSP06	PoStoSP06_ FP	CACCATGATCTTGGATCTTAACTTTTCTAAAAGAAAAT	
	PoStoSP06_ RP	TTAGTCTTCATTTTCTTTAGGTTTATCTTGAACAT	
PoStoSP13	PoStoSP13_FP	CACCATGCATTTAAGAAAAAAACGCTTTTTT	
	PoStoSP13_ RP	TTAAAATAAAAAATTGTCTAAAAAAACGAAACAG	
PoStoSP14	PoStoSP14 _FP	CACCATGCATTTAAGAAAAAAACGCTTTTTT	
	PoStoSP14_ RP	TTATTTTTTAACTTTTTGAAAACAATAATATCCAG	
PoStoSP18	PoStoSP18_ FP	CACCATGGTATTAAAGGGTAAATTAAAAATAATATTTTTTTTTG	
	PoStoSP18_ RP	TTAAGAAAGAGCTTTTTTTTGAGCAGC	
PoStoSP28	PoStoSP28_ FP	CACCATGCAAAACACAAAAAAATCTTTAGTTATTAAATT	
	PoStoSP28_ RP	TTATTGTCTTCTTTTTTTAATAACAAAGAAAGCAAC	
Vitvi01g00455	Vv455_pENTR_Fw	CACCATGGTGATGTTCAAGGTTTCTCGAGTC	cloning into pENTR
	Vv455_pENTR_Rv_noSTOP	TGTGATAACAGTGGGAGCAGAACG	
Vitvi16g00891	Vv891_pENTR_Fw	CACCATGGATAGTGTCTTCTTTTCAACCCTCA	
	Vv891_pENTR_Rv_noSTOP	TGTGATCACTGTAGGTTTCTCCCTTC	
Vitvi01g00455	F_Vitvi01g00455_pho	ATGGTGATGTTCAAGGTTTCTCGAGTC	cloning into pJET
	R_Vitvi01g00455_pho	CTATGTGATAACAGTGGGAGCAGAAC	
Vitvi16g00891	F_Vitvi16g00891_pho	ATGGATAGTGTCTTCTTTTCAACCCTCA	
	R_Vitvi16g00891_pho	TTATGTGATCACTGTAGGTTTCTCCCTTC	
Vitvi01g00455	Vv455_middle_Fw	CTTTGAGGTCCCAACTGGCTGG	sequencing
	Vv455_middle_Rv	CGCTTTGCATAAGCTCCACCAA	
Vitvi16g00891	Vv891_middle_Fw	TTGAGGTCCCCACTGGATGGAAATTT	
	Vv891_middle_Rv	AGAAAGATCAATGTCTGGGATGTCTACC	
pENTR	M13 F	GTAAAACGACGGCCAGT	
	M13 R	CAGGAAACAGCTATGACC	
pJET1.2/blunt	Forward Sequencing Primer	CGACTCACTATAGGGAGAGCGGC	
	Reverse Sequencing Primer	AAGAACATCGATTTTCCATGGCAG	

Supplemental Table 2. High-throughput RNA-Seq of uninfected and with '*Candidatus* Phytoplasma solani' infected grapevine cv. Zweigelt vein enriched leaf samples (available as an Excel file). Data are extracted and reused from Supplemental Table 1, Dermastia et al. 2021. For each mRNA sequence and sample, the differences in expression between phytoplasma-infected and uninfected plants were calculated as $\log_2$FC. Only mRNAs with FDR adjusted p-value < 0.05 were considered as differentially expressed (red – up-regulated; green – down-regulated). U – uninfected samples; I – samples infected with '*Ca*. P. solani'; E – early growing season; L – late growing season

Supplemental Table 3. Protein groups identified through protein pull-down and mass spectrometry analysis (available as an Excel file). A list of proteins with their UniProtKB identifiers and fasta headers, identified in the YFP-tagged candidate effector PoSTOSP28 sample, but not present in control sample after pull-down.

Supplemental Table 2. High-throughput RNA-Seq of uninfected and with '*Candidatus* Phytoplasma solani' infected grapevine cv. Zweigelt vein enriched leaf samples. Data are extracted and reused from Supplemental Table 1, Dermastia et al. 2021. For each mRNA sequence and sample, the differences in expression between phytoplasma-infected and uninfected plants were calculated as log_2FC. Only mRNAs with FDR adjusted p-value < 0.05 were considered as differentially expressed (red – up-regulated; green – down-regulated). U – uninfected samples; I – samples infected with '*Ca*. P. solani'; E – early growing season; L – late growing season

Gene ID	DESCRIPTION	BINCODE	NAME	log_2 FC E-I : E-U	adjusted p-value (FDR)	log_2 FC L-I : L-U	adjusted p-value (FDR)
CARBOHYDRATE METABOLISM							
Vitvi18g02758	ADPGLC-PPase large subunit (AT1G27680 similarity)	2.1.2.1	major CHO metabolism.synthesis.starch.AGPase	2.99	0.000	1.25	0.002
Vitvi14g01873	ADP-glucose pyrophosphorylase family protein (AT1G74910 similarity)	2.1.2.1	major CHO metabolism.synthesis.starch.AGPase	0.48	0.003	0.20	0.191
Vitvi01g01824	classical AGP protein (AT5G11680 similarity)	35.2	not assigned.unknown	0.11	0.498	0.39	0.016
Vitvi17g00178	ketose-bisphosphate aldolase class-II family protein (AT1G18270 similarity)	35.1	not assigned.no ontology	0.86	0.039	0.09	0.828
Vitvi09g01500	Aldolase-type TIM barrel family protein (AT3G14130 similarity)	1.2.2	PS.photorespiration.glycolate oxydase	0.78	0.001	1.08	0.000
Vitvi19g01733	Aldolase-type TIM barrel family protein (AT1G16350 similarity)	23.1.2.30	nucleotide metabolism.synthesis.purine.IMP dehydrogenase	0.47	0.013	0.01	0.978
Vitvi08g01506	Aldolase superfamily protein (AT2G36460 similarity)	1.3.6	PS.calvin cycle.aldolase	0.68	0.001	1.04	0.000
Vitvi01g00360	Aldolase superfamily protein (AT2G01140 similarity)	1.3.6	PS.calvin cycle.aldolase	0.67	0.001	1.10	0.000
Vitvi04g02304	Aldolase-type TIM barrel family protein (AT4G21320 similarity)	35.2	not assigned.unknown	0.65	0.064	0.72	0.081
Vitvi09g01499	Aldolase-type TIM barrel family protein (AT3G14130 similarity)	1.2.2	PS.photorespiration.glycolate oxydase	0.62	0.039	0.39	0.187
Vitvi14g00619	threonine aldolase 1 (AT1G08630 similarity)	13.2.5.2	amino acid metabolism.degradation.serine-glycine-cysteine group.glycine	0.41	0.206	-1.01	0.014
Vitvi19g01724	Aldolase-type TIM barrel family protein (AT1G79500 similarity)	10.1.12.2	cell wall.precursor synthesis.KDO pathway.KDO-8-phosphate synthase	0.37	0.085	0.23	0.344
Vitvi03g01173	Aldolase-type TIM barrel family protein (AT5G64250 similarity)	26.1	misc.misc2	0.57	0.238	-0.42	0.490
Vitvi14g01571	Aldolase-type TIM barrel family protein (AT3G01850 similarity)	7.2.3	OPP.non-reductive PP.ribulose-phosphate 3-epimerase	0.11	0.592	0.18	0.403
Vitvi01g01779	Aldolase-type TIM barrel family protein (AT5G13420 similarity)	7.2.2	OPP.non-reductive PP.transaldolase	0.09	0.511	0.78	0.000
Vitvi08g00972	Aldolase superfamily protein (AT2G26800 similarity)	13.2.4.4	amino acid metabolism.degradation.branched chain group.leucine	0.02	0.935	0.21	0.190
Vitvi08g01624	D-tagatose-1%2C6-bisphosphate aldolase subunit (AT2G36895 similarity)	35.2	not assigned.unknown	0.00	1.000	-0.94	0.008
Vitvi07g00047	Aldolase-type TIM barrel family protein (AT4G02610 similarity)	13.1.6.5.5	amino acid metabolism.synthesis.aromatic aa.tryptophan.tryptophan synthase	-0.02	0.870	-0.13	0.309
Vitvi07g00718	Dihydroneopterin aldolase (AT3G11750 similarity)	25.9	C1-metabolism.dihydroneopterin aldolase	-0.06	0.759	0.00	0.992
Vitvi08g01457	Aldolase-type TIM barrel family protein (AT2G36230 similarity)	13.1.7.3	amino acid metabolism.synthesis.histidine.phosphoribosyl-AMP cyclohydrolase	-0.07	0.685	-0.03	0.874
Vitvi02g01702	Aldolase superfamily protein (AT1G12230 similarity)	7.2.2	OPP.non-reductive PP.transaldolase	-0.09	0.681	0.11	0.633
Vitvi12g00745	Aldolase-type TIM barrel family protein (AT5G48220 similarity)	13.1.6.5.4	amino acid metabolism.synthesis.aromatic aa.tryptophan.indole-3-glycerol phosphate synthase	-0.16	0.258	0.20	0.148
Vitvi06g01161	Aldolase superfamily protein (AT2G26800 similarity)	13.2.4.4	amino acid metabolism.degradation.branched chain group.leucine	-0.24	0.511	-0.21	0.580
Vitvi02g00513	Aldolase superfamily protein (AT1G12230 similarity)	7.2.2	OPP.non-reductive PP.transaldolase	-0.36	0.121	-0.30	0.204
Vitvi10g02382	Aldolase-type TIM barrel family protein (AT4G21320 similarity)	20.2.1	stress.abiotic.heat	-0.47	0.105	0.05	0.921
Vitvi18g00261	Aldolase superfamily protein (AT1G69740 similarity)	19.4	tetrapyrrole synthesis.ALA dehydratase	-0.47	0.019	-0.14	0.510
Vitvi03g00048	fructose-bisphosphate aldolase 2 (AT4G38970 similarity)	1.3.6	PS.calvin cycle.aldolase	-0.72	0.063	-1.35	0.001
Vitvi19g01352	Aldolase-type TIM barrel family protein (AT3G14420 similarity)	1.2.2	PS.photorespiration.glycolate oxydase	-1.04	0.013	-1.40	0.002
Vitvi04g01421	fructose-bisphosphate aldolase 2 (AT4G38970 similarity)	1.3.6	PS.calvin cycle.aldolase	-1.10	0.009	-1.66	0.001
Vitvi19g00904	Aldolase superfamily protein (AT4G26530 similarity)	1.3.6	PS.calvin cycle.aldolase	-2.58	0.000	-2.99	0.000
Vitvi14g01938	phosphofructokinase 5 (AT2G22480 similarity)	4.2.4	glycolysis.plastid branch.phosphofructokinase (PFK)	2.21	0.000	0.61	0.014
Vitvi08g01307	fructokinase-like 1 (AT3G54090 similarity)	2.2.1.1	major CHO metabolism.degradation.sucrose.fructokinase	0.82	0.000	-1.77	0.000
Vitvi07g01463	PTHR13697//PTHR13697:SF20 - PHOSPHOFRUCTOKINASE // SUBFAMILY NOT NAMED (1 of 3)	35.2	not assigned.unknown	0.62	0.015	0.53	0.051
Vitvi11g00237	phosphofructokinase 3 (AT4G26270 similarity)	4.3.4	glycolysis.unclear/dually targeted.phosphofructokinase (PFK)	0.46	0.066	0.39	0.129
Vitvi04g00040	phosphofructokinase 3 (AT4G26270 similarity)	4.3.4	glycolysis.unclear/dually targeted.phosphofructokinase (PFK)	0.41	0.091	-0.06	0.825
Vitvi10g00129	Phosphofructokinase family protein (AT1G12000 similarity)	4.1.5	glycolysis.cytosolic branch.pyrophosphate-fructose-6-P phosphotransferase	0.41	0.016	0.45	0.009
Vitvi07g01462	phosphofructokinase 5 (AT2G22480 similarity)	4.2.4	glycolysis.plastid branch.phosphofructokinase (PFK)	0.34	0.097	0.20	0.348
Vitvi16g00381	phosphofructokinase 4 (AT5G61580 similarity)	4.3.4	glycolysis.unclear/dually targeted.phosphofructokinase (PFK)	0.26	0.203	0.65	0.004
Vitvi18g00037	Phosphofructokinase family protein (AT1G76550 similarity)	4.1.5	glycolysis.cytosolic branch.pyrophosphate-fructose-6-P phosphotransferase	0.25	0.227	0.12	0.585
Vitvi10g00212	phosphofructokinase 2 (AT5G47810 similarity)	4.3.4	glycolysis.unclear/dually targeted.phosphofructokinase (PFK)	-0.03	0.866	0.28	0.107
Vitvi15g01370	Phosphofructokinase family protein (AT1G20950 similarity)	35.2	not assigned.unknown	-0.22	0.405	-0.39	0.241
Vitvi01g00025	PTHR10584:SF155 - FRUCTOKINASE-LIKE 2 (1 of 1)	2.2.1.1	major CHO metabolism.degradation.sucrose.fructokinase	-0.25	0.231	-0.38	0.132
Vitvi16g01672	phosphofructokinase 4 (AT5G61580 similarity)	35.2	not assigned.unknown	-0.88	0.099	1.85	0.112
Vitvi15g00775	pfkB-like carbohydrate kinase family protein (AT5G51830 similarity)	2.2.1.1	major CHO metabolism.degradation.sucrose.fructokinase	1.03	0.011	0.99	0.037
Vitvi14g00980	pfkB-like carbohydrate kinase family protein (AT5G37850 similarity)	35.2	not assigned.unknown	0.90	0.000	0.30	0.052
Vitvi05g01516	pfkB-like carbohydrate kinase family protein (AT3G59480 similarity)	2.2.1.1	major CHO metabolism.degradation.sucrose.fructokinase	0.64	0.018	0.64	0.019

Vitvi16g00895	pfkB-like carbohydrate kinase family protein (AT5G51830 similarity)	2.2.1.1	major CHO metabolism.degradation.sucrose.fructokinase	0.55	0.001	0.19	0.178
Vitvi18g02757	pfkB-like carbohydrate kinase family protein (AT1G19600 similarity)	3.5	minor CHO metabolism.others	0.48	0.016	-0.05	0.828
Vitvi07g00016	pfkB-like carbohydrate kinase family protein (AT1G06730 similarity)	3.5	minor CHO metabolism.others	0.33	0.150	0.15	0.512
Vitvi05g00436	pfkB-like carbohydrate kinase family protein (AT5G19150 similarity)	3.5	minor CHO metabolism.others	0.22	0.176	0.14	0.477
Vitvi18g02417	pfkB-like carbohydrate kinase family protein (AT4G10260 similarity)	2.2.1.1	major CHO metabolism.degradation.sucrose.fructokinase	0.19	0.632	-0.06	0.958
Vitvi03g00063	pfkB-like carbohydrate kinase family protein (AT1G49350 similarity)	3.5	minor CHO metabolism.others	0.01	0.984	-0.23	0.418
Vitvi14g00950	pfkB-like carbohydrate kinase family protein (AT1G66430 similarity)	2.2.1.1	major CHO metabolism.degradation.sucrose.fructokinase	-0.16	0.489	-0.25	0.277
Vitvi19g00016	pfkB-like carbohydrate kinase family protein (AT1G17160 similarity)	3.5	minor CHO metabolism.others	-0.17	0.336	-0.24	0.210
Vitvi06g00640	pfkB-like carbohydrate kinase family protein (AT5G58730 similarity)	3.5	minor CHO metabolism.others	-0.54	0.071	1.13	0.004
Vitvi16g00086	pfkB-like carbohydrate kinase family protein (AT4G27600 similarity)	3.5	minor CHO metabolism.others	-0.59	0.038	-1.38	0.000
Vitvi07g00721	pfkB-like carbohydrate kinase family protein (AT4G28706 similarity)	3.5	minor CHO metabolism.others	-0.82	0.001	-0.62	0.017
Vitvi10g00038	hexokinase-1 protein (AT1G05205 similarity)	35.2	not assigned.unknown	0.28	0.140	0.24	0.302
Vitvi06g01272	hexokinase-like 1 (AT1G50460 similarity)	2.2.1.4	major CHO metabolism.degradation.sucrose.hexokinase	0.02	0.977	-1.21	0.038
Vitvi11g00260	hexokinase 1 (AT4G29130 similarity)	2.2.1.4	major CHO metabolism.degradation.sucrose.hexokinase	-0.33	0.048	-0.24	0.148
Vitvi18g01114	hexokinase 3 (AT1G47840 similarity)	2.2.1.4	major CHO metabolism.degradation.sucrose.hexokinase	-0.73	0.006	-0.81	0.020
Vitvi18g00504	phosphoglucose isomerase 1 (AT4G24620 similarity)	4.2.3	glycolysis.plastid branch.glucose-6-phosphate isomerase	-0.23	0.108	0.26	0.074
Vitvi01g00455	Phosphoglucomutase/phosphomannomutase family protein (AT1G70730 similarity)	4.3.2	glycolysis.unclear/dually targeted.phosphoglucomutase (PGM)	0.39	0.125	0.80	0.006
Vitvi16g00891	phosphoglucomutase (AT5G51820 similarity)	4.3.2	glycolysis.unclear/dually targeted.phosphoglucomutase (PGM)	-0.57	0.014	-0.45	0.054
Vitvi01g00442	phosphoglucomutase%2C putative / glucose phosphomutase (AT1G70820 similarity)	4.2.2	glycolysis.plastid branch.phosphoglucomutase (PGM)	-1.30	0.045	-1.39	0.039
Vitvi04g01633	UDP-glucose pyrophosphorylase 2 (AT5G17310 similarity)	4.1.1	glycolysis.cytosolic branch.UGPase	-0.01	0.980	0.91	0.001
Vitvi13g00087	UDP-glucose pyrophosphorylase 3 (AT3G56040 similarity)	35.2	not assigned.unknown	-0.68	0.174	-0.91	0.041
Vitvi04g02166	ascorbate peroxidase 3 (AT4G35000 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	2.00	0.000	2.51	0.004
Vitvi06g00358	ascorbate peroxidase 1 (AT1G07890 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	0.76	0.005	0.82	0.002
Vitvi18g00256	thylakoidal ascorbate peroxidase (AT1G77490 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	0.08	0.718	-0.36	0.073
Vitvi08g01143	ascorbate peroxidase 2 (AT3G09640 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	-0.02	0.970	-0.26	0.518
Vitvi03g00137	ascorbate peroxidase 3 (AT4G35000 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	-0.31	0.148	-0.65	0.006
Vitvi04g00484	ascorbate peroxidase 6 (AT4G32320 similarity)	26.12	misc.peroxidases	-1.00	0.003	-1.31	0.002
Vitvi18g00445	ascorbate peroxidase 4 (AT4G09010 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	-1.28	0.000	-1.64	0.000
OXIDATIVE STRESS							
Vitvi04g02274	catalase 2 (AT4G35090 similarity)	21.6	redox.dismutases and catalases	-0.06	0.915	0.50	0.374
Vitvi18g00095	catalase 2 (AT4G35090 similarity)	21.6	redox.dismutases and catalases	-0.40	0.503	-1.14	0.050
Vitvi04g01564	catalase 2 (AT4G35090 similarity)	21.6	redox.dismutases and catalases	-0.75	0.050	-0.20	0.627
Vitvi13g00241	dehydroascorbate reductase 2 (AT1G75270 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	0.05	0.884	0.59	0.048
Vitvi14g00840	dehydroascorbate reductase 1 (AT5G16710 similarity)	35.2	not assigned.unknown	-0.79	0.011	-0.88	0.007
Vitvi08g01483	monodehydroascorbate reductase 1 (AT3G52880 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	0.28	0.041	0.06	0.691
Vitvi02g00030	monodehydroascorbate reductase 6 (AT1G63940 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	-0.05	0.796	0.28	0.164
Vitvi14g01751	monodehydroascorbate reductase 4 (AT3G27820 similarity)	21.2.1	redox.ascorbate and glutathione.ascorbate	-0.76	0.002	-0.91	0.001
Vitvi07g00037	glutathione reductase (AT3G54660 similarity)	21.2.2	redox.ascorbate and glutathione.glutathione	-0.49	0.046	-0.44	0.077
Vitvi14g02616	copper/zinc superoxide dismutase 1 (AT1G08830 similarity)	21.6	redox.dismutases and catalases	0.39	0.121	0.48	0.067
Vitvi13g00177	manganese superoxide dismutase 1 (AT3G10920 similarity)	21.6	redox.dismutases and catalases	0.39	0.011	0.74	0.000
Vitvi14g02607	copper/zinc superoxide dismutase 1 (AT1G08830 similarity)	21.6	redox.dismutases and catalases	-0.08	0.849	0.11	0.798
Vitvi10g01476	Fe superoxide dismutase 3 (AT5G23310 similarity)	21.6	redox.dismutases and catalases	-0.27	0.424	-0.04	0.935
Vitvi08g01802	copper/zinc superoxide dismutase 3 (AT5G18100 similarity)	21.6	redox.dismutases and catalases	-0.54	0.010	-0.59	0.010
Vitvi06g01349	copper/zinc superoxide dismutase 2 (AT2G28190 similarity)	21.6	redox.dismutases and catalases	-0.77	0.002	-0.69	0.004
Vitvi16g00280	Fe superoxide dismutase 2 (AT5G51100 similarity)	21.6	redox.dismutases and catalases	-1.22	0.000	-0.97	0.005
ABC TRANSPORTERS							
Vitvi01g01917	ABC-2 type transporter family protein (AT3G25620 similarity)	35.2	not assigned.unknown	4.61	0.000	1.85	0.302
Vitvi01g01915	ABC-2 type transporter family protein (AT3G25620 similarity)	35.2	not assigned.unknown	2.95	0.038	4.63	0.174
Vitvi17g00972	ABC transporter family protein (AT3G28345 similarity)	34.16	transport.ABC transporters and multidrug resistance systems	2.16	0.017	0.42	0.691

