Dynamics of Phytoplasma Effector Interactions with Plant TCP Transcription Factor (sub)classes

by

Sylvain Capdevielle

September 2019

A thesis submitted to the University of East Anglia
for the degree of
Doctor of Philosophy

University of East Anglia
John Innes Centre

This thesis is dedicated to my godfather Patrick.

Abstract

Phytoplasmas are bacterial plant pathogens vectored by sap-feeding leafhoppers. These bacteria induce dramatic changes in plant development, such as witches' brooms and secrete effectors that modulate host plant processes. Aster Yellows phytoplasma strain Witches Broom (AY-WB) secretes the effector Secreted AY-WB Protein 11 (SAP11), which binds and destabilizes plant TCP (TEOSINTE BRANCHED1, CYCLOIDEA, PROLIFERATING CELL FACTORS 1 and 2) transcription factors, particularly members of the class II TCPs of the CYCLOIDEA/TEOSINTE BRANCHED 1 (CYC/TB1) and CINCINNATA (CIN)-TCP (sub)classes. Stable production of $SAP11_{AYWB}$ in *A. thaliana* and maize leads to the induction of specific developmental phenotypes, such as witches' brooms and altered leaf shapes, consistent with the destabilisation of Class II CYC/TB1 and CIN-TCPs. SAP11 effector homologs were found in other phytoplasmas, but their binding specificities toward the TCP (sub)classes are unclear. This thesis shows that four divergent SAP11 effector homologs from various phytoplasma groups differentially bind class II TCP (sub)classes, specifically the helix-loop-helix motif of the conserved TCP domain. I extended the analysis to include SAP11 homologs from divergent phytoplasma groups and investigated their phylogeny. I showed that the SAP11 proteins group in five distinct clades. The SAP11 phylogeny is different from the phytoplasma 16S rDNA phylogeny, suggesting horizontal exchange of SAP11 genes among phytoplasmas. Interestingly, the SAP11 effector homologs within each clade bind TCP members from three TCP (sub)classes, including class I TCPs. I elucidated the SAP11 region involved in the binding specificities toward class I and class II TCPs. Finally, SAP effectors of AY-WB phytoplasma genes that lie on the genetic island of $SAP11_{AYWB}$ also interact with TCP (sub)classes, including Class I TCPs, thereby expanding the interaction range of AY-WB SAPs to all TCP (sub)classes. The work enables predictions of binding specificities to TCP (sub)classes of SAP11 effector homologs that may be discovered in the future.

Acknowledgements

The story of my PhD started on the other side of the world, in New Zealand, where I spent a six months master internship in the group of Dr Kee Hoon Sohn. From there, I had a first experience in working in the plant-microbe interaction field in an international lab. I want to thank the "KSL" lab members, Kee, Cécile, Jay, Toby, Maxim and Sera for their warm welcome. A special mention to Jay Jayaraman; I am grateful to have met you, my friend. This Kiwi experience strengthened my desire to explore science outside my comfort zone and I decided to pursue my PhD in the UK, in the Hogenhout Lab.

I would like to particularly thank my primary supervisor Prof. Saskia Hogenhout for the opportunity to work in her lab. Saskia, thank you very much for your support and guidance along the years, and for helping me develop both my scientific and personal growth. I thank my two secondary supervisors, Prof. Sophien Kamoun and Prof. Mark Banfield for their valuable input and their mentoring along my PhD. I couldn't have hoped for a better supervisory team. I also thank the BBSRC and NRP DTP programme for funding my PhD.

I was lucky to find a brilliant lab crew in the UK too, which improved my experience considerably here. Many thanks to Sam, James, Weijie, Yazhou, Adi, Roland, Beccy, Qun, Josh, Matteo, Archana, Michael, Roberto for being great colleagues. I also want to thank the insectary dream team, past member Jake Stone and Ian Bedford, and current members Ana Jordan, Darrell Bean and Susie Gill. I would like to thank James Canham especially for his friendship and his support.

This work would not have been possible without the help and input of the phytoplasma team. I am grateful to have learned from Dr Pascal Pecher, who supervised me during the first two years of my PhD. Thank you, Pascal, for being patient with me, but also for your positive attitude and energy. I am also very grateful to Dr Weijie Huang, who received the "phytoplasma torch" from Pascal. Thank you very much Weijie, for your great help in the experiments, for helping me develop my critical thinking and keeping me motivated. I also want to acknowledge Gabriele

Moro, Ali-Al-Subhi, Thomas Mathers and Thorsten Langner from Sophien's lab for their help in the interaction assays and phylogenetic analyses. Thank you to Dr Chih-Horng Kuo, Prof. Russell Groves and Agnieszka Zwolińska for providing me new phytoplasma sequences for my PhD.

I want to thank all my British and international friends here. First the British, Tom, Billy, Sam and Jessica. For my international friends, I would like to thank Andre and Marie for their friendship since the beginning of my PhD. Thank you to Philipp, Pawel, Baptiste, Valeria, Matina, Marco, Roger and Estela for enriching my Norwich life.

Last but not least, I would like to thank my parents and my sister for their unwavering support along the years. Thank you for your encouragements, your advices and for always believing in me, even when I doubted myself. I will always be grateful for this.

Table of Contents

Abstract I

Acknowledgements III

List of Figures and Tables VII

List of abbreviations X

Chapter 1 1

General Introduction 1

1.1 Overview 3
1.2 A brief history 4
1.3 Phytoplasma phylogeny 5
1.4 Phytoplasma genomes and genetic features 8
1.5 The biology of phytoplasma association with their insect vectors 13
1.6 Phytoplasma-host interactions 15
1.7 Phytoplasma effectors 20
1.8 The TCP family 24
1.9 Analysis of $SAP11_{AYWB}$ 29
1.10 The role of SAP11 in the phytoplasma invasion 31
1.12 Aims, outlines and outcomes of the PhD thesis 32

Chapter 2 35

Material and Methods 35

2.1 Molecular Cloning of the genes 37
2.2 Yeast two-hybrid analysis 39
2.3 Transient expression assays in *Arabidopsis thaliana* protoplasts 41
2.4 Phylogenetic trees 41

Chapter 3 43

SAP11 effector binding specificity to plant class II TCPs involves the helix-loop-helix region of the conserved TCP