Vitvi06g01423	PF00005//PF01061//PF08370 - ABC transporter (ABC_tran) // ABC-2 type transporter (ABC2_membrane) // Plant PDR ABC transporter associated (PDR_assoc) (1 of 2)	34.16	transport.ABC transporters and multidrug resistance systems	1.87	0.425	2.70	0.003
Vitvi06g01423	PF00005//PF01061//PF08370 - ABC transporter (ABC_tran) // ABC-2 type transporter (ABC2_membrane) // Plant PDR ABC transporter associated (PDR_assoc) (1 of 2)	34.16	transport.ABC transporters and multidrug resistance systems	1.87	0.425	2.70	0.003
Vitvi01g01918	ABC-2 type transporter family protein (AT3G25620 similarity)	35.2	not assigned.unknown	1.45	0.111	3.01	0.205
Vitvi09g00470	PTHR19241:SF219 - ABC TRANSPORTER G FAMILY MEMBER 38 (1 of 30)	34.16	transport.ABC transporters and multidrug resistance systems	1.25	0.051	-0.23	0.762
Vitvi09g00294	PF00005//PF08370//PF14510 - ABC transporter (ABC_tran) // Plant PDR ABC transporter associated (PDR_assoc) // ABC-transporter extracellular N-terminal (ABC_trans_N) (1 of 1)	34.16	transport.ABC transporters and multidrug resistance systems	1.06	0.095	-1.69	0.022
Vitvi13g00638	ABC transporter family protein (AT5G60740 similarity)	34.16	transport.ABC transporters and multidrug resistance systems	0.94	0.003	0.49	0.099
Vitvi01g01908	ABC-2 type transporter family protein (AT3G25620 similarity)	35.2	not assigned.unknown	0.93	0.175	0.95	0.338
AUTOPHAGY							
Vitvi13g00112	autophagy-like protein (AT2G40316 similarity)	35.2	not assigned.unknown	0.39	0.027	0.45	0.014
Vitvi07g00210	yeast autophagy 18 G-like protein (AT1G03380 similarity)	29.5.2	protein.degradation.autophagy	0.30	0.085	0.35	0.042
Vitvi14g00317	autophagy 18h-like protein (AT1G54710 similarity)	29.5.2	protein.degradation.autophagy	0.22	0.266	0.28	0.158
Vitvi16g01309	Autophagy-related protein 13 (AT3G49590 similarity)	35.2	not assigned.unknown	0.21	0.133	0.44	0.003
Vitvi04g01617	Autophagy-related protein 13 (AT3G18770 similarity)	35.2	not assigned.unknown	0.19	0.224	0.22	0.205
Vitvi15g00986	AUTOPHAGY 8 (AT3G61710 similarity)	31.1	cell.organisation	0.05	0.653	-0.16	0.145
Vitvi11g00813	yeast autophagy 18 B-like protein (AT4G30510 similarity)	34.99	transport.misc	0.02	0.933	1.19	0.000
Vitvi11g00777	autophagy-like protein (AT4G30790 similarity)	35.2	not assigned.unknown	-0.01	0.980	0.11	0.550
Vitvi04g01615	Autophagy-related protein 13 (AT3G18770 similarity)	35.2	not assigned.unknown	-0.07	0.891	-0.53	0.285
Vitvi03g00492	autophagy 2 (AT3G19190 similarity)	35.2	not assigned.unknown	-0.11	0.540	0.29	0.093
Vitvi07g00580	autophagy 9 (APG9) (AT2G31260 similarity)	29.5.2	protein.degradation.autophagy	-0.11	0.507	0.16	0.331
Vitvi19g01985	yeast autophagy 18 F-like protein (AT5G54730 similarity)	29.5.2	protein.degradation.autophagy	-0.14	0.393	0.25	0.136
Vitvi02g00798	autophagy protein Apg5 family (AT5G17290 similarity)	29.5.2	protein.degradation.autophagy	-0.18	0.239	0.21	0.196
Vitvi13g00009	yeast autophagy-like protein (AT2G40810 similarity)	29.5.2	protein.degradation.autophagy	-0.29	0.221	-0.14	0.687
Vitvi04g01616	Autophagy-related protein 13 (AT3G18770 similarity)	35.2	not assigned.unknown	0.85	0.159	-1.02	0.180

Supplemental Table 3. Protein groups identified through protein pull-down and mass spectrometry analysis. A list of proteins with their UniProtKB identifiers and fasta headers, identified in the YFP-tagged candidate effector PoSTOSP28 sample, but not present in control sample after pull-down.

Protein IDs	Fasta headers	Number of proteins	Peptides
tr|Q8LSZ3|Q8LSZ3_TOBAC;tr|A0A1S4A5R2|A0A1S4A5R2_TOBAC;tr|A0A1S4B618|A0A1S4B618_TOBAC;tr|A0A1S3ZIP8|A0A1S3ZIP8_TOBAC	tr|Q8LSZ3|Q8LSZ3_TOBAC NADPH-protochlorophyllide oxidoreductase OS=Nicotiana tabacum OX=4097 GN=POR1 PE=2 SV=1;tr|A0A1S4A5R2 A0A1S4A5R2_TOBAC NADPH-protochlorophyllide oxidoreductase OS=Nicotiana tabacum OX=4097 GN=LOC107793976 PE=3 SV=1;tr|A0A1S4B618|A0A1	4	3
tr|A0A140G1T9|A0A140G1T9_TOBAC	tr|A0A140G1T9|A0A140G1T9_TOBAC ribosomal protein S18 OS=Nicotiana tabacum OX=4097 GN=rps18 PE=3 SV=1	1	2
tr|A0A1S3ZRR1 A0A1S3ZRR1_TOBAC;tr|A0A1S3YLV9|A0A1S3YLV9_TOBAC	tr|A0A1S3ZRR1|A0A1S3ZRR1_TOBAC nucleoid-associated protein At4g30620, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107789621 PE=4 SV=1;tr|A0A1S3YLV9 A0A1S3YLV9_TOBAC nucleoid-associated protein At4g30620, chloroplastic-like OS=Nicotiana tabacum OX	2	2
tr|A0A1S4B4M6|A0A1S4B4M6_TOBAC;tr|A0A1S4BHP4|A0A1S4BHP4_TOBAC;tr|A0A1S3Z8E1 A0A1S3Z8E1_TOBAC	tr|A0A1S4B4M6|A0A1S4B4M6_TOBAC 26S proteasome non-ATPase regulatory subunit 1 homolog OS=Nicotiana tabacum OX=4097 GN=LOC107804491 PE=3 SV=1;tr|A0A1S4BHP4|A0A1S4BHP4_TOBAC 26S proteasome non-ATPase regulatory subunit 1 homolog OS=Nicotiana tabacum OX=4097	3	5
tr|A0A1S4AZ72|A0A1S4AZ72_TOBAC;tr|A0A1S3YKL7 A0A1S3YKL7_TOBAC	tr|A0A1S4AZ72|A0A1S4AZ72_TOBAC thioredoxin-like 3-1, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107802806 PE=4 SV=1;tr|A0A1S3YKL7 A0A1S3YKL7_TOBAC thioredoxin-like 3-1, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107777304 PE=4 SV=1	2	2
tr|A0A1S4C4F5|A0A1S4C4F5_TOBAC	tr|A0A1S4C4F5|A0A1S4C4F5_TOBAC glutamate-1-semialdehyde 2,1-aminomutase, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107815022 PE=4 SV=1	1	3
tr|A0A1S4CKA0|A0A1S4CKA0_TOBAC	tr|A0A1S4CKA0|A0A1S4CKA0_TOBAC uncharacterized protein LOC107819940 isoform X1 OS=Nicotiana tabacum OX=4097 GN=LOC107819940 PE=4 SV=1	1	1
tr|Q9AT16|Q9AT16_TOBAC;tr|Q8H6P9|Q8H6P9_TOBAC;tr|Q84P55|Q84P55_TOBAC;tr|Q84P56|Q84P56_TOBAC	tr|Q9AT16|Q9AT16_TOBAC Ankyrin-repeat protein HBP1 OS=Nicotiana tabacum OX=4097 GN=LOC107793888 PE=2 SV=1;tr|Q8H6P9|Q8H6P9_TOBAC Ankyrin domain protein OS=Nicotiana tabacum OX=4097 GN=ANK1 PE=2 SV=1;tr|Q84P55|Q84P55_TOBAC TGB12K interacting protein 3 OS=Ni	4	4
tr|A0A1S4B341|A0A1S4B341_TOBAC;tr|A0A1S4AGR7|A0A1S4AGR7_TOBAC	tr|A0A1S4B341|A0A1S4B341_TOBAC cell division protein FtsY homolog, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107804044 PE=3 SV=1;tr|A0A1S4AGR7|A0A1S4AGR7_TOBAC cell division protein FtsY homolog, chloroplastic-like OS=Nicotiana tabacum OX=4097	2	3
tr|A0A1S4D985|A0A1S4D985_TOBAC;tr|A0A1S4ABS2|A0A1S4ABS2_TOBAC	tr|A0A1S4D985|A0A1S4D985_TOBAC serine--glyoxylate aminotransferase-like OS=Nicotiana tabacum OX=4097 GN=LOC107827351 PE=3 SV=1;tr|A0A1S4ABS2|A0A1S4ABS2_TOBAC serine–glyoxylate aminotransferase OS=Nicotiana tabacum OX=4097 GN=LOC107795801 PE=3 SV=1	2	3
tr|A0A1S4DJC0|A0A1S4DJC0_TOBAC;tr|A0A1S3Z5N4|A0A1S3Z5N4_TOBAC;sp|Q9ZS34|CHLP_TOBAC;tr|A0A1S4DJC8|A0A1S4DJC8_TOBAC;tr|A0A1S4AZJ0|A0A1S4AZJ0_TOBAC;tr|A0A1S3XXZ2|A0A1S3XXZ2_TOBAC	tr|A0A1S4DJC0|A0A1S4DJC0_TOBAC geranylgeranyl diphosphate reductase, chloroplastic-like isoform X1 OS=Nicotiana tabacum OX=4097 GN=LOC107830453 PE=4 SV=1;tr|A0A1S3Z5N4|A0A1S3Z5N4_TOBAC geranylgeranyl diphosphate reductase, chloroplastic OS=Nicotiana tabacu	6	4
tr|A0A140G1V2|A0A140G1V2_TOBAC	tr|A0A140G1V2|A0A140G1V2_TOBAC ribosomal protein S8 OS=Nicotiana tabacum OX=4097 GN=rps8 PE=3 SV=1	1	2
tr|A0A1S4A2J0|A0A1S4A2J0_TOBAC;tr|A0A1S3ZTF8|A0A1S3ZTF8_TOBAC	tr|A0A1S4A2J0|A0A1S4A2J0_TOBAC cysteine desulfurase 1, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107793129 PE=4 SV=1;tr|A0A1S3ZTF8|A0A1S3ZTF8_TOBAC cysteine desulfurase 1, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107790197 PE=4 SV=	2	2
tr|A0A1S4CZJ7|A0A1S4CZJ7_TOBAC;tr|A0A1S3YAN6|A0A1S3YAN6_TOBAC	tr|A0A1S4CZJ7|A0A1S4CZJ7_TOBAC COBW domain-containing protein 1-like OS=Nicotiana tabacum OX=4097 GN=LOC107824330 PE=4 SV=1;tr|A0A1S3YAN6|A0A1S3YAN6_TOBAC COBW domain-containing protein 1-like OS=Nicotiana tabacum OX=4097 GN=LOC107774293 PE=4 SV=1	2	2
tr|A0A1S3Y035|A0A1S3Y035_TOBAC;tr|A0A1S4A0C9|A0A1S4A0C9_TOBAC	tr|A0A1S3Y035|A0A1S3Y035_TOBAC ribulose bisphosphate carboxylase/oxygenase activase, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107770767 PE=4 SV=1;tr|A0A1S4A0C9|A0A1S4A0C9_TOBAC ribulose bisphosphate carboxylase/oxygenase activase, chloroplasti	2	3
tr|Q9FES7|Q9FES7_TOBAC;tr|A0A1S3XEV7|A0A1S3XEV7_TOBAC;sp|O22436|CHLI_TOBAC;tr|A0A1S3Z7S1|A0A1S3Z7S1_TOBAC	tr|Q9FES7|Q9FES7_TOBAC Mg-protoporphyrin IX chelatase OS=Nicotiana tabacum OX=4097 GN=Su-s PE=2 SV=1;tr|A0A1S3XEV7|A0A1S3XEV7_TOBAC Mg-protoporphyrin IX chelatase OS=Nicotiana tabacum OX=4097 GN=LOC107764312 PE=3 SV=1;sp|O22436|CHLI_TOBAC Magnesium-chelata	4	3
tr|A0A1S4DKC4|A0A1S4DKC4_TOBAC;tr|A0A1S3YLS7|A0A1S3YLS7_TOBAC;tr|A0A1S4CLW9|A0A1S4CLW9_TOBAC;tr|A0A1S4AFU6|A0A1S4AFU6_TOBAC	tr|A0A1S4DKC4|A0A1S4DKC4_TOBAC chaperone protein dnaJ A6, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107830729 PE=4 SV=1;tr|A0A1S3YLS7 A0A1S3YLS7_TOBAC chaperone protein dnaJ A6, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107777513 PE	4	2
tr|A0A1S4BJ77|A0A1S4BJ77_TOBAC	tr|A0A1S4BJ77|A0A1S4BJ77_TOBAC Tryptophan synthase OS=Nicotiana tabacum OX=4097 GN=LOC107808897 PE=3 SV=1	1	2
tr|A0A1S4CMV5 A0A1S4CMV5_TOBAC;tr|A0A1S4BWP8|A0A1S4BWP8_TOBAC;tr|A0A1S3ZRW6 A0A1S3ZRW6_TOBAC;tr|A0A1S4A6I2|A0A1S4A6I2_TOBAC;tr|A0A1S3ZTA5|A0A1S3ZTA5_TOBAC	tr|A0A1S4CMV5|A0A1S4CMV5_TOBAC 26S proteasome non-ATPase regulatory subunit 2 homolog OS=Nicotiana tabacum OX=4097 GN=LOC107820776 PE=3 SV=1;tr|A0A1S4BWP8|A0A1S4BWP8_TOBAC 26S proteasome non-ATPase regulatory subunit 2 homolog A-like OS=Nicotiana tabacum O	5	4
tr|A0A1S4AZ18|A0A1S4AZ18_TOBAC;tr|A0A1S3YZJ0|A0A1S3YZJ0_TOBAC	tr|A0A1S4AZ18|A0A1S4AZ18_TOBAC ATP sulfurylase 1, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107802851 PE=4 SV=1;tr|A0A1S3YZJ0|A0A1S3YZJ0_TOBAC ATP sulfurylase 1, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107781534 PE=4 SV=1	2	5
tr|A0A076L2E2|A0A076L2E2_TOBAC;tr|A0A1S3ZK73|A0A1S3ZK73_TOBAC	tr|A0A076L2E2|A0A076L2E2_TOBAC 40S ribosomal protein S25 OS=Nicotiana tabacum OX=4097 GN=LOC107768022 PE=2 SV=1;tr|A0A1S3ZK73|A0A1S3ZK73_TOBAC 40S ribosomal protein S25-like OS=Nicotiana tabacum OX=4097 GN=LOC107787568 PE=4 SV=1	2	2
tr|A0A1S4CGP3|A0A1S4CGP3_TOBAC;tr|A0A1S3X430|A0A1S3X430_TOBAC;tr|A0A1S4A807|A0A1S4A807_TOBAC;tr|A0A1S4AP69|A0A1S4AP69_TOBAC;tr|A0A1S4B9L9|A0A1S4B9L9_TOBAC	tr|A0A1S4CGP3|A0A1S4CGP3_TOBAC 60S ribosomal protein L14-2-like OS=Nicotiana tabacum OX=4097 GN=LOC107818785 PE=4 SV=1;tr|A0A1S3X430|A0A1S3X430_TOBAC probable 60S ribosomal protein L14 OS=Nicotiana tabacum OX=4097 GN=LOC107760861 PE=4 SV=1;tr|A0A1S4A807|A0	5	4

tr|A0A1S3ZEQ4|A0A1S3ZEQ4_TOBAC	tr|A0A1S3ZEQ4|A0A1S3ZEQ4_TOBAC GDSL esterase/lipase At1g29670-like OS=Nicotiana tabacum OX=4097 GN=LOC107785876 PE=4 SV=1	1	1
tr|A0A1S4AJL5|A0A1S4AJL5_TOBAC;tr|A0A1S4CRI6|A0A1S4CRI6_TOBAC;tr|A0A1S4AJC6|A0A1S4AJC6_TOBAC;tr|A0A1S4CRF2|A0A1S4CRF2_TOBAC;tr|A0A1S4BGN4|A0A1S4BGN4_TOBAC;tr|A0A1S3XYY2|A0A1S3XYY2_TOBAC	tr|A0A1S4AJL5|A0A1S4AJL5_TOBAC 26S proteasome non-ATPase regulatory subunit 4 homolog isoform X1 OS=Nicotiana tabacum OX=4097 GN=LOC107798338 PE=4 SV=1;tr|A0A1S4CRI6|A0A1S4CRI6_TOBAC 26S proteasome non-ATPase regulatory subunit 4 homolog isoform X2 OS=Nico	6	3
tr|A0A1S4CQ71|A0A1S4CQ71_TOBAC;tr|A0A1S4AA87|A0A1S4AA87_TOBAC;sp|Q05214|ACT1_TOBAC;tr|A0A1S4BAV7|A0A1S4BAV7_TOBAC;tr|A0A1S4C557|A0A1S4C557_TOBAC;tr|A0A1S4C1C0|A0A1S4C1C0_TOBAC;tr|A0A1S3ZT51|A0A1S3ZT51_TOBAC;sp|P93371|ACT5_TOBAC;tr|A0A1S3X7J8|A0A1S3X7J8_TOBAC;tr|A0A1S3YSA4|A0A1S3YSA4_TOBAC	tr|A0A1S4CQ71|A0A1S4CQ71_TOBAC actin-like OS=Nicotiana tabacum OX=4097 GN=LOC107821481 PE=3 SV=1;tr|A0A1S4AA87|A0A1S4AA87_TOBAC actin OS=Nicotiana tabacum OX=4097 GN=LOC107795436 PE=3 SV=1;sp|Q05214|ACT1_TOBAC Actin OS=Nicotiana tabacum OX=4097 PE=3 SV=1;t	10	11
tr|A0A1S3XE72|A0A1S3XE72_TOBAC;tr|A0A1S4DPW9|A0A1S4DPW9_TOBAC	tr|A0A1S3XE72|A0A1S3XE72_TOBAC 60S ribosomal protein L22-3 OS=Nicotiana tabacum OX=4097 GN=LOC107764146 PE=4 SV=1;tr|A0A1S4DPW9|A0A1S4DPW9_TOBAC 60S ribosomal protein L22-3-like OS=Nicotiana tabacum OX=4097 GN=LOC107832173 PE=4 SV=1	2	1
tr|A0A1S3YFA7|A0A1S3YFA7_TOBAC;tr|A0A1S3YF92|A0A1S3YF92_TOBAC;tr|A0A1S3ZYZ3|A0A1S3ZYZ3_TOBAC	tr|A0A1S3YFA7|A0A1S3YFA7_TOBAC flagellar radial spoke protein 5-like isoform X2 OS=Nicotiana tabacum OX=4097 GN=LOC107775472 PE=4 SV=1;tr|A0A1S3YF92|A0A1S3YF92_TOBAC flagellar radial spoke protein 5-like isoform X1 OS=Nicotiana tabacum OX=4097 GN=LOC107775	3	3
tr|Q6QND3|Q6QND3_TOBAC;tr|A0A1S4B745|A0A1S4B745_TOBAC;tr|A0A1S4CJY5|A0A1S4CJY5_TOBAC;tr|A0A1S3XXA7|A0A1S3XXA7_TOBAC	tr|Q6QND3|Q6QND3_TOBAC Putative pyridoxine biosyntheais protein isoform A OS=Nicotiana tabacum OX=4097 GN=Pdx1-B PE=3 SV=1;tr|A0A1S4B745|A0A1S4B745_TOBAC pyridoxal 5-phosphate synthase subunit PDX1 OS=Nicotiana tabacum OX=4097 GN=LOC107805188 PE=3 SV=1	4	3
tr|A0A1S4DD77|A0A1S4DD77_TOBAC;tr|A0A1S4CCG1|A0A1S4CCG1_TOBAC;sp|O80361|RK4_TOBAC	tr|A0A1S4DD77|A0A1S4DD77_TOBAC 50S ribosomal protein L4, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107828567 PE=3 SV=1;tr|A0A1S4CCG1|A0A1S4CCG1_TOBAC 50S ribosomal protein L4, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107817573 PE=3 SV=1	3	2
tr|A0A1S4D6G4|A0A1S4D6G4_TOBAC;tr|A0A1S3XGJ0|A0A1S3XGJ0_TOBAC	tr|A0A1S4D6G4|A0A1S4D6G4_TOBAC 40S ribosomal protein S30 OS=Nicotiana tabacum OX=4097 GN=LOC107826518 PE=3 SV=1;tr|A0A1S3XGJ0|A0A1S3XGJ0_TOBAC 40S ribosomal protein S30 OS=Nicotiana tabacum OX=4097 GN=LOC107764876 PE=3 SV=1	2	1
tr|A0A1S4CLB7|A0A1S4CLB7_TOBAC	tr|A0A1S4CLB7|A0A1S4CLB7_TOBAC aspartate-semialdehyde dehydrogenase-like OS=Nicotiana tabacum OX=4097 GN=LOC107820046 PE=3 SV=1	1	1
tr|A0A1S4A4D2|A0A1S4A4D2_TOBAC;tr|A0A1S4ATT2|A0A1S4ATT2_TOBAC;tr|A0A1S4A4N5|A0A1S4A4N5_TOBAC	tr|A0A1S4A4D2|A0A1S4A4D2_TOBAC fe-S cluster assembly factor HCF101, chloroplastic-like isoform X2 OS=Nicotiana tabacum OX=4097 GN=LOC107793652 PE=3 SV=1;tr|A0A1S4ATT2|A0A1S4ATT2_TOBAC fe-S cluster assembly factor HCF101, chloroplastic-like OS=Nicotiana tab	3	2
tr|A0A1S3XBY2|A0A1S3XBY2_TOBAC;tr|A0A1S3X8Q5|A0A1S3X8Q5_TOBAC;sp|O80362|RK10_TOBAC	tr|A0A1S3XBY2|A0A1S3XBY2_TOBAC 50S ribosomal protein L10, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107763454 PE=3 SV=1;tr|A0A1S3X8Q5|A0A1S3X8Q5_TOBAC 50S ribosomal protein L10, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107762422 PE=3 SV	3	2
tr|A0A1S3ZVR6|A0A1S3ZVR6_TOBAC;sp|P27494|CB23_TOBAC;tr|A0A1S3YPE9|A0A1S3YPE9_TOBAC;tr|A0A1S3ZSU2|A0A1S3ZSU2_TOBAC;tr|A0A1S3ZVQ8|A0A1S3ZVQ8_TOBAC;tr|A0A1S4BMB0|A0A1S4BMB0_TOBAC;tr|A0A1S3ZYV8|A0A1S3ZYV8_TOBAC;tr|A0A1S3XEU1|A0A1S3XEU1_TOBAC;tr|A0A1S3YWG7|A0A1S3YWG7_TOBAC;tr|A0A1S3Y974|A0A1S3Y974_TOBAC;sp|P27491|CB27_TOBAC;tr|A0A1S3Y8B9|A0A1S3Y8B9_TOBAC;tr|A0A1S3Y8Q2|A0A1S3Y8Q2_TOBAC;tr|A0A1S4BC47|A0A1S4BC47_TOBAC:sp|P27495|CB24_TOBAC;sp|P27496|CB25_TOBAC;tr|A0A1S3Y100|A0A1S3Y100_TOBAC;sp|P27492|CB21_TOBAC;tr|A0A1S3YD73|A0A1S3YD73_TOBAC;tr|A0A1S4BG28|A0A1S4BG28_TOBAC;tr|A0A1S4BG47|A0A1S4BG47_TOBAC;tr|A0A1S4BPJ2|A0A1S4BPJ2_TOBAC;tr|A0A1S4A8I3|A0A1S4A8I3_TOBAC;tr|A0A1S4BG25|A0A1S4BG25_TOBAC:tr|A0A1S4BPJ4|A0A1S4BPJ4_TOBAC;sp|P27493|CB22_TOBAC;tr|A0A1S4BPM1|A0A1S4BPM1_TOBAC;tr|A0A1S3Y6H9|A0A1S3Y6H9_TOBAC;tr|A0A1S3Y8R2|A0A1S3Y8R2_TOBAC;tr|A0A1S3Y8G4|A0A1S3Y8G4_TOBAC;tr|A0A1S3Y820|A0A1S3Y820_TOBAC	tr|A0A1S3ZVR6|A0A1S3ZVR6_TOBAC Chlorophyll a-b binding protein, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107790955 PE=3 SV=1;sp|P27494|CB23_TOBAC Chlorophyll a-b binding protein 36, chloroplastic OS=Nicotiana tabacum OX=4097 GN=CAB36 PE=3 SV=1;tr|A	31	2
tr|A0A1S4D708|A0A1S4D708_TOBAC;tr|A0A1S4D6S4|A0A1S4D6S4_TOBAC;tr|A0A1S4D2I0|A0A1S4D2I0_TOBAC;tr|A0A1S4D2Z3|A0A1S4D2Z3_TOBAC	tr|A0A1S4D708|A0A1S4D708_TOBAC uncharacterized protein ycf39-like isoform X2 OS=Nicotiana tabacum OX=4097 GN=LOC107826653 PE=4 SV=1;tr|A0A1S4D6S4 A0A1S4D6S4_TOBAC uncharacterized protein ycf39-like isoform X1 OS=Nicotiana tabacum OX=4097 GN=LOC107826653 PE	4	2
tr|A0A1S4CJ62|A0A1S4CJ62_TOBAC;tr|A0A1S4C068|A0A1S4C068_TOBAC;tr|A0A1S4AYN5|A0A1S4AYN5_TOBAC;tr|A0A1S3XV78|A0A1S3XV78_TOBAC	tr|A0A1S4CJ62|A0A1S4CJ62_TOBAC 60S ribosomal protein L9-1-like OS=Nicotiana tabacum OX=4097 GN=LOC107819630 PE=4 SV=1;tr|A0A1S4C068|A0A1S4C068_TOBAC 60S ribosomal protein L9-1-like OS=Nicotiana tabacum OX=4097 GN=LOC107813756 PE=4 SV=1;tr|A0A1S4AYN5|A0A1S4	4	2
tr|A0A1S4D9E4|A0A1S4D9E4_TOBAC;tr|A0A1S4CYL2|A0A1S4CYL2_TOBAC	tr|A0A1S4D9E4|A0A1S4D9E4_TOBAC pheophytinase, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107827246 PE=4 SV=1;tr|A0A1S4CYL2|A0A1S4CYL2_TOBAC pheophytinase, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107824051 PE=4 SV=1	2	2
tr|A0A1S3ZX13|A0A1S3ZX13_TOBAC;tr|A0A1S3WYR8|A0A1S3WYR8_TOBAC	tr|A0A1S3ZX13|A0A1S3ZX13_TOBAC Fructose-bisphosphate aldolase OS=Nicotiana tabacum OX=4097 GN=LOC107791431 PE=3 SV=1;tr|A0A1S3WYR8|A0A1S3WYR8_TOBAC Fructose-bisphosphate aldolase OS=Nicotiana tabacum OX=4097 GN=LOC107759369 PE=3 SV=1	2	7
tr|A0A1S3XHG4|A0A1S3XHG4_TOBAC;tr|A0A1S4AIH6|A0A1S4AIH6_TOBAC;tr|A0A1S4A6S4|A0A1S4A6S4_TOBAC	tr|A0A1S3XHG4|A0A1S3XHG4_TOBAC phosphoglucomutase, cytoplasmic-like OS=Nicotiana tabacum OX=4097 GN=LOC107765291 PE=4 SV=1;tr|A0A1S4AIH6|A0A1S4AIH6_TOBAC phosphoglucomutase, cytoplasmic-like OS=Nicotiana tabacum OX=4097 GN=LOC107798006 PE=4 SV=1;tr|A0A1S4A	3	1
tr|A0A1S3X0Y5|A0A1S3X0Y5_TOBAC	tr|A0A1S3X0Y5|A0A1S3X0Y5_TOBAC Ferredoxin--NADP reductase, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107760150 PE=3 SV=1	1	1
tr|A0A1S3Z6E6|A0A1S3Z6E6_TOBAC;tr|A0A1S4DC16|A0A1S4DC16_TOBAC	tr|A0A1S3Z6E6|A0A1S3Z6E6_TOBAC V-type proton ATPase subunit H OS=Nicotiana tabacum OX=4097 GN=LOC107783503 PE=3 SV=1;tr|A0A1S4DC16|A0A1S4DC16_TOBAC V-type proton ATPase subunit H OS=Nicotiana tabacum OX=4097 GN=LOC107828022 PE=3 SV=1	2	2
tr|A0A1S4CW25|A0A1S4CW25_TOBAC	tr|A0A1S4CW25|A0A1S4CW25_TOBAC probable inactive shikimate kinase like 2, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107823221 PE=4 SV=1	1	2

tr|A0A1S4DNT1 A0A1S4DNT1_TOBAC;tr|A0A1S3ZGF1|A0A1S3ZGF1_TOBAC;tr|A0A1S4DNS6|A0A1S4DNS6_TOBAC	tr|A0A1S4DNT1|A0A1S4DNT1_TOBAC Eukaryotic translation initiation factor 3 subunit I OS=Nicotiana tabacum OX=4097 GN=LOC107831796 PE=3 SV=1;tr|A0A1S3ZGF1|A0A1S3ZGF1_TOBAC Eukaryotic translation initiation factor 3 subunit I OS=Nicotiana tabacum OX=4097 GN=L	3	2
tr|A0A140G1Q1 A0A140G1Q1_TOBAC	tr|A0A140G1Q1|A0A140G1Q1_TOBAC ribosomal protein S2 OS=Nicotiana tabacum OX=4097 GN=rps2 PE=3 SV=1	1	2
tr|A0A1S4DC28|A0A1S4DC28_TOBAC;tr|A0A1S3X053|A0A1S3X053_TOBAC;tr|A0A1S3WZM7|A0A1S3WZM7_TOBAC;tr|A0A1S4DC59|A0A1S4DC59_TOBAC	tr|A0A1S4DC28|A0A1S4DC28_TOBAC glutathionyl-hydroquinone reductase YqjG-like isoform X2 OS=Nicotiana tabacum OX=4097 GN=LOC107828221 PE=4 SV=1;tr|A0A1S3X053|A0A1S3X053_TOBAC glutathionyl-hydroquinone reductase YqjG-like isoform X2 OS=Nicotiana tabacum OX=4	4	2
tr|A0A1S3XCY9|A0A1S3XCY9_TOBAC;tr|A0A1S4C7S3|A0A1S4C7S3_TOBAC	tr|A0A1S3XCY9|A0A1S3XCY9_TOBAC Protein transport protein Sec61 subunit beta OS=Nicotiana tabacum OX=4097 GN=LOC107763814 PE=3 SV=1;tr|A0A1S4C7S3|A0A1S4C7S3_TOBAC Protein transport protein Sec61 subunit beta OS=Nicotiana tabacum OX=4097 GN=LOC107816002 PE=3	2	1
tr|A0A1S4C8W0|A0A1S4C8W0_TOBAC;tr|A0A1S4C9D8|A0A1S4C9D8_TOBAC;tr|A0A1S4CX44|A0A1S4CX44_TOBAC;tr|A0A1S4C8W4|A0A1S4C8W4_TOBAC;tr|A0A1S3Z7W7|A0A1S3Z7W7_TOBAC;tr|A0A1S4AZJ1|A0A1S4AZJ1_TOBAC;tr|A0A1S3Z2N3|A0A1S3Z2N3_TOBAC;tr|A0A1S3Z750|A0A1S3Z750_TOBAC	tr|A0A1S4C8W0|A0A1S4C8W0_TOBAC dihydrolipoyl dehydrogenase 2, chloroplastic-like isoform X3 OS=Nicotiana tabacum OX=4097 GN=LOC107816379 PE=3 SV=1;tr|A0A1S4C9D8|A0A1S4C9D8_TOBAC dihydrolipoyl dehydrogenase 2, chloroplastic-like isoform X2 OS=Nicotiana taba	8	1
tr|A0A1S4ARQ1|A0A1S4ARQ1_TOBAC;tr|A0A1S3Z100|A0A1S3Z100_TOBAC	tr|A0A1S4ARQ1|A0A1S4ARQ1_TOBAC calcium sensing receptor, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107800572 PE=4 SV=1;tr|A0A1S3Z100|A0A1S3Z100_TOBAC calcium sensing receptor, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107781793 PE=4	2	2
tr|A0A1S4A473|A0A1S4A473_TOBAC;tr|A0A1S3XMD5|A0A1S3XMD5_TOBAC;tr|A0A1S4AA32|A0A1S4AA32_TOBAC;tr|A0A1S3YRY3|A0A1S3YRY3_TOBAC;tr|A0A1S3YRG8|A0A1S3YRG8_TOBAC	tr|A0A1S4A473|A0A1S4A473_TOBAC ATP-citrate synthase beta chain protein 2-like OS=Nicotiana tabacum OX=4097 GN=LOC107793546 PE=4 SV=1;tr|A0A1S3XMD5|A0A1S3XMD5_TOBAC ATP-citrate synthase beta chain protein 2-like OS=Nicotiana tabacum OX=4097 GN=LOC107766582	5	3
tr|A0A1S3ZKZ4|A0A1S3ZKZ4_TOBAC;tr|A0A1S3Z828|A0A1S3Z828_TOBAC;sp|Q42962|PGKY_TOBAC;tr|A0A1S4CGR1|A0A1S4CGR1_TOBAC;tr|A0A1S4BTZ7|A0A1S4BTZ7_TOBAC	tr|A0A1S3ZKZ4|A0A1S3ZKZ4_TOBAC Phosphoglycerate kinase OS=Nicotiana tabacum OX=4097 GN=LOC107787830 PE=3 SV=1;tr|A0A1S3Z828|A0A1S3Z828_TOBAC Phosphoglycerate kinase OS=Nicotiana tabacum OX=4097 GN=LOC107783850 PE=3 SV=1;sp|Q42962|PGKY_TOBAC Phosphoglycerat	5	3
tr|A0A1S3Z6N1|A0A1S3Z6N1_TOBAC;tr|A0A1S3YPJ3|A0A1S3YPJ3_TOBAC	tr|A0A1S3Z6N1|A0A1S3Z6N1_TOBAC glyoxylate/succinic semialdehyde reductase 2, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107783419 PE=4 SV=1;tr|A0A1S3YPJ3|A0A1S3YPJ3_TOBAC glyoxylate/succinic semialdehyde reductase 2, chloroplastic-like OS=Nicoti	2	2
tr|A0A1S4AFK0 A0A1S4AFK0_TOBAC;tr|A0A1S4AKS8|A0A1S4AKS8_TOBAC	tr|A0A1S4AFK0|A0A1S4AFK0_TOBAC Thiamine thiazole synthase, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107797059 PE=3 SV=1;tr|A0A1S4AKS8|A0A1S4AKS8_TOBAC Thiamine thiazole synthase, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107798781 PE=3 SV=1	2	2
tr|A0A1S4DI23|A0A1S4DI23_TOBAC;tr|A0A1S4BRE4|A0A1S4BRE4_TOBAC;tr|A0A1S4BRG6|A0A1S4BRG6_TOBAC;tr|A0A1S4BRE8|A0A1S4BRE8_TOBAC;tr|A0A1S4BRE2|A0A1S4BRE2_TOBAC	tr|A0A1S4DI23|A0A1S4DI23_TOBAC glycine--tRNA ligase, chloroplastic/mitochondrial 2-like OS=Nicotiana tabacum OX=4097 GN=LOC107830076 PE=3 SV=1;tr|A0A1S4BRE4|A0A1S4BRE4_TOBAC glycine--tRNA ligase, chloroplastic/mitochondrial 2 isoform X4 OS=Nicotiana tabacu	5	2
tr|A0A1S4CNI5|A0A1S4CNI5_TOBAC;tr|A0A1S4C789|A0A1S4C789_TOBAC	tr|A0A1S4CNI5|A0A1S4CNI5_TOBAC presequence protease 1, chloroplastic/mitochondrial-like OS=Nicotiana tabacum OX=4097 GN=LOC107820951 PE=4 SV=1;tr|A0A1S4C789|A0A1S4C789_TOBAC presequence protease 1, chloroplastic/mitochondrial-like OS=Nicotiana tabacum OX=4	2	1
tr|A0A1S4DGZ9|A0A1S4DGZ9_TOBAC;tr|A0A1S3X395|A0A1S3X395_TOBAC	tr|A0A1S4DGZ9|A0A1S4DGZ9_TOBAC probable plastid-lipid-associated protein 13, chloroplastic OS=Nicotiana tabacum OX=4097 GN=LOC107829677 PE=4 SV=1;tr|A0A1S3X395|A0A1S3X395_TOBAC probable plastid-lipid-associated protein 13, chloroplastic OS=Nicotiana tabacu	2	2
tr|A0A1S3ZKM6|A0A1S3ZKM6_TOBAC;tr|A0A1S3YCR6|A0A1S3YCR6_TOBAC;tr|Q84QE8|Q84QE8_TOBAC;tr|A0A1S4CLK5|A0A1S4CLK5_TOBAC;tr|A0A1S3ZHB8|A0A1S3ZHB8_TOBAC;tr|A0A1S3XLS3|A0A1S3XLS3_TOBAC;sp|Q40458|PSBO_TOBAC,tr|A0A1S4BSA2|A0A1S4BSA2_TOBAC	tr|A0A1S3ZKM6|A0A1S3ZKM6_TOBAC oxygen-evolving enhancer protein 1, chloroplastic-like OS=Nicotiana tabacum OX=4097 GN=LOC107787712 PE=4 SV=1;tr|A0A1S3YCR6|A0A1S3YCR6_TOBAC oxygen-evolving enhancer protein 1, chloroplastic-like OS=Nicotiana tabacum OX=4097	8	1
tr|A0A1S3XIX6|A0A1S3XIX6_TOBAC;tr|A0A1S4BMB4|A0A1S4BMB4_TOBAC	tr|A0A1S3XIX6|A0A1S3XIX6_TOBAC 2,3-dimethylmalate lyase-like OS=Nicotiana tabacum OX=4097 GN=LOC107765651 PE=4 SV=1;tr|A0A1S4BMB4|A0A1S4BMB4_TOBAC 2,3-dimethylmalate lyase-like OS=Nicotiana tabacum OX=4097 GN=LOC107809819 PE=4 SV=1	2	1
tr|A0A1S4CIK7|A0A1S4CIK7_TOBAC;tr|A0A1S3YQ39|A0A1S3YQ39_TOBAC;tr|A0A1S3ZJY4|A0A1S3ZJY4_TOBAC;tr|A0A1S3Z039|A0A1S3Z039_TOBAC;tr|A0A1S4ANL0|A0A1S4ANL0_TOBAC;tr|A0A1S3YHK3|A0A1S3YHK3_TOBAC	tr|A0A1S4CIK7|A0A1S4CIK7_TOBAC ATP-citrate synthase alpha chain protein 2-like OS=Nicotiana tabacum OX=4097 GN=LOC107819406 PE=4 SV=1;tr|A0A1S3YQ39|A0A1S3YQ39_TOBAC ATP-citrate synthase alpha chain protein 2 OS=Nicotiana tabacum OX=4097 GN=LOC107778559 PE=	6	2

C.3 Supplementary Material for Publication 2.3

Supplementary material is available at https://www.frontiersin.org/articles/10.3389/fpls.2022.935819/full. Accessed: 26-July-2022.

Š. Tomaž, K. Gruden and A. Coll. "TGA Transcription Factors – Structural Characteristics as Basis for Functional Variability," *Frontiers in Plant Science*, vol. 13, no. 935819, 2022.

C.4 Supplementary Material for Publication 2.4

Supplementary Information

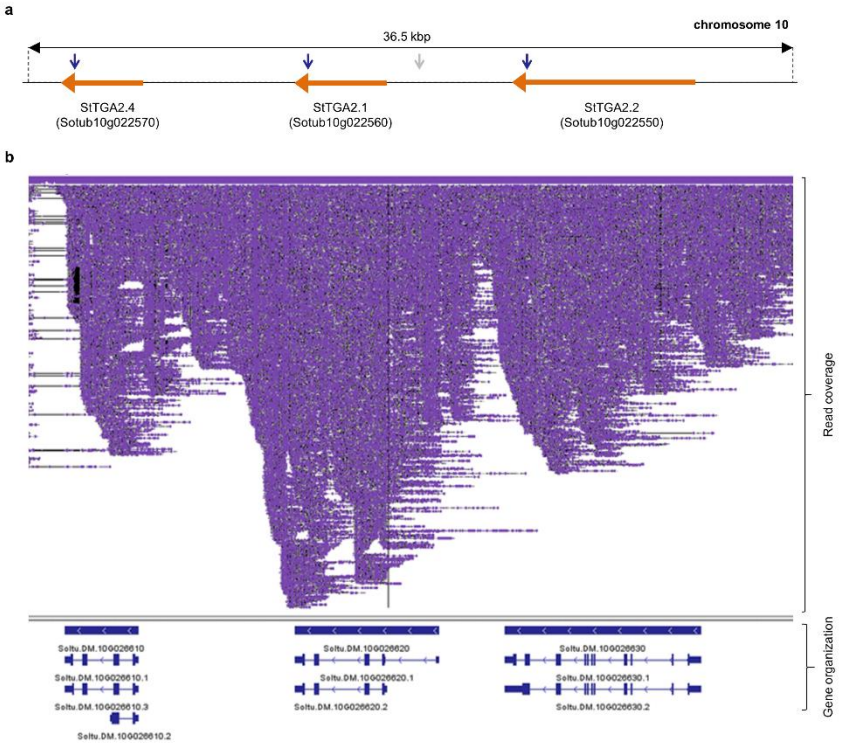
Title: A mini-TGA protein modulates gene expression through heterogeneous association with transcription factors

Authors: Špela Tomaž*, Marko Petek, Tjaša Lukan, Karmen Pogačar, Katja Stare, Erica Teixeira Prates, Daniel A. Jacobson, Jan Zrimec, Gregor Bajc, Matej Butala, Maruša Pompe Novak, Quentin Dudley, Nicola Patron, Ajda Taler-Verčič, Aleksandra Usenik, Dušan Turk, Salomé Prat, Anna Coll** & Kristina Gruden**

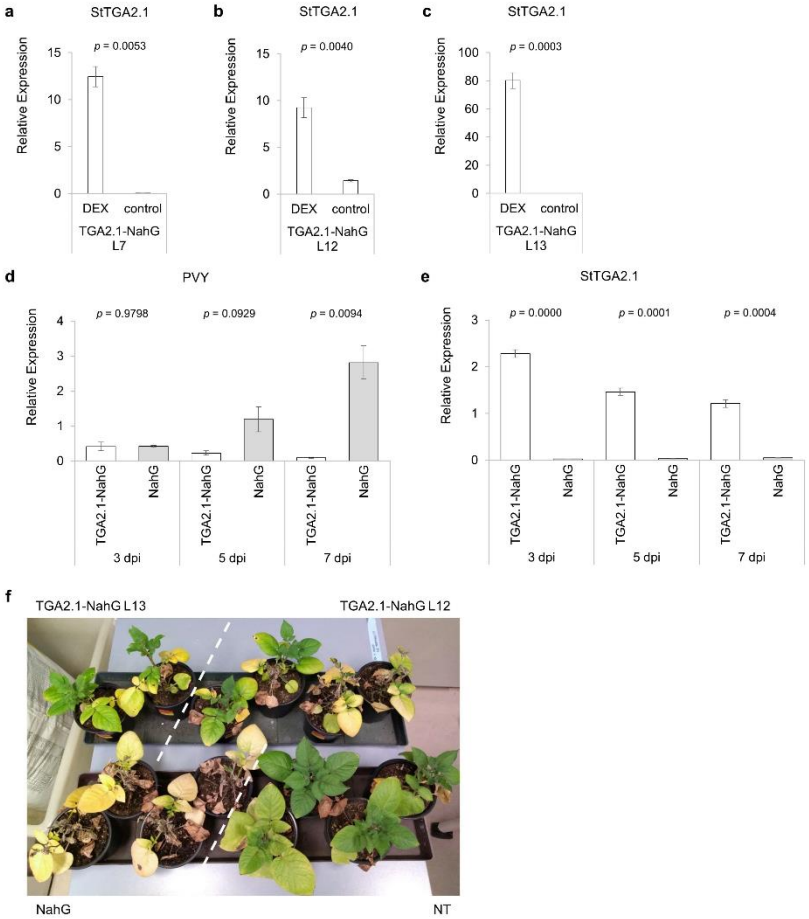
*Corresponding author: Š. Tomaž (spela.tomaz@nib.si), +386 (0)59 232 833.

**Authors contributed equally and share the last authorship.

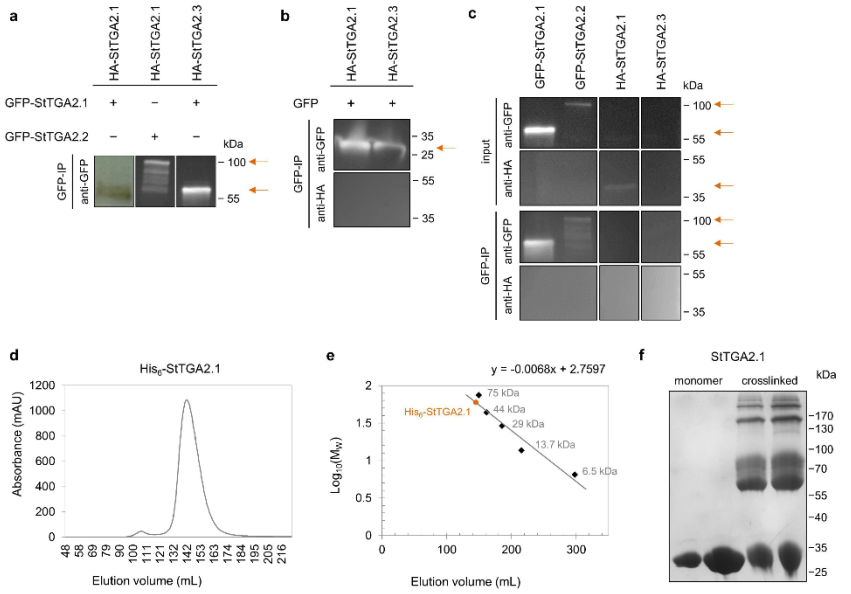
Supplementary Figures



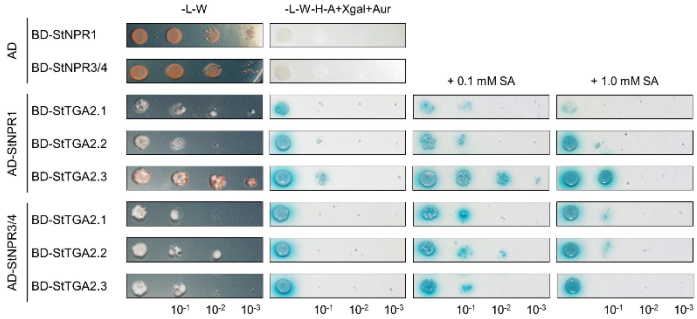
Supplementary Fig. 1. Targeted long-read sequencing confirms the presence of mini-TGAs in potato genome. a, Schematic representation of a ~36.5 kbp region on chromosome 10 in the double-monoploid (DM) potato reference genome v6.1 (Pham et al., 2020), encompassing the three tandemly repeated StTGA genes, *StTGA2.1*, *StTGA2.2* and *StTGA2.4* (ROI). All three genes are located on the negative strand. Approximate annealing sites of primer pairs used in targeted long-read sequencing are indicated, with primer pair A (blue arrows) targeting all three genes, and primer pair B (grey arrow) targeting a central part of the ROI. The primers are listed in Supplementary Table 11. **b,** Screenshot, showing the read mapping (purple), obtained through targeted long-read sequencing of the ROI in the tetraploid potato cultivar Rywal, to the potato reference genome v6.1 (Pham et al., 2020). Complete sequence coverage (mean depth 569 at 99.6 % coverage of the ROI) confirms the presence and organization of the three StTGAs. Gene organization with several predicted alternative splicing variants, as represented in the DM genome with corresponding transcript IDs, is depicted at the bottom of the figure.



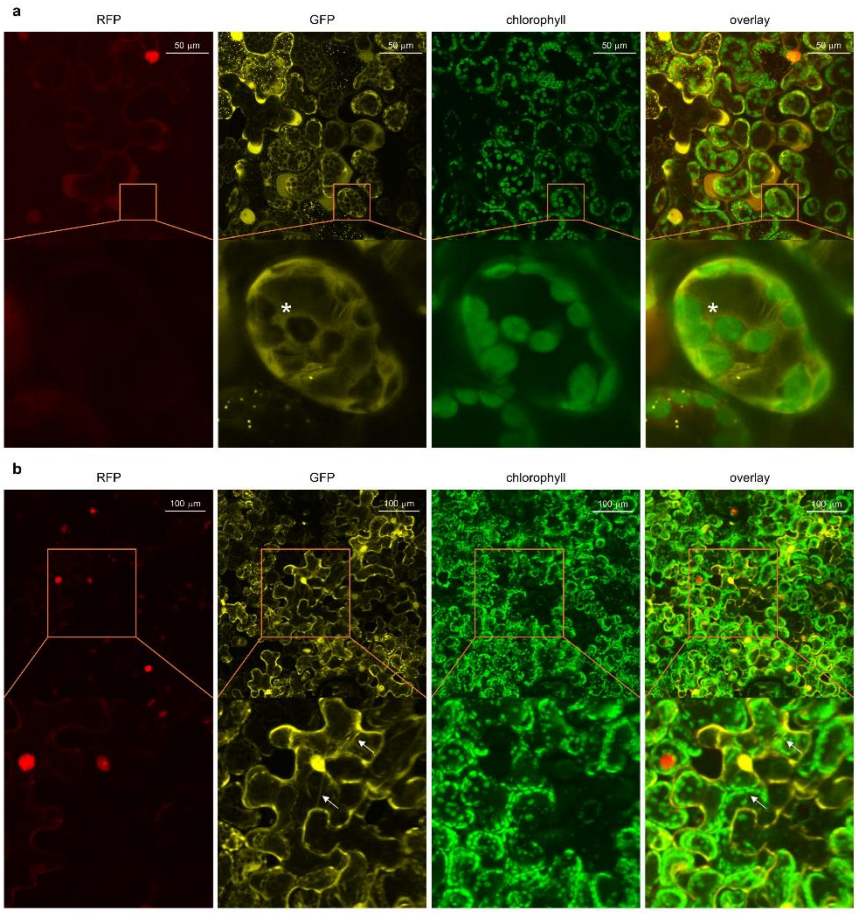
Supplementary Fig. 2. PVY replication in the second salicylic acid-deficient transgenic line overexpressing *StTGA2.1*. Relative expression levels of *StTGA2.1* in three NahG transgenic lines overexpressing *StTGA2.1* (TGA2.1-NahG), **a,** line 7 (L7), **b,** line 12 (L12) and **c**, line 13 (L13), three hours after dexamethasone (DEX) treatment, compared to non-treated plants (control). Average values ± standard error from three biological replicates for DEX treatment and **a,** two or **b, c,** three replicates for control are shown. Relative expression levels of **d,** PVY and **e,** *StTGA2.1* in PVY-infected leaves of DEX-treated TGA2.1-NahG L7 (white) and NahG (light grey) plants at 3, 5 and 7 days post infection (dpi). Average values ± standard error from three biological replicates are shown. **a – e,** Significance was determined using a two-tailed *t*-test. **f,** Phenotypic differences in TGA2.1-NahG L12 and L13, NahG and non-transgenic potato (NT) plants at 32 dpi. Plants from different genotypes are delineated with white dotted lines.



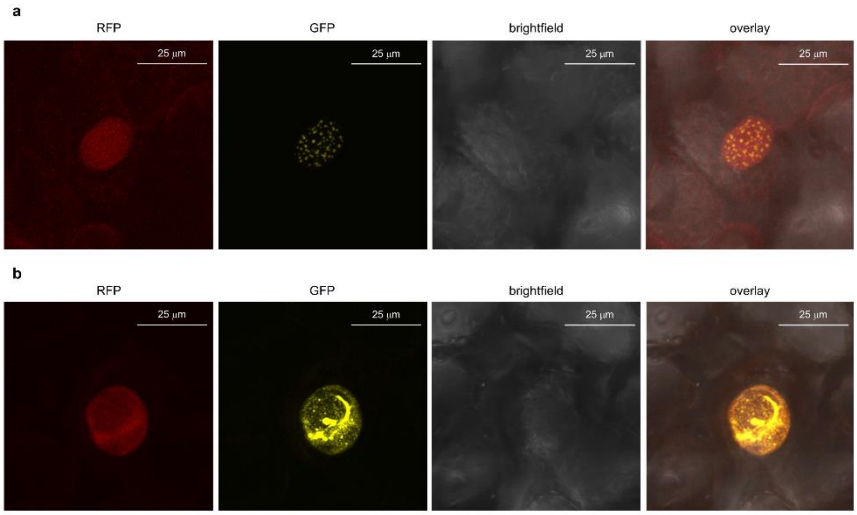
Supplementary Fig. 3. Protein interaction analysis shows the mini-TGA StTGA2.1 can form homodimers *in planta* and *in vitro*. a, Co-immunoprecipitation assay positive controls for GFP, showing the detection of GFP-tagged StTGAs after immunoprecipitation (GFP-IP) in the protein-interaction analysis samples (Fig. 3b). The combination of GFP and HA-tagged proteins expressed in *N. benthamiana* is indicated for each sample (+/−). **b,** Negative controls for GFP, showing the HA-tagged proteins do not interact with GFP alone. **c**, Additional controls, showing individual GFP-tagged StTGAs could be detected in the leaf protein extracts (input) and GFP-IP samples. The HA-tagged StTGA2.1 alone was detected in the input, while none of the individual HA-tagged StTGAs were detected in after GFP-IP. **a–c,** Arrows indicate expected bands. **d,** Size exclusion chromatography elution volume of a purified His_6-tagged StTGA2.1. **e,** The protein standard calibration curve used to calculate its oligomeric state. The StTGA2.1 elution volume (orange dot) in relation to the protein standard (black diamonds) corresponds to an ~63 kDa large protein, which is twice the size of a monomer (~33 kDa). Protein sizes are depicted next to each standard marker. **f,** StTGA2.1 oligomeric state on SDS-PAGE gel, before (monomer) and after crosslinking (crosslinked), showing the formation of homodimers at ~60 kDa and higher order complexes at ~90, ~160 kDa and above. Two protein amounts per treatment were loaded.



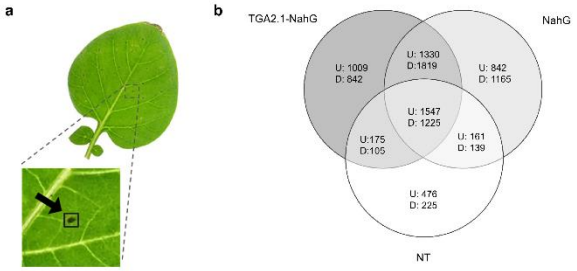
Supplementary Fig. 4. Protein interactions between StTGAs and StNPR cofactors in yeast. StTGA2.1, StTGA2.2, and StTGA2.3 interactions with StNPR1 and StNPR3/4 in the yeast two-hybrid assay. Yeast were co-transformed with bait (BD) and prey (AD) construct combinations and selected on control media without Leu and Trp (-L-W). Positive interactions were determined by yeast growth on selection media without Leu, Trp, His and Ade, with added X-α-galactosidase and Aureobasidin A (-L-W-H-A+Xgal+Aur). The effect of salicylic acid (SA) on interaction strength was determined by yeast growth on the same selection media with added 0.1 mM or 1.0 mM SA.



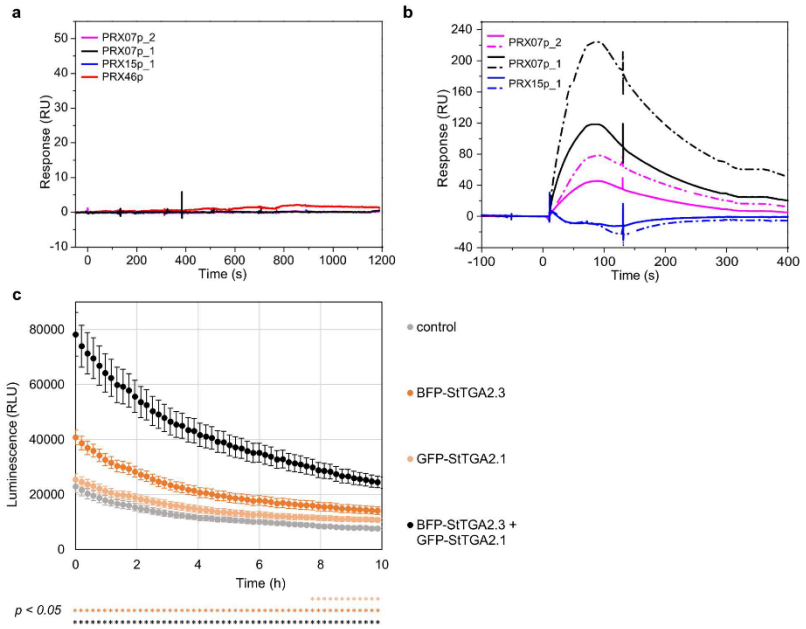
Supplementary Fig. 5. Diverse localization patterns of StTGA2.1. Subcellular localization of GFP-tagged StTGA2.1 (yellow) with H2B-RFP nuclear marker (red) and chloroplast autofluorescence (green) in *N. benthamiana* leaves, showing **a,** StTGA2.1 enrichment around chloroplasts (asterisk) and **b,** localization in the ER (white arrows). **a, b,** Protein fluorescence is represented as the z-stack maximum projection. Orange lines depict the image section close-up. Scale bars, **a,** 50 µm and **b,** 100 µm.



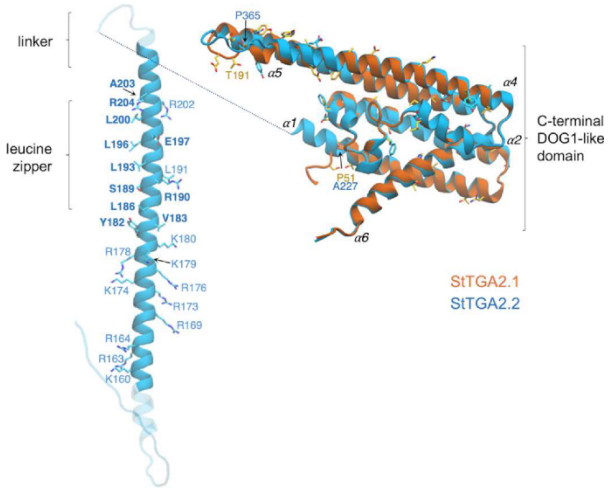
Supplementary Fig. 6. StTGA2.2 and StTGA2.3 subnuclear formations. Subnuclear localization of GFP-tagged **a,** StTGA2.2 and **b,** StTGA2.3 (yellow) with H2B-RFP nuclear marker (red) in *N. benthamiana* leaves. **a, b,** Protein fluorescence is represented as the z-stack maximum projection. Scale bars, 25 µm.



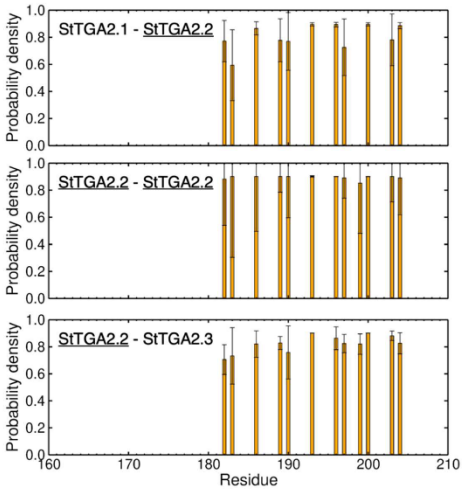
Supplementary Fig. 7. RNA-sequencing sampling procedure and gene expression analysis Venn diagram. a, Lesions with their immediate surrounding area were cut-out from PVY-inoculated leaves at four days post infection and pooled for RNA-sequencing analysis. Mock-inoculated leaf sections of similar size were sampled as controls. **b,** Venn diagram of differentially expressed genes from RNA sequencing between PVY and mock-inoculated TGA2.1-NahG, NahG and NT plants. Genes with adjusted p-value < 0.05 and $|\log_2 FC| \leq -1$ were considered significantly differentially expressed. U, up-regulated genes; D, down-regulated genes.



Supplementary Fig. 8. Interaction between StTGA2.1 or StTGA2.1 premixed with StTGA2.3 and selected TGA-binding sites and transactivation assay repetition. a, Surface plasmon resonance results, showing no interaction between the StTGA2.1 protein and the chip-immobilized PRX15p_1, PRX46p, PRX07p_1 or PRX07p_2 DNA fragments, bound to the chip at ~77, 42, 55, or 60 response units (RU), respectively. Representative sensorgrams are shown. **b,** Interaction analysis between StTGA2.3 alone (solid line) or premixed with StTGA2.1 (dashed line) and the chip-immobilized PRX15p_1, PRX07p_1 or PRX07p_2 DNA fragments, bound to the chip at ~105, 105 or 48 RU, respectively. **c,** Transactivation assay repetition, showing *in planta StPRX07* promoter activation by GFP-tagged StTGA2.1 (light orange), BFP-tagged StTGA2.3 (dark orange) or a combination of both (black). BFP or GFP-tagged controls and their combination (control) were used to detect the basal promoter activity (grey). Average values ± standard error of 17 biological replicates in the first 10 h of measurement are shown. Significance was determined using a two-tailed *t*-test and is shown below the response curve for GFP-tagged StTGA2.1 (light orange), BFP-tagged StTGA2.3 (dark orange) or a combination of both (black) compared with control.



Supplementary Fig. 9. Comparative structural analysis and persistent contacts in dimers of StTGA2.1 and StTGA2.2. Molecular architectures of StTGA2.1 (orange) and StTGA2.2 (blue) proteins. The StTGA2.2 bZIP domain (aa 158-211) is shown. The N-termini of StTGA2.1 (aa 1-45) and StTGA2.2 (aa 1-129) are not shown for visual clarity. The C-terminal part is highly conserved between StTGA2.1 (aa 47-240), StTGA2.2 (aa 222-446), and StTGA2.3 (aa 105-327) (Fig. 5a). Amino acid residues forming persistent contacts in the leucine zipper, according to molecular dynamics simulations, are shown in bold. Basic amino acid residues that may contribute to DNA-binding are depicted and labelled. Non-conservative substitution sites in the putative DOG1 domains are also represented as liquorice.



Supplementary Fig. 10. Probability density of residues of StTGA2.2 forming contacts with a dimer partner. The truncated protein forming a dimer with StTGA2.2 is specified in each plot (i.e., StTGA2.1, StTGA2.2, or StTGA2.3). A maximum distance of 7 Å between C_α atoms in a pair of residues was established. Bars with a standard deviation > 50% of the probability density are considered transient contacts in the simulations and are not included in these plots. In the StTGA2.2-StTGA2.2 and StTGA2.2-StTGA2.3 dimers the main interacting sites in StTGA2.2 are Tyr182, Val183, Leu186, Ser189, Arg190, Leu193, Leu196, Glu197, Leu200, Gln201, Ala203 and Arg204. The same interacting sites are kept in StTGA2.1-StTGA2.2, except for Glu197. In the simulations, StTGA2.1, StTGA2.2 and StTGA2.3, are truncated, keeping the amino acids 1-43, 159-206, and 42-89, respectively.

Supplementary Tables

Supplementary Table 1. A list of identified StTGA orthologues including basic protein information. Protein sequence lengths, molecular weight (M_W) and theoretical pI, calculated with the ProtParam tool (Gasteiger et al., 2005), and StTGA domain prediction based on Prosite (de Castro et al., 2006).

No.	Gene ID	Chromosome	Length (aa)	M_W (kDa)	pI	Domain 1	Domain 2
1	**Sotub01g009430**	1	327	36.34	8.61	bZIP	DOG1
2	**Sotub04g010500**	4	369	41.66	6.20	bZIP	DOG1
3	**Sotub04g022350**	4	369	41.57	5.46	bZIP	DOG1
4	**Sotub04g027470**	4	361	40.88	6.68	bZIP	DOG1
5	**Sotub05g007640**	5	483	53.10	6.51	bZIP	DOG1
6	**Sotub06g031310**	6	503	55.94	6.53	bZIP	DOG1
7	**Sotub10g020240**	10	488	55.53	6.26	bZIP	DOG1
8	**Sotub10g022140**	10	461	52.24	6.47	bZIP	DOG1
9	**Sotub10g022550**	10	438	48.23	7.18	bZIP	DOG1
10	**Sotub10g022560**	10	270	30.49	5.69	-	DOG1
11	**Sotub10g022570**	10	271	30.51	6.00	-	DOG1
12	**Sotub11g020650**	11	324	36.17	8.86	bZIP	DOG1
13	**Sotub11g025820**	11	435	48.60	6.91	bZIP	DOG1
14	**Sotub12g025350**	12	348	39.56	6.92	bZIP	DOG1

Supplementary Table 2. Differential expression of StTGAs in NT and NahG genotypes after viral infection. Microarray data showing StTGA gene expression comparisons between PVY- and mock-inoculated plants at 1, 3 and 6 days post inoculation (dpi), adapted from Baebler *et al.* (2014). Only statistically significant values (FDR adjusted p-value < 0.05) are shown, given as log_2FC. Cell shading based on log_2FC values: blue, down-regulated; orange, up-regulated.

		NT			NahG		
Gene ID	**Microarray ID**	**1 dpi**	**3 dpi**	**6 dpi**	**1 dpi**	**3 dpi**	**6 dpi**
Sotub04g010500	MICRO.14906.C1	-	-	-	1.04	-	-
	MICRO.14906.C2	-	-	-	0.77	-	-0.79
	MICRO.14906.C3	-	-0.93	-0.68	0.49	-0.90	-1.35
	POAED18TP	-	-0.45	-	-	-	-
Sotub04g022350	MICRO.7564.C1	0.96	-	-	-	-	-
Sotub01g009430	MICRO.14316.C1	-	-	-	-	-	-
	MICRO.14867.C1	0.34	0.68	0.42	0.54	-	0.28
	MICRO.8878.C1	-0.62	-	-	-	-	0.53
	POAC439TP	-0.64	-0.39	-	-	-	0.60
Sotub10g022550	MICRO.7441.C1	-1.56	-1.03	0.82	-0.80	1.05	2.63
	MICRO.7474.C3	-0.95	-1.06	-	-	-	-
Sotub10g022560	bf_arrayxxx_0075b06.t7m.scf	-	-	-	-	2.11	1.98
	MICRO.7474.C2	-0.41	-	-	0.34	-	0.80
	MICRO.7474.C5	-	-	-	-	0.44	0.87
	STMHL35TV	-0.90	-	-	-	0.70	1.72
Sotub10g022570	MICRO.7474.C6	1.14	1.23	-	0.68	-	-
Sotub04g027470	bf_arrayxxx_0086h05.t7m.scf	0.82	-	-	-	-	-
	MICRO.11212.C1	-	-	-0.41	-	-0.46	-0.66
	MICRO.9918.C1	-0.37	-0.46	-	-	-0.48	-0.47
	STMGM64TH	-0.91	-1.53	-	-1.00	-	0.44
Sotub06g031310	bf_suspxxxx_0045g08.t7m.scf	-	-	-	-	-	0.71
Sotub10g020240	bf_ivrootxx_0054a07.t3m.scf	0.49	-0.59	-	-	-	-
Sotub11g025820	MICRO.14778.C1	0.67	0.78	1.15	0.71	1.23	2.22
	MICRO.3439.C1	-0.99	-1.29	-0.57	-0.99	-0.41	0.35
Sotub05g007640	MICRO.15058.C1	-	-	-	-	-	-0.74
	POAD811TP	-	-	-	-	-	-
	POAD811TV	-	-	-	-	-	-

Supplementary Table 3. Differential gene expression in salicylic acid-deficient transgenic plants overexpressing StTGA2.1 after viral infection (available as an Excel file). RNA sequencing results table showing gene expression comparisons between dexamethasone (DEX)-treated mock-inoculated plants (TGA2.1-NahG vs. NahG and TGA2.1-NahG vs. NT) and comparisons of DEX-treated PVY- vs mock-inoculated plants of all three genotypes, given as log_2FC and adjusted p-value. Additionally, comparisons between DEX- and control-treatments (DEX vs. control) are shown for PVY and mock-inoculated TGA2.1-NahG plants, as controls. For each gene, the ITAG or PGSC identifier, description and a MapMan ontology functional gene group (BIN) (Ramšak et al., 2014) with BIN number, is listed. Only the first BIN mapping is shown. FDR adjusted p-values below 0.05 are depicted in red. TMM normalized reads counts and raw read counts for each of the three biological replicates per genotype per treatment are shown. Cell shading based on log_2FC values: blue, down-regulated; orange, up-regulated.

Supplementary Table 4. Technical validation of RNA sequencing results with qPCR. Comparison of five differentially expressed genes obtained from RNA sequencing results (RNA-Seq) with qPCR analysis for TGA2.1-NahG, NahG and NT plants 4 days after viral infection (PVY- vs. mock-inoculated plants comparison), given as log_2FC. Gene expression ratios with FDR adjusted p-value < 0.05 are underlined. Cell shading based on log_2FC values: blue, down-regulated; orange, up-regulated.

		TGA2.1-NahG		NahG		NT	
Gene Name	**Gene ID**	**RNA-Seq**	**qPCR**	**RNA-Seq**	**qPCR**	**RNA-Seq**	**qPCR**
StACX3	Sotub10g008540	2.05	2.07	2.21	2.11	1.46	1.50
StCS	Sotub01g027350	1.55	1.17	1.49	1.80	0.52	0.67
StPti5	Sotub02g020180	4.37	3.85	5.36	8.25	2.14	5.33
StPRX28	Sotub01g042120 Sotub02g011690	7.47	6.29	6.91	5.69	2.89	2.48
StTGA2.1	Sotub10g022560	-1.45	-1.54	0.21	0.95	-0.26	-0.37

Supplementary Table 5. Enrichment of differentially regulated genes in salicylic acid-deficient transgenic plants overexpressing StTGA2.1 after viral infection (available as an Excel file). Gene Set Enrichment Analysis results table showing MapMan ontology functional gene groups (BINs) (Ramšak et al., 2014) enriched in up-regulated or down-regulated genes in TGA2.1-NahG, NahG or NT plants after PVY infection. Functional groups, significantly enriched (FDR corrected q-value < 0.05) in at least one of the three genotypes, are listed. (+), enriched in up-regulated genes; (−), enriched in down-regulated genes.

Supplementary Table 6. Classification of selected potato peroxidases. Genes from the MapMan peroxidase functional group (BIN 26.12) (Ramšak et al., 2014) were named and classified according to BLAST results in the RedoxiBase database (Savelli et al., 2019). All hits with E-value equal to zero or best hits with E-value above zero (*) are listed for each gene. The presence of secretory signal peptides was predicted with SignalP 5.0 (Almagro Armenteros et al., 2019). (+), signal peptide; (−), no signal peptide predicted; (+/−) signal peptide predicted in at least one of several sequences assigned to the same gene ID; bold, the three peroxidases included in target confirmation.

Gene ID	BLAST Hits	Classification	Secretory signal Peptide
PGSC0003DMG400020437	StPRX05, StPRX52, StPRX72	class III peroxidase	–
PGSC0003DMG400025492	StPRX72, StPRX52	class III peroxidase	+
PGSC0003DMG400025621	StPRX20, StPRX09, StPRX28, StPRX50	class III peroxidase	–
PGSC0003DMG400026575	StPRX19, StPRX02, StPRX15	class III peroxidase	+/–
PGSC0003DMG400032147	StPRX12, StPRX10	class III peroxidase	+
Sotub01g006690	StPRX20, StPRX09	class III peroxidase	+
Sotub01g042120	StPRX28, StPRX50	class III peroxidase	+
Sotub02g022480	StPRX72, StPRX52	class III peroxidase	+
Sotub02g022490	StPRX72, StPRX52	class III peroxidase	+
Sotub02g023700	StPRX05	class III peroxidase	+
Sotub02g027910	StPRX12, StPRX10	class III peroxidase	–
Sotub02g031070	StPRX19, StPRX02	class III peroxidase	+
Sotub02g035680	**StPRX15**	**class III peroxidase**	**+**
Sotub02g037370	StPRX19*	class III peroxidase	+
Sotub03g007840	**StPRX46**	**class III peroxidase**	**+**
Sotub03g010400	StPRX19, StPRX02	class III peroxidase	+
Sotub03g015450	StPRX24	class III peroxidase	+
Sotub04g026530	StPRX14	class III peroxidase	+
Sotub04g026540	StPRX11, StPRX17	class III peroxidase	+
Sotub04g026550	StPRX11, StPRX17	class III peroxidase	+
Sotub05g006810	StPRX75	class III peroxidase	–
Sotub05g024920	StPRX47, StPRX16	class III peroxidase	+
Sotub06g009790	StPRX47, StPRX16	class III peroxidase	+
Sotub06g032420	StPRX71	class III peroxidase	+
Sotub06g033210	StPRX25	class III peroxidase	+
Sotub08g006050	StPRX27	class III peroxidase	+
Sotub08g011590	StAPx-R	ascorbate peroxidase related	–
Sotub08g017100	StPRX40	class III peroxidase	+
Sotub09g007100	StPRX03, StPRX04	class III peroxidase	+
Sotub09g020950	**StPRX07**	**class III peroxidase**	**+**
Sotub10g024760	StPRX03	class III peroxidase	+
Sotub11g009570	StPRX58	class III peroxidase	+
Sotub11g029720	StPRX51*	class III peroxidase	+

Supplementary Table 7. Biological validation of RNA sequencing results with qPCR. Relative expression of three class III peroxidase genes as obtained by RNA sequencing (RNA-Seq) and qPCR analyses for both TGA2.1-NahG lines, line 7 (L7) and line 12 (L12), NahG and NT plants 4 days after viral infection (PVY- vs. mock-inoculated plants comparison), given as log_2FC. Gene expression ratios with FDR adjusted p-value < 0.05 are underlined. Cell shading based on log_2FC values: blue, down-regulated; orange, up-regulated.

		TGA2.1-NahG			NahG		NT	
Gene Name	Gene ID	RNA-Seq	qPCR L7	qPCR L12	RNA-Seq	qPCR	RNA-Seq	qPCR
StPRX07	Sotub09g020950	10.78	6.87	3.86	7.13	3.54	5.19	3.43
StPRX15	Sotub02g035680	1.90	1.14	1.59	-0.15	0.20	4.27	3.42
StPRX46	Sotub03g007840	6.79	3.92	3.24	2.88	2.21	4.23	1.84
StTGA2.1	Sotub10g022560	-1.45	-3.67	0.10	0.21	-0.07	-0.26	-0.50

Supplementary Table 8. Statistics of interactions between the StTGA2.2 homodimer and DNA in molecular dynamics simulations. Total frequency of hydrogen bond, salt bridge interactions, and hydrophobic contacts (*) between the protomers of the StTGA2.2 homodimer and the two DNA strands (DNA1 and DNA2) in five independent molecular dynamics simulations of 200 ns. For hydrophobic contacts, a maximum distance of 4 Å between C_β atoms in the protein and the methyl group in a thymine was established.

StTGA2.2	DNA1	Frequency / %					StTGA2.2	DNA2	Frequency / %				
K160	A-8	19	6	8	-	25	K160	-	-	-	-	-	-
R163	T-7	90	-	82	63	59	R163	T-7	33	34	-	20	41
R163	A3	-	28	20	-	-	R163	-	-	-	-	-	-
R164	A-8	27	28	1	19	-	R164	T-7	36	18	19	10	1
Q167	T-7	11	-	22	20	17	Q167	T-6	11	5	14	8	10
N168	A3	4	15	9	7	7	N168	C3	30	29	19	22	16
N168	T-4	26	16	24	23	26	-	-	-	-	-	-	-
R169	T2	22	22	30	14	19	R169	T2	19	13	20	24	24
A171*	T-4	18	20	38	21	24	A171*	T-5	21	34	38	33	44
A172*	T2	43	24	21	21	43	A172*	T2	37	38	36	32	34
R173	C-1	22	65	17	32	24	R173	G1	39	9	9	12	20
R173	G1	12	35	16	17	12	R173	C-1	-	6	7	11	12
S175	T-4	78	80	54	59	63	S175	C-4	60	59	58	49	57
R176	G1	26	38	31	41	39	R176	G1	13	34	20	17	13
R176	-	-	-	-	-	-	R176	C-1	22	24	22	24	27
R178	T-4	37	29	33	33	22	R178	C-4	59	45	26	41	35
R178	G-5	18	-	58	17	24	R178	T-5	13	14	28	24	22
K179	G-3	18	7	31	17	10	K179	T-3	30	29	33	35	38
K180	A-2	22	38	30	28	31	K180	A-2	37	48	30	33	37

Supplementary Table 9. Statistics of interactions between the StTGA2.1 in heterodimer and DNA in molecular dynamics simulations. Total frequency of hydrogen bond, salt bridge interactions, and hydrophobic contacts (*) between StTGA2.1 in the StTGA2.1-StTGA2.2 heterodimer and DNA in five independent molecular dynamics simulations of 200 ns. For hydrophobic contacts, a maximum distance of 4 Å between C_β atoms in the protein and the methyl group in a thymine was established.

StTGA2.1	DNA	Frequency / %				
R7	G1	-	10	-	22	-
R7	T2	22	13	22	17	25
R11	G1	31	12	20	27	12
R11	C-1	13	15	39	13	12
R20	T-5	27	5	22	15	0
R20	C-4	28	24	33	37	32
R24	T-3	28	22	15	24	22
A13*	T-3	12	32	6	6	27

Supplementary Table 10. Molecular dynamics simulations-based protocol for structure refinement of free and DNA-bound TGA dimers. In the equilibration phases, there is a gradual change in temperature, time step, and the position restraint potentials. Atom velocities are re-distributed using five different seed numbers to initiate Equilibration_6. With that, five independent trajectories are generated for conformational sampling. Applied potentials are either simple harmonic restraints (H) or flat-bottom potentials (FB). During and after Equilibration_5, the position restraint potentials are applied only to the residues forming the leucine heptads. "Backbone" refers to both the backbone atoms of TGA proteins and DNA.

Simulation Phases	Position Restraints				Time Step	Total Time	Temperature
	Backbone	Type	Side Chain	Type			
Equilibration_1	1.000	H	0.100	H	1 fs	2 ns	100.15 K
Equilibration_2	1.000	H	0.050	H	2 fs	1 ns	200.15 K
Equilibration_3	1.000	H	0.025	H	2 fs	1 ns	250.15 K
Equilibration_4	1.000	H	0.000	H	2 fs	1 ns	300.15 K
Equilibration_5	0.500*	H	0.000	-	2 fs	4 ns	320.15 K
Equilibration_6	0.250*	H	0.000	-	2 fs	5 x 2 ns	340.15 K
Equilibration_7	0.125*	H	0.000	-	4 fs	5 x 12 ns	360.15 K
Equilibration_8	0.050*	H	0.000	-	4 fs	5 x 128 ns	340.15 K
Equilibration_9	0.025*	H	0.000	-	4 fs	5 x 64 ns	320.15 K
Equilibration_10	0.250*	FB	0.000	-	4 fs	5 x 100 ns	298.15 K
Sampling	0.000	-	0.000	-	4 fs	5 x 128 ns	298.15 K

*Position restraints applied to Cα only.

Supplementary Table 11. Primers used for cloning and sequencing (available as an Excel file).

Supplementary Table 12. Primers and probes used for qPCR analysis. F, forward; R, reverse; P, probe.

Gene Name	Gene Description	Gene ID	Primer/Probe Sequence (5'-3')	Amplicon Efficiency
StACX3	acyl-CoA oxidase	Sotub10g008540	Baebler *et al.* (2014)	
StCS	citrate synthase	Sotub01g027350		
StPRX28*	class III peroxidase	Sotub01g042120, Sotub02g011690		
StPti5	ethylene responsive factor	Sotub02g020180	F: GCAAGAAATACAGAGGCGTACGA R: CATACTCTCGCACCGTGTCTAG P: FAM-TTCGGCAGCGTATTTT-NFQ	93%
StTGA2.1	TGA transcription factor	Sotub10g022560	F: ATCTCTTGCTACTGAAGGGTCAT R: CCATAGCCATTGCCATCTGACTTAT P: FAM-CCGGAGATGCAGCTTATAAT-NFQ	90%
StPRX07	class III peroxidase	Sotub09g020950 (CLCdnDe14_116472, CLCdnRY2_25372, CLCdnRY3_29096, VdnDe5_126154)**	F: GTTGGAATCGCACATTGTAATTTT R: AGCTGCTAATGTAGGATCCATAGTCTT P: FAM-TCCAAGATAGGCTATCGCCCGTACCTG-Zen Iowa Black TM FQ	92%
StPRX15	class III peroxidase	Sotub02g035680 (CLCdnDe10_14846, CLCdnRY1_7366, CLCdnRY10_8871, VdnDe1_213633, VdnDe1_213635, VdnDe5_46588)**	F: CACTACAAACAAACAGATCACTTCACCTA R: GACGAAGCAGTCGTGGAAGAA P: FAM-TACCGCCGCCGCCACCCT-Zen Iowa Black TM FQ	80%
StPRX46	class III peroxidase	Sotub03g007840 (CLCdnDe11_47734, CLCdnRY1_4154, PBdnRY1_4594, VdnDe1_299904, VdnDe1_299906, VdnDe5_44388)**	F: CGGCGGTTGGGTCAGTT R: GAATGAGAGATGCTCCCATTCG P: FAM-TGGCGGAAGCTATTGCGAGGGA-Zen Iowa Black TM FQ	91%
StCOX1	cytochrome oxidase	Sotub04g015050	Weller *et al.* (2000)	
StEF-1	elongation factor	Sotub06g010680	Baebler *et al.* (2009)	
PVY Univ	PVY coat protein	AJ390390	Kogovšek *et al.* (2008)	

* Previously named POX (Baebler et al., 2014).

** For qPCR assay design we used sequences obtained from cultivar Rywal and cultivar Désirée reference transcriptome(Petek et al., 2020) as template. Only transcripts, targeted by the assay, are listed.

Supplementary Table 13. Primers used for targeted genome sequencing.

Primer Pair	Primer Name	Sequence (5'-3')
A	WO100007-NIB-1_7F	GAGGCAAAATACGTGAAGCTAAAAG
	WO100007-NIB-1_7R	AGAGCACGGGCTGATTGTCT
B	WO100007-1_3F	GCTACGCCTTTGCAGCATAT
	WO100007-1_3R	ACAGGAAAAGCAACTCTGGCT

Supplementary Table 14. Complementary primers used for preparation of promoter DNA fragments.

Gene Name and ID	Promoter Fragment	Primer Name	Sequence (5'-3')
StPRX07 Sotub09g020950	PRX07p_1	per950p_1F	GTTACTACTCGAGCGTGTGCCCAACGTCACAATCC
		per950p_1R	GGATTGTGACGTTGGGCACA
	PRX07p_2	per950p_2F	GTTACTACTCGAGCGTTAGGGGTGACGTTTCCAAT
		per950p_2R	ATTGGAAACGTCACCCCTAA
StPRX15 Sotub02g035680	PRX15p_1	per680p_1F	GTTACTACTCGAGCGTTTAATAATGATGACATTTG
		per680p_1R	CAAATGTCATCATTATTAAA
StPRX46 Sotub03g007840	PRX46p	per840p_1F	GTTACTACTCGAGCGTGAACCTGAGGTCAACCGTT
		per840p_1R	AACGGTTGACCTCAGGTTCA

Supplementary References

Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol **37**: 420–423

Aoyama T, Chua NH (1997) A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J **11**: 605–612

Baebler Š, Krečič-Stres H, Rotter A, Kogovšek P, Cankar K, Kok EJ, Gruden K, Kovač M, Žel J, Pompe-Novak M, et al (2009) PVYNTN elicits a diverse gene expression response in different potato genotypes in the first 12 h after inoculation. Mol Plant Pathol **10**: 263–275

Baebler Š, Witek K, Petek M, Stare K, Tušek-Žnidarič M, Pompe-Novak M, Renaut J, Szajko K, Strzelczyk-Zyta D, Marczewski W, et al (2014) Salicylic acid is an indispensable component of the Ny-1 resistance-gene-mediated response against Potato virus Y infection in potato. J Exp Bot **65**: 1095–1109

de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006) ScanProsite: Detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res **34**: 362–365

Dudley QM, Cai YM, Kallam K, Debreyne H, Carrasco Lopez JA, Patron NJ (2021) Biofoundry-assisted expression and characterization of plant proteins. Synth Biol **6**: ysab029

Eschenfeldt WH, Stols L, Sanville Millard C, Joachimiak A, Donnelly MI (2009) A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. Methods Mol Biol **498**: 105–115

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. *In* JM Walker, ed, Proteomics Protoc. Handb., 1st ed. Humana Press, pp 571–607

Kogovšek P, Gow L, Pompe-Novak M, Gruden K, Foster GD, Boonham N, Ravnikar M (2008) Single-step RT real-time PCR for sensitive detection and discrimination of Potato virus Y isolates. J Virol Methods **149**: 1–11

Petek M, Zagorščak M, Ramšak Ž, Sanders S, Tomaž Š, Tseng E, Zouine M, Coll A, Gruden K (2020) Cultivar-specific transcriptome and pan-transcriptome reconstruction of tetraploid potato. Sci Data **7**: 249

Pham GM, Hamilton JP, Wood JC, Burke JT, Zhao H, Vaillancourt B, Ou S, Jiang J, Robin Buell C (2020) Construction of a chromosome-scale long-read reference genome assembly for potato. Gigascience **9**: giaa100

Ramšak Ž, Baebler Š, Rotter A, Korbar M, Mozetič I, Usadel B, Gruden K (2014) GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology. Nucleic Acids Res **42**: 1167–1175

Savelli B, Li Q, Webber M, Jemmat AM, Robitaille A, Zamocky M, Mathé C, Dunand C (2019) RedoxiBase: A database for ROS homeostasis regulated proteins. Redox Biol **26**: 101247

Stark JC, Huang A, Nguyen PQ, Dubner RS, Hsu KJ, Ferrante TC, Anderson M, Kanapskyte A, Mucha Q, Packett JS, et al (2018) BioBits^{TM} Bright: A fluorescent synthetic biology education kit. Sci Adv **4**: eaat5107

Weller SA, Elphinstone JG, Smith NC, Boonham N, Stead DE (2000) Detection of Ralstonia solanacearum strains with a quantitative, multiplex, real-time, fluorogenic PCR (TaqMan) assay. Appl Environ Microbiol **66**: 2853–2858

Supplementary Table 3 is available at http://projects.nib.si/sensors/results/. Accessed: 13-October-2022.

Supplementary Table 5. Enrichment of differentially regulated genes in salicylic acid-deficient transgenic plants overexpressing StTGA2.1 after viral infection. Gene Set Enrichment Analysis results table showing MapMan ontology functional gene groups (BINs) (Ramšak et al., 2014) enriched in up-regulated or down-regulated genes in TGA2.1-NahG, NahG or NT plants after PVY infection. Functional groups, significantly enriched (FDR corrected q-value < 0.05) in at least one of the three genotypes, are listed. (+), enriched in up-regulated genes; (−), enriched in down-regulated genes.

			PVY vs. mock		
BIN	Functional group	Size	TGA2.1-NahG	NahG	NT
1	PS	426	−	−	−
1.1	PS.LIGHTREACTION	330	−	−	−
1.1.1	PS.LIGHTREACTION.PHOTOSYSTEM II	141	−	−	−
1.1.1.1	PS.LIGHTREACTION.PHOTOSYSTEM II.LHC-II	32	−	−	−
1.1.1.2	PS.LIGHTREACTION.PHOTOSYSTEM II.PSII POLYPEPTIDE SUBUNITS	108	−	−	−
1.1.2	PS.LIGHTREACTION.PHOTOSYSTEM I	45	−	−	−
1.1.2.2	PS.LIGHTREACTION.PHOTOSYSTEM I.PSI POLYPEPTIDE SUBUNITS	32	−	−	−
1.1.4	PS.LIGHTREACTION.ATP SYNTHASE	30		−	
1.1.6	PS.LIGHTREACTION.NADH DH	42	−	−	
1.3	PS.CALVIN CYCLE	56	−	−	−
2.2.1.3	MAJOR CHO METABOLISM.DEGRADATION.SUCROSE.INVERTASES	29			+
2.2.2	MAJOR CHO METABOLISM.DEGRADATION.STARCH	37		−	
8.1	TCA /ORGANIC TRANSFORMATION.TCA	52		+	
9.1	MITOCHONDRIAL ELECTRON TRANSPORT / ATP SYNTHESIS.NADH-DH (TYPE I)	74		+	+
9.1.1	MITOCHONDRIAL ELECTRON TRANSPORT / ATP SYNTHESIS.NADH-DH (TYPE I).COMPLEX I	21	+		
10.5.1	CELL WALL.CELL WALL PROTEINS.AGPS	37			+
10.5.1.1	CELL WALL.CELL WALL PROTEINS.AGPS.AGP	30			+
10.6	CELL WALL.DEGRADATION	243			+
10.6.2	CELL WALL.DEGRADATION.MANNAN-XYLOSE-ARABINOSE-FUCOSE	74			+
11.2.4	LIPID METABOLISM.FA DESATURATION.OMEGA 6 DESATURASE	28	+		
11.6	LIPID METABOLISM.LIPID TRANSFER PROTEINS ETC	65		−	−
13.2.7	AMINO ACID METABOLISM.DEGRADATION.HISTIDINE	33	+		
16.1.2	SECONDARY METABOLISM.ISOPRENOIDS.MEVALONATE PATHWAY	47	+		
16.1.5	SECONDARY METABOLISM.ISOPRENOIDS.TERPENOIDS	155	+		
16.10	SECONDARY METABOLISM.SIMPLE PHENOLS	38			+
16.5.99	SECONDARY METABOLISM.SULFUR-CONTAINING.MISC	24		−	−
16.5.99.1	SECONDARY METABOLISM.SULFUR-CONTAINING.MISC.ALLIINASE	24		−	−
17.1.3	HORMONE METABOLISM.ABSCISIC ACID.INDUCED-REGULATED-RESPONSIVE-ACTIVATED	32	−	−	
17.2	HORMONE METABOLISM.AUXIN	351	−	−	
17.2.3	HORMONE METABOLISM.AUXIN.INDUCED-REGULATED-RESPONSIVE-ACTIVATED	287	−	−	
17.4.1	HORMONE METABOLISM.CYTOKININ SYNTHESIS-DEGRADATION	85		−	−
17.7	HORMONE METABOLISM.JASMONATE	64		+	+
17.8	HORMONE METABOLISM.SALICYLIC ACID	22			+
20.1.7	STRESS.BIOTIC.PR-PROTEINS	208	+	+	+
20.1.7.1	STRESS.BIOTIC.PR-PROTEINS.PR1 (ANTIFUNGAL)	33	+		
20.1.7.3	STRESS.BIOTIC.PR-PROTEINS.PR3/4/8/11 (CHITINASES AND CHITIN BINDING PROTEINS)	44	+	+	+
26.12	MISC.PEROXIDASES	139	+		
26.19	MISC.PLASTOCYANIN-LIKE	44	+		
26.21	MISC.PROTEASE INHIBITOR/SEED STORAGE/LIPID TRANSFER PROTEIN (LTP) FAMILY PROTEIN	117		−	
26.25	MISC.SULFOTRANSFERASE	49			−
26.28	MISC.GDSL-MOTIF LIPASE	134		−	
26.3.2	MISC.GLUCO-, GALACTO- AND MANNOSIDASES.BETA-GALACTOSIDASE	22		−	
26.9	MISC.GLUTATHIONE S TRANSFERASES	90	+	+	+
27.3.32	RNA.REGULATION OF TRANSCRIPTION.WRKY DOMAIN TRANSCRIPTION FACTOR FAMILY	97		+	+
27.3.63	RNA.REGULATION OF TRANSCRIPTION.PHD FINGER TRANSCRIPTION FACTOR	49	+		
27.3.64	RNA.REGULATION OF TRANSCRIPTION.PHOR1	21	+		
27.3.8	RNA.REGULATION OF TRANSCRIPTION.C2C2(ZN) DOF ZINC FINGER FAMILY	41	−		
28.1.3	DNA.SYNTHESIS/CHROMATIN STRUCTURE.HISTONE	83	+	+	+
28.1.3.2	DNA.SYNTHESIS/CHROMATIN STRUCTURE.HISTONE.CORE	76	+	+	+
28.1.3.2.1	DNA.SYNTHESIS/CHROMATIN STRUCTURE.HISTONE.CORE.H2A	26	+	+	
28.1.3.2.3	DNA.SYNTHESIS/CHROMATIN STRUCTURE.HISTONE.CORE.H3	20	+	+	+
29.2.1.1.1	PROTEIN.SYNTHESIS.RIBOSOMAL PROTEIN.PROKARYOTIC.CHLOROPLAST	103		−	−
29.2.1.1.1.1	PROTEIN.SYNTHESIS.RIBOSOMAL PROTEIN.PROKARYOTIC.CHLOROPLAST.30S SUBUNIT	35		−	−
29.2.1.1.1.2	PROTEIN.SYNTHESIS.RIBOSOMAL PROTEIN.PROKARYOTIC.CHLOROPLAST.50S SUBUNIT	68		−	−
29.2.1.2	PROTEIN.SYNTHESIS.RIBOSOMAL PROTEIN.EUKARYOTIC	324	+	+	
29.2.1.2.1	PROTEIN.SYNTHESIS.RIBOSOMAL PROTEIN.EUKARYOTIC.40S SUBUNIT	123	+	+	
29.2.1.2.2	PROTEIN.SYNTHESIS.RIBOSOMAL PROTEIN.EUKARYOTIC.60S SUBUNIT	200	+	+	
29.5.11.20	PROTEIN.DEGRADATION.UBIQUITIN.PROTEASOM	73	+	+	
29.8	PROTEIN.ASSEMBLY AND COFACTOR LIGATION	75	−	−	
30.1	SIGNALLING.IN SUGAR AND NUTRIENT PHYSIOLOGY	66	+	+	+
30.1.1	SIGNALLING.IN SUGAR AND NUTRIENT PHYSIOLOGY.MISC	63	+	+	+
30.2.8.1	SIGNALLING.RECEPTOR KINASES.LEUCINE RICH REPEAT VIII.TYPE 1	20			−
30.2.16	SIGNALLING.RECEPTOR KINASES.CATHARANTHUS ROSEUS-LIKE RLK1	79			+
30.2.17	SIGNALLING.RECEPTOR KINASES.DUF 26	94			+
30.2.19	SIGNALLING.RECEPTOR KINASES.LEGUME-LECTIN	38	+		+
30.2.99	SIGNALLING.RECEPTOR KINASES.MISC	206	+		+
34.10	TRANSPORT.NUCLEOTIDES	47	−		−
34.13	TRANSPORT.PEPTIDES AND OLIGOPEPTIDES	123		−	
35.1.5	NOT ASSIGNED.NO ONTOLOGY.PENTATRICOPEPTIDE (PPR) REPEAT-CONTAINING PROTEIN	416	+		−
35.1.27	NOT ASSIGNED.NO ONTOLOGY.TETRATRICOPEPTIDE REPEAT (TPR)	289			−

Supplementary Table 11. Primers used for cloning and sequencing.

Target Type	Target Name	Target ID/Reference	Application	Assay	Primer Name	Sequence (5'-3')
coding sequence	StNPR1	Sotub07g011600	confirmation/sequencing	na	StNPR1_middle_Fw	GTCTGAATTCCCCCTGGCTAG
coding sequence	StNPR1	Sotub07g011600	confirmation/sequencing	na	StNPR1_middle_Rv	ACAAATAGGCGAGCACACTGA
coding sequence	StNPR3/4	Sotub02g015550	confirmation/sequencing	na	StNPR3/4a_middle_Fw	GCAAGTGCTGAGACAACAGC
coding sequence	StNPR3/4	Sotub02g015550	confirmation/sequencing	na	StNPR3/4a_middle_Rv	ACAAATAGCTTAAGAAAGTACGGAAAG
coding sequence	StTGA2.2	Sotub10g022550	confirmation/sequencing	na	StTGA2.2_middle_Fw	GCTTCTAGTCAATCAGTTGGAGC
coding sequence	StTGA2.2	Sotub10g022550	confirmation/sequencing	na	StTGA2.2_middle_Rv	TTTGATTTCTATCATCTCCATCCATGT
coding sequence	StTGA2.3	Sotub01g009430	confirmation/sequencing	na	StTGA2.3_middle_Fw	TCTCAGGGAATCGAATCACTGC
coding sequence	StTGA2.3	Sotub01g009430	confirmation/sequencing	na	StTGA2.3_middle_Rv	TGATTGATCTCCTGAACTTGAAACAT
coding sequence	StTGA2.3	Sotub01g009430	digestion-ligation cloning	cell-free protein synthesis	MPN011_TGA2.3_F	AGAGGTCTCtAATGGCAGACGTCAGTGCTATG
coding sequence	StTGA2.3	Sotub01g009430	digestion-ligation cloning	cell-free protein synthesis	MPN012_TGA2.3_mid1_R	AGAGGTCTCtTGATCTCTGGTTTTGTCACTTGCATCG
coding sequence	StTGA2.3	Sotub01g009430	digestion-ligation cloning	cell-free protein synthesis	MPN013_TGA2.3_mid1_F	AGAGGTCTCtATCAAAAGACGCTTCGTAGGCTTTCC
coding sequence	StTGA2.3	Sotub01g009430	digestion-ligation cloning	cell-free protein synthesis	MPN014_TGA2.3_mid2_R	AGAGGTCTCtGAGTCTCAGCCAATGACTGTTGCAG
coding sequence	StTGA2.3	Sotub01g009430	digestion-ligation cloning	cell-free protein synthesis	MPN015_TGA2.3_mid2_F	AGAGGTCTCtACTCTGGCAGGCTCTCTTGGACC
coding sequence	StTGA2.3	Sotub01g009430	digestion-ligation cloning	cell-free protein synthesis	MPN016_TGA2.3_R	AGAGGTCTCtCGAACCCTCTCGGGGACGGGCAAGC
coding sequence	mTagBFP2	Stark *et al.* (2018)	fusion cloning	transactivation assay	C_mTAGBFP2_linker2	GGTGGTGGCGGTGGAGGTATGGTGTCTAAGGGCGAAGAGCTGATT
coding sequence	mTagBFP2	Stark *et al.* (2018)	fusion cloning	transactivation assay	D_mTAGBFP2	ATTAAGCTTGTGCCCCAGTTTGCTAGG
coding sequence	StTGA2.3	Sotub01g009430	fusion cloning	transactivation assay	A_TGA2.3	CACCATGGCAGACGTCAGTGCTATGACTGAT
coding sequence	StTGA2.3	Sotub01g009430	fusion cloning	transactivation assay	B_TGA2.3_linker2	TCCACCGCCACCACCCTCTCGGGGACGGGCAAGCCA
coding sequence	StTGA2.1	Sotub10g022560	ligation independent cloning	protein production in *E. coli*	TGA560_LIC_F	TACTTCCAATCCAATGCCGCAGCAATGGATGCAATGGGTGATAGGGA
coding sequence	StTGA2.1	Sotub10g022560	ligation independent cloning	protein production in *E. coli*	TGA550,560_LICR	TTATCCACTTCCAATGTTATTATTGCTCTCGTGGTCTGGC
coding sequence	StTGA2.1	Sotub10g022560	pENTR cloning	co-immunoprecipitation, localization, transactivation assay	Sotub10g022560_pENTR_Fw	CACCATGGATGCAATGGGTGATAGGGA
coding sequence	StTGA2.1	Sotub10g022560	pENTR cloning	co-immunoprecipitation, localization, transactivation assay	Sotub10g022550_560_no_stop_Rv	TTGCTCTCGTGGTCTGGCA
coding sequence	StTGA2.2	Sotub10g022550	pENTR cloning	co-immunoprecipitation, localization, transactivation assay	Sotub10g022550_pENTR_Fw	CACCATGCCCAGTTTCATTTCTCAGATTC
coding sequence	StTGA2.2	Sotub10g022550	pENTR cloning	co-immunoprecipitation, localization, transactivation assay	Sotub10g022550_560_no_stop_Rv	TTGCTCTCGTGGTCTGGCA
coding sequence	StTGA2.3	Sotub01g009430	pENTR cloning	co-immunoprecipitation, localization, transactivation assay	Sotub01g009430_pENTR_Fw	CACCATGGCAGACGTCAGTGCTATG
coding sequence	StTGA2.3	Sotub01g009430	pENTR cloning	co-immunoprecipitation, localization, transactivation assay	Sotub01g009430_no_stop_Rv	CTCTCGGGGACGGGCAAG
coding sequence	StNPR1	Sotub07g011600	pJET cloning	sequence analysis, further cloning	StNPR1_F	ATGGATAGTAGGACTGCTTTTTCG
coding sequence	StNPR1	Sotub07g011600	pJET cloning	sequence analysis, further cloning	StNPR1_R	CTATTTCCTAAAAGGGAGATTATTGGGC
coding sequence	StNPR3/4	Sotub02g015550	pJET cloning	sequence analysis, further cloning	StNPR3/4a_F	ATGGGAAGTTCTGCTGAACC
coding sequence	StNPR3/4	Sotub02g015550	pJET cloning	sequence analysis, further cloning	StNPR3/4a_R	TCATAGGTTCCTAGCTTTGACACTTG
coding sequence	StTGA2.1	Sotub10g022560	pJET cloning	sequence analysis, further cloning	tga2_ Sotub560_F	ATGGATGCAATGGGTGATAGG
coding sequence	StTGA2.1	Sotub10g022560	pJET cloning	sequence analysis, further cloning	tga2_ Sotub550_560_R	TTATTGCTCTCGTGGTCTGG
coding sequence	StTGA2.2	Sotub10g022550	pJET cloning	sequence analysis, further cloning	tga2_Sotub550_F	ATGCCCAGTTTCATTTCTCAGA
coding sequence	StTGA2.2	Sotub10g022550	pJET cloning	sequence analysis, further cloning	tga2_ Sotub550_F1	ATGGCAGATTCTGGTTCTCG
coding sequence	StTGA2.2	Sotub10g022550	pJET cloning	sequence analysis, further cloning	tga2_ Sotub550_560_R	TTATTGCTCTCGTGGTCTGG
coding sequence	StTGA2.3	Sotub01g009430	pJET cloning	sequence analysis, further cloning	tga2_Sotub430_F	ATGGCAGACGTCAGTGCTATGA
coding sequence	StTGA2.3	Sotub01g009430	pJET cloning	sequence analysis, further cloning	tga2_Sotub430_R	TTACTCTCGGGGACGGGCAA
coding sequence	StNPR1	Sotub07g011600	in vivo cloning	yeast two-hybrid assay	NPR1_pGADT7_Fw	ATGGCCATGGAGGCCAGTGAATTCCACATGGATAGTAGGACTGCTTTTTCG GATTCGAAT
coding sequence	StNPR1	Sotub07g011600	in vivo cloning	yeast two-hybrid assay	pGBKT7_NPR1_F	ATCTCAGAGGAGGACCTGCATATGGCCATGGATAGTAGGACTGCTTTTTCG GATTC
coding sequence	StNPR1	Sotub07g011600	in vivo cloning	yeast two-hybrid assay	NPR1_pGADT7_new_Rv	ATCTGCAGCTCGAGCTCGATGGATCTATTTCCTAAAAGGGAGATTATTGGG CTTATCTAC
coding sequence	StNPR1	Sotub07g011600	in vivo cloning	yeast two-hybrid assay	NPR1_pGBKT7_new_Rv	AGAGGCCCCAAGGGGTTATGCTAGTTATGCCTATTTCCTAAAAGGGAGATT ATTGGGCTTATCTAC
coding sequence	StNPR3/4	Sotub02g015550	in vivo cloning	yeast two-hybrid assay	NPR3/4A_pGADT7_Fw	ATGGCCATGGAGGCCAGTGAATTCCACATGGGAAGTTCTGCTGAACCATCA TCTTCT
coding sequence	StNPR3/4	Sotub02g015550	in vivo cloning	yeast two-hybrid assay	NPR3/4A_pGBKT7_Fw	ATCTCAGAGGAGGACCTGCATATGGCCATGGGAAGTTCTGCTGAACCATCA TCTTCT
coding sequence	StNPR3/4	Sotub02g015550	in vivo cloning	yeast two-hybrid assay	NPR3/4A_pGADT7_Rv	ATCTGCAGCTCGAGCTCGATGGATTCATAGGTTCCTAGCTTTGACACTTGCA CC
coding sequence	StNPR3/4	Sotub02g015550	in vivo cloning	yeast two-hybrid assay	NPR3/4A_pGBKT7_Rv	AGAGGCCCCAAGGGGTTATGCTAGTTATGCTCATAGGTTCCTAGCTTTGAC ACTTGCACC
coding sequence	StTGA2.1	Sotub10g022560	in vivo cloning	yeast two-hybrid assay	TGA2.1_pGADT7_Fw	ATGGCCATGGAGGCCAGTGAATTCCACATGGATGCAATGGGTGATAGGGAA GGCCC
coding sequence	StTGA2.1	Sotub10g022560	in vivo cloning	yeast two-hybrid assay	TGA2.1_pGBKT7_Fw	ATCTCAGAGGAGGACCTGCATATGGCCATGGATGCAATGGGTGATAGGGAA GGCCC
coding sequence	StTGA2.1	Sotub10g022560	in vivo cloning	yeast two-hybrid assay	TGA2.1_2.2_pGADT7_Rv	ATCTGCAGCTCGAGCTCGATGGATTTATTGCTCTCGTGGTCTGGCAAGCC
coding sequence	StTGA2.1	Sotub10g022560	in vivo cloning	yeast two-hybrid assay	TGA2.1_2.2_pGBKT7_Rv	AGAGGCCCCAAGGGGTTATGCTAGTTATGCTTATTGCTCTCGTGGTCTGGC AAGCC
coding sequence	StTGA2.2	Sotub10g022550	in vivo cloning	yeast two-hybrid assay	TGA2.2_pGADT7_Fw	ATGGCCATGGAGGCCAGTGAATTCCACATGCCCAGTTTCATTTCTCAGATTC CAGTTTC
coding sequence	StTGA2.2	Sotub10g022550	in vivo cloning	yeast two-hybrid assay	TGA2.2_pGBKT7_Fw	ATCTCAGAGGAGGACCTGCATATGGCCATGCCCAGTTTCATTTCTCAGATTC CAGTTTC
coding sequence	StTGA2.2	Sotub10g022550	in vivo cloning	yeast two-hybrid assay	TGA2.1_2.2_pGADT7_Rv	ATCTGCAGCTCGAGCTCGATGGATTTATTGCTCTCGTGGTCTGGCAAGCC
coding sequence	StTGA2.2	Sotub10g022550	in vivo cloning	yeast two-hybrid assay	TGA2.1_2.2_pGBKT7_Rv	AGAGGCCCCAAGGGGTTATGCTAGTTATGCTTATTGCTCTCGTGGTCTGGC AAGCC
coding sequence	StTGA2.3	Sotub01g009430	in vivo cloning	yeast two-hybrid assay	TGA2.3_pGADT7_Fw	ATGGCCATGGAGGCCAGTGAATTCCACATGGCAGACGTCAGTGCTATGACT GATAC
coding sequence	StTGA2.3	Sotub01g009430	in vivo cloning	yeast two-hybrid assay	TGA2.3_pGBKT7_Fw	ATCTCAGAGGAGGACCTGCATATGGCCATGGCAGACGTCAGTGCTATGACT GATAC
coding sequence	StTGA2.3	Sotub01g009430	in vivo cloning	yeast two-hybrid assay	TGA2.3_pGADT7_Rv	ATCTGCAGCTCGAGCTCGATGGATTTACTCTCGGGGACGGGCAAGCCA
coding sequence	StTGA2.3	Sotub01g009430	in vivo cloning	yeast two-hybrid assay	TGA2.3_pGBKT7_Rv	AGAGGCCCCAAGGGGTTATGCTAGTTATGCTTACTCTCGGGGACGGGCAA GCCA
promoter	StPRX07	Sotub09g020950	confirmation/sequencing	na	p7_per950_F1	CACCGTCAAATTTGAACGGACCATAATGGG
promoter	StPRX07	Sotub09g020950	confirmation/sequencing	na	P950(f2)_1000	ACCAAACTACATACCTTAGGATAATGACT
promoter	StPRX07	Sotub09g020950	confirmation/sequencing	na	P950(f4)_2460	CCAACGTCACAATCCTTAACCTCC
promoter	StPRX07	Sotub09g020950	confirmation/sequencing	na	P950(r2)_2500	GTGTGGGCTCGAGACATCG
promoter	StPRX07	Sotub09g020950	confirmation/sequencing	na	P950(r3)_1500	ATCTCTCAATCATTCCATCTAAAACCC
promoter	StPRX07	Sotub09g020950	confirmation/sequencing	na	P950(r4)_800	GGTTTATTCAGAAAGTCAGCATCCCG
promoter	StPRX15	Sotub02g035680	confirmation/sequencing	na	p2_per680_F1	CACCTCCTGCTAATGACACGTATCCCT
promoter	StPRX15	Sotub02g035680	confirmation/sequencing	na	P680(f2)_1100	CCTAACTTTAAATTGGGAGGACAAAACA
promoter	StPRX15	Sotub02g035680	confirmation/sequencing	na	P680(f4)_2650	CACTCATATTTTCAGAGCTCTAATATCA
promoter	StPRX15	Sotub02g035680	confirmation/sequencing	na	P680(r2)_2300	GGATATTCTGTATGTGAATGAGGCAA
promoter	StPRX15	Sotub02g035680	confirmation/sequencing	na	P680(r3)_1400	ATGCTGCTTCTATTATTTTCTCTGC
promoter	StPRX15	Sotub02g035680	confirmation/sequencing	na	P680(r4)_500	ACAAGAGTGTTGGGGATAGGACG
promoter	StPRX46	Sotub03g007840	confirmation/sequencing	na	P840(f2)_750	CCCAACCACTCTTAATACAATCATATTTAC
promoter	StPRX46	Sotub03g007840	confirmation/sequencing	na	P840(r2)_750	GTAAAATAGATTCTATTAAGAGTGGTTGGG
promoter	StPRX07	Sotub09g020950	pENTR cloning	promoter analysis, transactivation assay	p7_per950_F2	CACCGTCGACTAGGCTGTTGTGTCTTA
promoter	StPRX07	Sotub09g020950	pENTR cloning	promoter analysis, transactivation assay	p7_per950_R1	GATGTTAATGAAATCAACACACCCAAACT
promoter	StPRX07	Sotub09g020950	pENTR cloning	promoter analysis, transactivation assay	p7_per950_R2	CAATAGTTTCAGTTTGTGGACATGTTG
promoter	StPRX15	Sotub02g035680	pENTR cloning	promoter analysis	p2_per680_F2	CACCTCAAACAACAGAGGTGTCAGAGG
promoter	StPRX15	Sotub02g035680	pENTR cloning	promoter analysis	p2_per680_R2	TGGACATGATCTCCGGTAATAAGC
promoter	StPRX46	Sotub03g007840	pENTR cloning	promoter analysis	p3_per840_F2	CACCCAATCAGCAGCCATATAAAGTAAGC
promoter	StPRX46	Sotub03g007840	pENTR cloning	promoter analysis	p3_per840_R2	CTTGAGTTAACATTTCTGAATTTTGCTC
vector	pENTR	Invitrogen (USA)	confirmation/sequencing	na	M13 Forward (-20)	GTAAAACGACGGCCAG
vector	pENTR	Invitrogen (USA)	confirmation/sequencing	na	M13 Reverse	CAGGAAACAGCTATGAC
vector	pJET1.2/blunt	Thermo Scientific (USA)	confirmation/sequencing	na	pJET1.2 Forward Sequencing Primer	CGACTCACTATAGGGAGAGCGGC
vector	pJET1.2/blunt	Thermo Scientific (USA)	confirmation/sequencing	na	pJET1.2 Reverse Sequencing Primer	AAGAACATCGATTTTCCATGGCAG
vector	pMCSG7	Eschenfeldt *et al.* (2009)	confirmation/sequencing	na	T7 prom primer (forward)	TAATACGACTCACTATAGGG
vector	pMCSG7	Eschenfeldt *et al.* (2009)	confirmation/sequencing	na	T7 ter primer (reverse)	GCTAGTTATTGCTCAGCGG
vector	pTA7002	Aoyama & Chua (1997)	confirmation/sequencing	na	pTA7002_F	ACCTCGATCGAGATCTTCGCA
vector	pTA7002	Aoyama & Chua (1997)	confirmation/sequencing	na	pTA7002_R	GTGTGGGCAATGAAACTGATGC
vector	pEPQD0KN0025	Dudley *et al.* (2021)	confirmation/sequencing	na	QMD060	TAATACGACTCACTATAGGG
vector	pEPQD0KN0025	Dudley *et al.* (2021)	confirmation/sequencing	na	QMD061	CTAGTTATTGCTCAGCGGT
vector	pEPQD0KN0025	Dudley *et al.* (2021)	confirmation/sequencing	na	QMD062	TCGCCACCTCTGACTTGAG
vector	pEPQD0KN0025	Dudley *et al.* (2021)	confirmation/sequencing	na	QMD063	TTTCAAAAATATGGTATTGATAATCCTG

Appendix D

Permission for Reproduction of Included Publications

D.1 Permission for Reproduction of Publication 2.1

Open Access Article

Differential Response of Grapevine to Infection with '*Candidatus* Phytoplasma solani' in Early and Late Growing Season through Complex Regulation of mRNA and Small RNA Transcriptomes

by Marina Dermastia [1,*], Blaž Škrlj [2,3], Rebeka Strah [1,3], Barbara Anžič [1], Špela Tomaž [1,3], Maja Križnik [1], Christina Schönhuber [4], Monika Riedle-Bauer [5], Živa Ramšak [1], Marko Petek [1], Aleš Kladnik [6], Nada Lavrač [2], Kristina Gruden [1], Thomas Roitsch [7], Günter Brader [4] and Maruša Pompe-Novak [1,8]

[1] National Institute of Biology, 1000 Ljubljana, Slovenia
[2] Jožef Stefan Institute, 1000 Ljubljana, Slovenia
[3] Jožef Stefan International Postgraduate School, 1000 Ljubljana, Slovenia
[4] Bioresources Unit, Austrian Institute of Technology, 3430 Tulln, Austria
[5] Federal College and Research Institute for Viticulture and Pomology, 3400 Klosterneuburg, Austria
[6] Department of Biology, Biotechnical Faculty, University of Ljubljana, 1000 Ljubljana, Slovenia
[7] Department of Plant and Environmental Sciences, University of Copenhagen, 2630 Taastrup, Denmark
[8] School of Viticulture and Enology, University of Nova Gorica, 5271 Vipava, Slovenia
[*] Author to whom correspondence should be addressed.

Academic Editors: Giorgio Gambino and Chiara Pagliarani

Int. J. Mol. Sci. **2021**, *22*(7), 3531; https://doi.org/10.3390/ijms22073531

Received: 23 February 2021 / Revised: 24 March 2021 / Accepted: 25 March 2021 / Published: 29 March 2021

(This article belongs to the Special Issue Molecular Bases of Stress Adaptation in Plants: The Contribute of Regulation by Small RNAs in Plant Development and Stress Response)

View Full-Text | Download PDF | Browse Figures | Citation Export

Abstract

Bois noir is the most widespread phytoplasma grapevine disease in Europe. It is associated with '*Candidatus* Phytoplasma solani', but molecular interactions between the causal pathogen and its host plant are not well understood. In this work, we combined the analysis of high-throughput RNA-Seq and sRNA-Seq data with interaction network analysis for finding new cross-talks among pathways involved in infection of grapevine cv. Zweigelt with '*Ca.* P. solani' in early and late growing seasons. While the early growing season was very dynamic at the transcriptional level in asymptomatic grapevines, the regulation at the level of small RNAs was more pronounced later in the season when symptoms developed in infected grapevines. Most differentially expressed small RNAs were associated with biotic stress. Our study also exposes the less-studied role of hormones in disease development and shows that hormonal balance was already perturbed before symptoms development in infected grapevines. Analysis at the level of communities of genes and mRNA-microRNA interaction networks revealed several new genes (e.g., expansins and cryptdin) that have not been associated with phytoplasma pathogenicity previously. These novel actors may present a new reference framework for research and diagnostics of phytoplasma diseases of grapevine. View Full-Text

Keywords: '*Candidatus* Phytoplasma solani'; grapevine; bois noir; RNA-Seq; sRNA-Seq; miRNA; phasiRNA; hormones; interaction network

► Show Figures

D.2 Permission for Reproduction of Publication 2.3

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 26 July 2022
DOI 10.3389/fpls.2022.935819

OPEN ACCESS

EDITED BY
Ian T. Major,
Canadian Forest Service, Canada

REVIEWED BY
Mark Zander,
Rutgers, The State University of New Jersey, United States
Shelley Hepworth,
Carleton University, Canada

*CORRESPONDENCE
Špela Tomaž
spela.tomaz@nib.si

SPECIALTY SECTION
This article was submitted to Plant Cell Biology, a section of the journal Frontiers in Plant Science

RECEIVED 04 May 2022
ACCEPTED 04 July 2022
PUBLISHED 26 July 2022

CITATION
Tomaž Š, Gruden K and Coll A (2022) TGA transcription factors—Structural characteristics as basis for functional variability.
Front. Plant Sci. 13:935819.
doi: 10.3389/fpls.2022.935819

TGA transcription factors—Structural characteristics as basis for functional variability

Špela Tomaž[1,2]*, Kristina Gruden[1] and Anna Coll[1]

[1]Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia, [2]Jožef Stefan International Postgraduate School, Ljubljana, Slovenia

TGA transcription factors are essential regulators of various cellular processes, their activity connected to different hormonal pathways, interacting proteins and regulatory elements. Belonging to the basic region leucine zipper (bZIP) family, TGAs operate by binding to their target DNA sequence as dimers through a conserved bZIP domain. Despite sharing the core DNA-binding sequence, the TGA paralogues exert somewhat different DNA-binding preferences. Sequence variability of their N- and C-terminal protein parts indicates their importance in defining TGA functional specificity through interactions with diverse proteins, affecting their DNA-binding properties. In this review, we provide a short and concise summary on plant TGA transcription factors from a structural point of view, including the relation of their structural characteristics to their functional roles in transcription regulation.

KEYWORDS

DOG1 domain, functional variability, intrinsically disordered regions, plant transcription regulation, post-translational modifications, structural characteristics, TGA transcription factors

Introduction

The *Arabidopsis thaliana* genome encodes for over 2,200 transcription factor genes, according to the Plant Transcription Factor Database[1] yet few of them have been thoroughly characterized. The TGACG-binding (TGA) transcription factors were among the first plant transcription factors ever studied, their discovery dating back to the year of 1989 (Katagiri et al., 1989). Named after their hallmark binding site, the TGA factors became known for their regulation of defense-related genes through interaction with NON-EXPRESSOR OF PR-1 (NPR1) cofactor (Zhang et al., 1999), a salicylic acid receptor and master regulator of plant immunity (Wu et al., 2012; Backer et al., 2019; Wang W. et al., 2020). Among dicot plant species, the ten Arabidopsis TGA factors,

1 http://planttfdb.gao-lab.org

D.3 Permission for Reproduction of Publication 2.4

Š. Tomaž, M. Petek, T. Lukan, K. Pogačar, K. Stare, E. Teixeira Prates, D. A. Jacobson, J. Zrimec, G. Bajc, M. Butala, M. Pompe Novak, Q. Dudley, N. Patron, A. Taler-Verčič, A. Usenik, D. Turk, S. Prat, A. Coll and K. Gruden, "A Mini-TGA Protein Modulates Gene Expression through Heterogeneous Association with Transcription Factors," *Plant Physiology,* Accepted under minor revisions, 2022.

References

Albertazzi, G., Milc, J., Caffagni, A., Francia, E., Roncaglia, E., Ferrari, F., … Pecchioni, N. (2009). Gene expression in grapevine cultivars in response to Bois Noir phytoplasma infection. *Plant Science*, *176*(6), 792–804. https://doi.org/10.1016/j.plantsci.2009.03.001

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Analyzing Protein Structure and Function. In *Molecular Biology of the Cell* (4th ed.). New York: Garland Science. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK26820/

Backer, R., Naidoo, S., & Berg, N. Van Den. (2019). The NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) and related family: Mechanistic insights in plant disease resistance. *Frontiers in Plant Science*, *10*, 102. https://doi.org/10.3389/fpls.2019.00102

Baebler, Š., Coll, A., & Gruden, K. (2020). Plant molecular responses to potato virus Y: A continuum of outcomes from sensitivity and tolerance to resistance. *Viruses*, *12*(2), 217. https://doi.org/10.3390/v12020217

Baebler, Š., Stare, K., Kovač, M., Blejec, A., Prezelj, N., Stare, T., … Gruden, K. (2011). Dynamics of responses in compatible potato - potato virus Y interaction are modulated by salicylic acid. *PLoS ONE*, *6*(12), e29009. https://doi.org/10.1371/journal.pone.0029009

Baebler, Š., Witek, K., Petek, M., Stare, K., Tušek-Žnidarič, M., Pompe-Novak, M., … Hennig, J. (2014). Salicylic acid is an indispensable component of the Ny-1 resistance-gene-mediated response against Potato virus Y infection in potato. *Journal of Experimental Botany*, *65*(4), 1095–1109. https://doi.org/10.1093/jxb/ert447

Bai, X., Correa, V. R., Toruño, T. Y., Ammar, E.-D., Kamoun, S., & Hogenhout, S. A. (2009). AY-WB phytoplasma secretes a protein that targets plant cell nuclei. *Molecular Plant-Microbe Interactions*, *22*(1), 18–30. https://doi.org/10.1094/MPMI-22-1-0018

Barbara, D. J., Morton, A., Clark, M. F., & Davies, D. L. (2002). Immunodominant membrane proteins from two phytoplasmas in the aster yellows clade (chlorante aster yellows and clover phyllody) are highly divergent in the major hydrophilic region. *Microbiology*, *148*(1), 157–167. https://doi.org/10.1099/00221287-148-1-157

Bertaccini, A., & Lee, I.-M. (2018). Phytoplasmas: An update. In G. P. Rao, A. Bertaccini, N. Fiore, & L. W. Liefting (Eds.), *Phytoplasmas: Plant Pathogenic Bacteria - I: Characterisation and Epidemiology of Phytoplasma - Associated Diseases* (1st ed., pp. 1–29). Springer, Singapore. https://doi.org/10.1007/978-981-13-0119-3

Bertaccini, A., Oshima, K., Maejima, K., & Namba, S. (2019). Phytoplasma effectors and

pathogenicity factors. In A. Bertaccini, K. Oshima, M. Kube, & G. P. Rao (Eds.), *Phytoplasmas: Plant Pathogenic Bacteria - III: Genomics, Host Pathogen Interactions and Diagnosis* (1st ed., pp. 17–34). Springer, Singapore. https://doi.org/10.1007/978-981-13-9632-8_2

Bertamini, M., Nedunchezhian, N., Tomasi, F., & Grando, M. S. (2002). Phytoplasma [Stolbur-subgroup (Bois Noir-BN)] infection inhibits photosynthetic pigments, ribulose-1,5-bisphosphate carboxylase and photosynthetic activities in field grown grapevine (Vitis vinifera L. cv. Chardonnay) leaves. *Physiological and Molecular Plant Pathology*, *61*(6), 357–366. https://doi.org/10.1006/pmpp.2003.0449

Bleau, J. R., & Spoel, S. H. (2021). Selective redox signaling shapes plant-pathogen interactions. *Plant Physiology*, *186*(1), 53–65. https://doi.org/10.1093/plphys/kiaa088

Boyle, P., Le Su, E., Rochon, A., Shearer, H. L., Murmu, J., Chu, J. Y., … Despres, C. (2009). The BTB/POZ domain of the Arabidopsis disease resistance protein NPR1 interacts with the repression domain of TGA2 to negate its function. *The Plant Cell*, *21*(11), 3700–3713. https://doi.org/10.1105/tpc.109.069971

Calil, I. P., & Fontes, E. P. B. (2017). Plant immunity against viruses: Antiviral immune receptors in focus. *Annals of Botany*, *119*(5), 711–723. https://doi.org/10.1093/aob/mcw200

Carra, A., Gambino, G., & Schubert, A. (2007). A cetyltrimethylammonium bromide-based method to extract low-molecular-weight RNA from polysaccharide-rich plant tissues. *Analytical Biochemistry*, *360*(2), 318–320. https://doi.org/10.1016/j.ab.2006.09.022

Chai, L.-X., Dong, K., Liu, S.-Y., Zhang, Z., Zhang, X.-P., Tong, X., … Wang, X.-B. (2020). A putative nuclear copper chaperone promotes plant immunity in Arabidopsis. *Journal of Experimental Botany*, *71*(20), 6684–6696. https://doi.org/10.1093/jxb/eraa401

Chaïb, J., Torregrosa, L., MacKenzie, D., Corena, P., Bouquet, A., & Thomas, M. R. (2010). The grape microvine - a model system for rapid forward and reverse genetics of grapevines. *Plant Journal*, *62*(6), 1083–1092. https://doi.org/10.1111/j.1365-313X.2010.04219.x

Chang, S., Puryear, J., & Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*, *11*(June), 113–116. https://doi.org/10.1007/BF02670468

Chen, J., Mohan, R., Zhang, Y., Li, M., Chen, H., Palmer, I. A., … Fu, Z. Q. (2019). NPR1 promotes its own and target gene expression in plant defense by recruiting CDK8. *Plant Physiology*, *181*(1), 289–304. https://doi.org/10.1104/pp.19.00124

Dermastia, M., Bertaccini, A., Constable, F., & Mehle, N. (2017). *Grapevine yellows diseases and their phytoplasma agents: Biology and detection* (1st ed.). Springer Cham, Switzerland. https://doi.org/https://doi.org/10.1007/978-3-319-50648-7

Ding, Y., Sun, T., Ao, K., Peng, Y., Zhang, Y., Li, X., & Zhang, Y. (2018). Opposite roles of salicylic acid receptors NPR1 and NPR3/NPR4 in transcriptional regulation of plant immunity. *Cell*, *173*(6), 1454–1467. https://doi.org/10.1016/j.cell.2018.03.044

Ekengren, S. K., Liu, Y., Schiff, M., Dinesh-Kumar, S. P., & Martin, G. B. (2003). Two MAPK cascades, NPR1, and TGA transcription factors play a role in Pto-mediated

disease resistance in tomato. *The Plant Journal*, *36*(6), 905–917. https://doi.org/10.1046/j.1365-313X.2003.01944.x

EPPO Global Database. (2022). Retrieved from https://gd.eppo.int

Fabre, A., Danet, J.-L., & Foissac, X. (2011). The stolbur phytoplasma antigenic membrane protein gene stamp is submitted to diversifying positive selection. *Gene*, *472*(1–2), 37–41. https://doi.org/10.1016/j.gene.2010.10.012

Fan, W., & Dong, X. (2002). In vivo interaction between NPR1 and transcription factor TGA2 leads to salicylic acid-mediated gene activation in Arabidopsis. *The Plant Cell*, *14*(6), 1377–1389. https://doi.org/10.1105/tpc.001628

Federici, F., Dupuy, L., Laplaze, L., Heisler, M., & Haseloff, J. (2012). Integrated genetic and computation methods for in planta cytometry. *Nature Methods*, *9*(5), 483–485. https://doi.org/10.1038/nmeth.1940

Feng, J., Cheng, Y., & Zheng, C. (2020). Expression patterns of octoploid strawberry TGA genes reveal a potential role in response to Podosphaera aphanis infection. *Plant Biotechnology Reports*, *14*(February), 55–67. https://doi.org/10.1007/s11816-019-00582-9

Flor, H. H. (1942). Inheritance of pathogenicity in Melampsora lini. *Phytoplathology*, *32*(8), 653–669.

Gambino, G., Boccacci, P., Margaria, P., Palmano, S., & Gribaudo, I. (2013). Hydrogen peroxide accumulation and transcriptional changes in grapevines recovered from Flavescence dorée disease. *Phytopathology*, *103*(8), 776–784. https://doi.org/10.1094/PHYTO-11-12-0309-R

Gambino, G., Perrone, I., & Gribaudo, I. (2008). A rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochemical Analysis*, *19*(6), 520–525. https://doi.org/10.1002/pca.1078

Gatz, C. (2013). From pioneers to team players: TGA transcription factors provide a molecular link between different stress pathways. *Molecular Plant-Microbe Interactions*, *26*(2), 151–159. https://doi.org/10.1094/MPMI-04-12-0078-IA

Gonzalez, D. H. (2016). Introduction to transcription factor structure and function. In D. H. Gonzalez (Ed.), *Plant Transcription Factors: Evolutionary, Structural and Functional Aspects* (pp. 3–11). Academic Press. https://doi.org/10.1016/B978-0-12-800854-6.00001-4

Guan, L., Haider, M. S., Khan, N., Nasim, M., Jiu, S., Fiaz, M., … Fang, J. (2018). Transcriptome sequence analysis elaborates a complex defensive mechanism of grapevine (Vitis vinifera L.) in response to salt stress. *International Journal of Molecular Sciences*, *19*(12), 4019. https://doi.org/10.3390/ijms19124019

Haider, M. S., Zhang, C., Kurjogi, M. M., Pervaiz, T., Zheng, T., Zhang, C., … Fang, J. (2017). Insights into grapevine defense response against drought as revealed by biochemical, physiological and RNA-Seq analysis. *Scientific Reports*, *7*, 13134. https://doi.org/10.1038/s41598-017-13464-3

Hily, J.-M., Candresse, T., Garcia, S., Vigne, E., Tannière, M., Komar, V., … Lemaire, O. (2018). High-throughput sequencing and the viromic study of grapevine leaves: From the detection of grapevine-infecting viruses to the description of a new environmental Tymovirales member. *Frontiers in Microbiology*, *9*, 1782.

https://doi.org/10.3389/fmicb.2018.01782

Hou, J., Sun, Q., Li, J., Ahammed, G. J., Yu, J., Fang, H., & Xia, X. (2019). Glutaredoxin S25 and its interacting TGACG motif-binding factor TGA2 mediate brassinosteroid-induced chlorothalonil metabolism in tomato plants. *Environmental Pollution*, *255*, 113256. https://doi.org/10.1016/j.envpol.2019.113256

Hren, M., Nikolić, P., Rotter, A., Blejec, A., Terrier, N., Ravnikar, M., … Gruden, K. (2009). "Bois noir" phytoplasma induces significant reprogramming of the leaf transcriptome in the field grown grapevine. *BMC Genomics*, *10*, 460. https://doi.org/10.1186/1471-2164-10-460

Huang, C.-Y., Wang, H., Hu, P., Hamby, R., & Jin, H. (2019). Small RNAs – Big players in plant-microbe interactions. *Cell Host and Microbe*, *26*(2), 173–182. https://doi.org/10.1016/j.chom.2019.07.021

Hussain, R. M. F., Sheikh, A. H., Haider, I., Quareshy, M., & Linthorst, H. J. M. (2018). Arabidopsis WRKY50 and TGA transcription factors synergistically activate expression of PR1. *Frontiers in Plant Science*, *9*, 930. https://doi.org/10.3389/fpls.2018.00930

IRPCM. (2004). "Candidatus Phytoplasma", a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. *International Journal of Systematic and Evolutionary Microbiology*, *54*(4), 1243–1255. https://doi.org/10.1099/ijs.0.02854-0

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

Jović, J., Cvrković, T., Mitrović, M., Krnjajić, S., Redinbaugh, M. G., Pratt, R. C., … Toševski, I. (2007). Roles of stolbur phytoplasma and Reptalus panzeri (Cixiinae, Auchenorrhyncha) in the epidemiology of Maize redness in Serbia. *European Journal of Plant Pathology*, *118*(May), 85–89. https://doi.org/10.1007/s10658-007-9105-0

Karasev, A. V., & Gray, S. M. (2013). Continuous and emerging challenges of Potato virus Y in potato. *Annual Review of Phytopathology*, *51*(1), 571–586. https://doi.org/10.1146/annurev-phyto-082712-102332

Karimi, M., Inzé, D., & Depicker, A. (2002). GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science*, *7*(5), 193–195. https://doi.org/10.1016/S1360-1385(02)02251-3

Kežar, A., Kavčič, L., Polák, M., Nováček, J., Gutiérrez-Aguirre, I., Žnidarič, M. T., … Podobnik, M. (2019). Structural basis for the multitasking nature of the potato virus Y coat protein. *Science Advances*, *5*(7), eaaw3808. https://doi.org/10.1126/sciadv.aaw3808

Klessig, D. F., Tian, M., & Choi, H. W. (2016). Multiple targets of salicylic acid and its derivatives in plants and animals. *Frontiers in Immunology*, *7*, 206. https://doi.org/10.3389/fimmu.2016.00206

Kreuze, J. F., Souza-Dias, J. A. C., Jeevalatha, A., Figueira, A. R., Valkonen, J. P. T., & Jones, R. A. C. (2019). Viral diseases in potato. In H. Campos & O. Ortiz (Eds.), *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind* (1st ed., pp. 389–430). Springer Cham, Switzerland. https://doi.org/10.1007/978-3-030-28683-5_11

Križnik, M., Gruden, K., & Baebler, Š. (2020). Molecular responses of plants to viruses with emphasis on small RNAs. In L. P. Awasthi (Ed.), *Applied Plant Virology: Advances, Detection, and Antiviral Strategies* (1st ed., pp. 141–157). Academic Press. https://doi.org/10.1016/b978-0-12-818654-1.00011-6

Križnik, M., Petek, M., Dobnik, D., Ramšak, Ž., Baebler, Š., Pollmann, S., … Gruden, K. (2017). Salicylic acid perturbs sRNA-gibberellin regulatory network in immune response of potato to potato virus Y infection. *Frontiers in Plant Science*, *8*(December), 1–14. https://doi.org/10.3389/fpls.2017.02192

Kroj, T., Chanclud, E., Michel-Romiti, C., Grand, X., & Morel, J.-B. (2016). Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *New Phytologist*, *210*(2), 618–626. https://doi.org/10.1111/nph.13869

Kumar, S., Zavaliev, R., Wu, Q., Zhou, Y., Cheng, J., & Dillard, L. (2022). Structural basis of NPR1 in activating plant immunity. *Nature*, *605*(May), 561–566. https://doi.org/10.1038/s41586-022-04699-w

Laimer, M., Lemaire, O., Herrbach, E., Goldschmidt, V., Minafra, A., Bianco, P., & Wetzel, T. (2009). Resistance to viruses, phytoplasmas and their vectors in the grapevine in europe: a review. *Journal of Plant Pathology*, *91*(1), 7–23.

Lazar, A., Coll, A., Dobnik, D., Baebler, Š., Bedina-Zavec, A., Žel, J., & Gruden, K. (2014). Involvement of potato (Solanum tuberosum L.) MKK6 in response to potato virus Y. *PLoS ONE*, *9*(8), e104553. https://doi.org/10.1371/journal.pone.0104553

Lefeuvre, P., Martin, D. P., Elena, S. F., Shepherd, D. N., Roumagnac, P., & Varsani, A. (2019). Evolution and ecology of plant viruses. *Nature Reviews Microbiology*, *17*(October), 632–644. https://doi.org/10.1038/s41579-019-0232-3

Lemaire-Chamley, M., Koutouan, C., Jorly, J., Assali, J., Yoshida, T., Nogueira, M., … Rothan, C. (2022). A chimeric TGA repressor slows down fruit maturation and ripening in tomato. *Plant and Cell Physiology*, *63*(1), 120–134. https://doi.org/10.1093/pcp/pcab150

Li, N., Muthreich, M., Huang, L.-J., Thurow, C., Sun, T., Zhang, Y., & Gatz, C. (2019). TGACG-BINDING FACTORs (TGAs) and TGA-interacting CC-type glutaredoxins modulate hyponastic growth in Arabidopsis thaliana. *New Phytologist*, *221*(4), 1906–1918. https://doi.org/10.1111/nph.15496

Lukan, T., Baebler, Š., Pompe-Novak, M., Guček, K., Zagorščak, M., Coll, A., & Gruden, K. (2018). Cell death is not sufficient for the restriction of potato virus Y spread in hypersensitive response-conferred resistance in potato. *Frontiers in Plant Science*, *9*, 168. https://doi.org/10.3389/fpls.2018.00168

Lukan, T., Machens, F., Coll, A., Baebler, Š., Messerschmidt, K., & Gruden, K. (2018). Plant X-tender: An extension of the AssemblX system for the assembly and expression of multigene constructs in plants. *PLoS ONE*, *13*(1), e0190526. https://doi.org/10.1371/journal.pone.0190526

Lukan, T., Pompe-Novak, M., Baebler, Š., Tušek-Žnidarič, M., Kladnik, A., Križnik, M., … Gruden, K. (2020). Precision transcriptomics of viral foci reveals the spatial regulation of immune-signaling genes and identifies RBOHD as an important player in the incompatible interaction between potato virus Y and potato. *The Plant Journal*,

104(3), 645–661. https://doi.org/10.1111/tpj.14953

Magnani, E., de Klein, N., Nam, H.-I., Kim, J.-G., Pham, K., Fiume, E., … Rhee, S. Y. (2014). A comprehensive analysis of microProteins reveals their potentially widespread mechanism of transcriptional regulation. *Plant Physiology*, *165*(1), 149–159. https://doi.org/10.1104/pp.114.235903

Malinova, I., Kunz, H.-H., Alseekh, S., Herbst, K., Fernie, A. R., Gierth, M., & Fettke, J. (2014). Reduction of the cytosolic phosphoglucomutase in Arabidopsis reveals impact on plant growth, seed and root development, and carbohydrate partitioning. *PLoS ONE*, *9*(11), e112468. https://doi.org/10.1371/journal.pone.0112468

Mandahar, C. L. (2006). *Multiplication of RNA plant viruses* (1st ed.). Springer Dordecht, Netherlands. https://doi.org/10.1007/1-4020-4725-8

Martin, R., Qi, T., Zhang, H., Liu, F., King, M., Toth, C., … Staskawicz, B. J. (2020). Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. *Science*, *370*(6521), eabd9993. https://doi.org/10.1126/science.abd9993

Martini, M., Murari, E., Mori, N., & Bertaccini, A. (1999). Identification and epidemic distribution of two flavescence doree-related phytoplasmas in Veneto (Italy). *Plant Disease*, *83*(10), 925–930. https://doi.org/10.1094/PDIS.1999.83.10.925

Mitrović, M., Jakovljević, M., Jović, J., Krstić, O., Kosovac, A., Trivellone, V., … Cvrković, T. (2016). 'Candidatus phytoplasma solani' genotypes associated with potato stolbur in Serbia and the role of Hyalesthes obsoletus and Reptalus panzeri (hemiptera, cixiidae) as natural vectors. *European Journal of Plant Pathology*, *144*(March), 619–630. https://doi.org/10.1007/s10658-015-0800-y

Moon, S.-J., Park, H. J., Kim, T.-H., Kang, J.-W., Lee, J.-Y., Cho, J. H., … Shin, D. (2018). OsTGA2 confers disease resistance to rice against leaf blight by regulating expression levels of disease related genes via interaction with NH1. *PloS ONE*, *13*(11), e0206910. https://doi.org/10.1371/journal.pone.0206910

Murmu, J., Bush, M. J., DeLong, C., Li, S., Xu, M., Khan, M., … Hepworth, S. R. (2010). Arabidopsis basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. *Plant Physiology*, *154*(3), 1492–1504. https://doi.org/10.1104/pp.110.159111

Murolo, S., Garbarino, M., Mancini, V., & Romanazzi, G. (2020). Spatial pattern of Bois noir: case study of a delicate balance between disease progression and recovery. *Scientific Reports*, *10*, 9801. https://doi.org/10.1038/s41598-020-66210-7

Music, M. S., Samarzija, I., Hogenhout, S. A., Haryono, M., Cho, S.-T., & Kuo, C.-H. (2019). The genome of 'Candidatus Phytoplasma solani' strain SA-1 is highly dynamic and prone to adopting foreign sequences. *Systematic and Applied Microbiology*, *42*(2), 117–127. https://doi.org/10.1016/j.syapm.2018.10.008

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

Niggeweg, R., Thurow, C., Weigel, R., Pfitzner, U., & Gatz, C. (2000). Tobacco TGA factors differ with respect to interaction with NPR1, activation potential and DNA-

binding properties. *Plant Molecular Biology*, *42*(March), 775–788. https://doi.org/10.1023/A:1006319113205

Nishiyama, E., Nonogaki, M., Yamazaki, S., Nonogaki, H., & Ohshima, K. (2021). Ancient and recent gene duplications as evolutionary drivers of the seed maturation regulators DELAY OF GERMINATION1 family genes. *New Phytologist*, *230*(3), 889–901. https://doi.org/10.1111/nph.17201

Oerke, E. C. (2006). Crop losses to pests. *Journal of Agricultural Science*, *144*(1), 31–43. https://doi.org/10.1017/S0021859605005708

Panchy, N., Lehti-Shiu, M., & Shiu, S.-H. (2016). Evolution of gene duplication in plants. *Plant Physiology*, *171*(4), 2294–2316. https://doi.org/10.1104/pp.16.00523

Peng, Y., Yang, J., Li, X., & Zhang, Y. (2021). Salicylic acid: Biosynthesis and signaling. *Annual Review of Plant Biology*, *72*, 761–791. https://doi.org/10.1146/annurev-arplant-081320-092855

Pracros, P., Renaudin, J., Eveillard, S., Mouras, A., & Hernould, M. (2006). Tomato flower abnormalities induced by stolbur phytoplasma infection are associated with changes of expression of floral development genes. *Molecular Plant-Microbe Interactions*, *19*(1), 62–68. https://doi.org/10.1094/MPMI-19-0062

Quaglino, F., Zhao, Y., Casati, P., Bulgari, D., Bianco, P. A., Wei, W., & Davis, R. E. (2013). "Candidatus Phytoplasma solani", a novel taxon associated with stolbur- and bois noir-related diseases of plants. *International Journal of Systematic and Evolutionary Microbiology*, *63*, 2879–2894. https://doi.org/10.1099/ijs.0.044750-0

Reiter, F., Wienerroither, S., & Stark, A. (2017). Combinatorial function of transcription factors and cofactors. *Current Opinion in Genetics and Development*, *43*, 73–81. https://doi.org/10.1016/j.gde.2016.12.007

Rochon, A., Boyle, P., Wignes, T., Fobert, P. R., & Després, C. (2006). The coactivator function of Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. *The Plant Cell*, *18*(12), 3670–3685. https://doi.org/10.1105/tpc.106.046953

Roeder, R. G. (2019). 50+ years of eukaryotic transcription: An expanding universe of factors and mechanisms. *Nature Structural and Molecular Biology*, *26*(September), 783–791. https://doi.org/10.1038/s41594-019-0287-x

Rubio, L., Galipienso, L., & Ferriol, I. (2020). Detection of plant viruses and disease management: Relevance of genetic diversity and evolution. *Frontiers in Plant Science*, *11*, 1092. https://doi.org/10.3389/fpls.2020.01092

Salzman, R. A., Fujita, T., Zhu-Salzman, K., Hasegawa, P. M., & Bressan, R. A. (1999). An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. *Plant Molecular Biology Reporter*, *17*(March), 11–17. https://doi.org/10.1023/A:1007520314478

Santi, S., De Marco, F., Polizzotto, R., Grisan, S., & Musetti, R. (2013). Recovery from stolbur disease in grapevine involves changes in sugar transport and metabolism. *Frontiers in Plant Science*, *4*, 171. https://doi.org/10.3389/fpls.2013.00171

Santi, S., Grisan, S., Pierasco, A., De Marco, F., & Musetti, R. (2013). Laser microdissection of grapevine leaf phloem infected by stolbur reveals site-specific gene responses associated to sucrose transport and metabolism. *Plant, Cell and*

Environment, *36*(2), 343–355. https://doi.org/10.1111/j.1365-3040.2012.02577.x

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology and Evolution*, *3*(March), 430–439. https://doi.org/10.1038/s41559-018-0793-y

Schiermeyer, A., Thurow, C., & Gatz, C. (2003). Tobacco bZIP factor TGA10 is a novel member of the TGA family of transcription factors. *Plant Molecular Biology*, *51*(April), 817–829. https://doi.org/10.1023/A:1023093101976

Sharma, A., Abrahamian, P., Carvalho, R., Choudhary, M., Paret, M. L., Vallad, G. E., & Jones, J. B. (2022). Future of bacterial disease management in crop production. *Annual Review of Phytopathology*, *60*, 259–282. https://doi.org/10.1146/annurev-phyto-021621-121806

Strange, R. N., & Scott, P. R. (2005). Plant disease: A threat to global food security. *Annual Review of Phytopathology*, *43*, 83–116. https://doi.org/10.1146/annurev.phyto.43.113004.133839

Sugio, A., Kingdom, H. N., MacLean, A. M., Grieve, V. M., & Hogenhout, S. A. (2011). Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(48), E1254–E1263. https://doi.org/10.1073/pnas.1105664108

Sugio, A., Maclean, A. M., & Hogenhout, S. A. (2014). The small phytoplasma virulence effector SAP11 contains distinct domains required for nuclear targeting and CIN-TCP binding and destabilization. *New Phytologist*, *202*(3), 838–848. https://doi.org/10.1111/nph.12721

Szajko, K., Chrzanowska, M., Witek, K., Strzelczyk-Zyta, D., Zagórska, H., Gebhardt, C., … Marczewski, W. (2008). The novel gene Ny-1 on potato chromosome IX confers hypersensitive resistance to Potato virus Y and is an alternative to Ry genes in potato breeding for PVY resistance. *Theoretical and Applied Genetics*, *116*(January), 297–303. https://doi.org/10.1007/s00122-007-0667-1

Thurow, C., Schiermeyer, A., Krawczyk, S., Butterbrodt, T., Nickolov, K., & Gatz, C. (2005). Tobacco bZIP transcription factor TGA2.2 and related factor TGA2.1 have distinct roles in plant defense responses and plant development. *The Plant Journal*, *44*(1), 100–113. https://doi.org/10.1111/j.1365-313X.2005.02513.x

Tsuda, K., & Somssich, I. E. (2015). Transcriptional networks in plant immunity. *New Phytologist*, *206*(3), 932–947. https://doi.org/10.1111/nph.13286

Valli, A. A., Gallo, A., Rodamilans, B., López-Moya, J. J., & García, J. A. (2018). The HCPro from the Potyviridae family: an enviable multitasking Helper Component that every virus would like to have. *Molecular Plant Pathology*, *19*(3), 744–763. https://doi.org/10.1111/mpp.12553

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Alfenas-Zerbini, P., … Zerbini, F. M. (2021). Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). *Archives of Virology*, *166*(September), 2633–2648. https://doi.org/10.1007/s00705-021-05156-1

Wang, J., Hu, M., Wang, J., Qi, J., Han, Z., Wang, G., … Chai, J. (2019). Reconstitution

and structure of a plant NLR resistosome conferring immunity. *Science*, *364*(6435), eaav5870. https://doi.org/10.1126/science.aav5870

Wang, W., Withers, J., Li, H., Zwack, P. J., Rusnac, D.-V., Shi, H., … Zheng, N. (2020). Structural basis of salicylic acid perception by Arabidopsis NPR proteins. *Nature*, *586*(October), 311–316. https://doi.org/10.1038/s41586-020-2596-y

Weintraub, P. G., & Beanland, L. A. (2006). Insect vectors of phytoplasmas. *Annual Review of Entomology*, *51*, 91–111. https://doi.org/10.1146/annurev.ento.51.110104.151039

Wu, C.-H., Derevnina, L., & Kamoun, S. (2018). Receptor networks underpin plant immunity. *Science*, *360*(6395), 1300–1301. https://doi.org/10.1126/science.aat2623

Wu, Y., Zhang, D., Chu, J. Y., Boyle, P., Wang, Y., Brindle, I. D., … Després, C. (2012). The Arabidopsis NPR1 protein is a receptor for the plant defense hormone salicylic acid. *Cell Reports*, *1*(6), 639–647. https://doi.org/10.1016/j.celrep.2012.05.008

Yuan, M., Jiang, Z., Bi, G., Nomura, K., Liu, M., Wang, Y., … Xin, X.-F. (2021). Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature*, *592*(April), 105–109. https://doi.org/10.1038/s41586-021-03316-6

Yuan, M., Ngou, B. P. M., Ding, P., & Xin, X.-F. (2021). PTI-ETI crosstalk: an integrative view of plant immunity. *Current Opinion in Plant Biology*, *62*(August), 102030. https://doi.org/10.1016/j.pbi.2021.102030

Zhang, Y., Cheng, Y. T., Qu, N., Zhao, Q., Bi, D., & Li, X. (2006). Negative regulation of defense responses in Arabidopsis by two NPR1 paralogs. *The Plant Journal*, *48*(5), 647–656. https://doi.org/10.1111/j.1365-313X.2006.02903.x

Zhang, Y., Fan, W., Kinkema, M., Li, X., & Dong, X. (1999). Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-1 gene. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(11), 6523–6528. https://doi.org/10.1073/pnas.96.11.6523

Zhang, Y., Tessaro, M. J., Lassner, M., & Li, X. (2003). Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *The Plant Cell*, *15*(11), 2647–2653. https://doi.org/10.1105/tpc.014894

Zhong, L., Chen, D., Min, D., Li, W., Xu, Z., Zhou, Y., … Ma, Y. (2015). AtTGA4, a bZIP transcription factor, confers drought resistance by enhancing nitrate transport and assimilation in Arabidopsis thaliana. *Biochemical and Biophysical Research Communications*, *457*(3), 433–439. https://doi.org/10.1016/j.bbrc.2015.01.009

Zhou, J.-M., Trifa, Y., Silva, H., Pontier, D., Lam, E., Shah, J., & Klessig, D. F. (2000). NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the PR-1 gene required for induction by salicylic acid. *Molecular Plant-Microbe Interactions*, *13*(2), 191–202. https://doi.org/10.1094/MPMI.2000.13.2.191

Bibliography

Publications Related to the Thesis

Journal Articles

M. Dermastia, B. Škrlj, R. Strah, B. Anžič, Š. Tomaž, M. Križnik, C. Schönhuber, M. Riedle-Bauer, Ž. Ramšak, M. Petek, A. Kladnik, N. Lavrač, K. Gruden, T. Roitsch, G. Brader and M. Pompe-Novak, "Differential Response of Grapevine to Infection with '*Candidatus* Phytoplasma solani' in Early and Late Growing Season Through Complex Regulation of mRNA and Small RNA Transcriptomes," *International Journal of Molecular Sciences,* vol. 22, no. 7:3531, 2021.

M. Dermastia, Š. Tomaž, R. Strah, T. Čepin, T. Lukan, A. Coll, B. Dušak, B. Anžič, T. Roitsch, S. Wienkoop, W. Weckwert, K. Gruden, M. Pompe-Novak and G. Brader, "Candidate Effector Proteins of '*Candidatus* Phytoplasma solani' are Associated with Modulation of Plant Carbohydrate Metabolism Toward Effective Glycolysis, with Accelerated Ascorbate-Glutathione Cycle, and with Induction of Autophagosomes," Manuscript draft.

Š. Tomaž, K. Gruden and A. Coll, "TGA Transcription Factors – Structural Characteristics as Basis for Functional Variability," *Frontiers in Plant Science,* vol. 13, no. 935819, 2022.

Š. Tomaž, M. Petek, T. Lukan, K. Pogačar, K. Stare, E. Teixeira Prates, D. A. Jacobson, J. Zrimec, G. Bajc, M. Butala, M. Pompe Novak, Q. Dudley, N. Patron, A. Taler-Verčič, A. Usenik, D. Turk, S. Prat, A. Coll and K. Gruden, "A Mini-TGA Protein Modulates Gene Expression through Heterogeneous Association with Transcription Factors," *Plant Physiology,* Accepted under minor revisions, 2022.

Biography

Špela Tomaž was born on 01.02.1993 in Trbovlje, Slovenia. She graduated from secondary school with honors in 2011 and enrolled into the Biochemistry study program at the Faculty of Chemistry and Chemical Technology (University of Ljubljana, Ljubljana, Slovenia). She received her bachelor's degree in 2014 with her thesis titled "Real-time reverse transcription polymerase chain reaction development for Dobrava-Kurkino virus detection" under the supervision of Prof. Dr. Tatjana Avišč-Županc (Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia). Špela completed her studies in 2017 with a master's thesis titled "Production of potato immune signalization and potato virus Y proteins in *Escherichia coli* and crystallization of potato protein TGA2.1" under the supervision of Prof. Dr. Kristina Gruden (National Institute of Biology, Ljubljana, Slovenia) and Prof. Dr. Dušan Turk (Jožef Stefan Institute, Ljubljana, Slovenia). Her master's thesis was later awarded the 2018 Faculty Prešeren Award for exceptional work. In 2017, she enrolled into the PhD study programme of Nanosciences and Nanotechnologies at the Jožef Stefan International Postgraduate School (Ljubljana, Slovenia) and began her work as a young researcher at the National Institute of Biology (Ljubljana, Slovenia), Department of Biotechnology and Systems Biology, under the supervision of Dr. Anna Coll. Her field of work involved studying the interactions between cultivated plants and their pathogens at the molecular level, specifically the relationship between potato and potato virus Y (PVY) and between grapevine and '*Candidatus* Phytoplasma solani' (phytoplasma). In studying the potato-PVY pathosystem, Špela focused on potato proteins involved in transcriptional reprogramming following infection, including the characterization of an atypical transcription factor protein and its role in plant immunity. She is the first author of the resulting research paper and of a review article, addressing the structural-functional features of the same group of transcription factors. In studying the grapevine-phytoplasma pathosystem, Špela was involved in optimization of the RNA isolation procedure from grapevine leaf veins and subsequently co-authored a research paper on transcriptomic changes in grapevine response. She was also involved in identification of phytoplasma effector targets. Špela presented the results of her work on many national and international conferences in the form of posters and lectures, and has frequently attended various training courses. She actively participated in the organization of the 7th Slovenian Symposium on Plant Biology with International Participation (2018) and functioned as editor for the symposium abstract book. Moreover, she participated in various promotional activities, such as "Dan očarljivih rastlin" and "Znanstival", where she has shared scientific knowledge with the public.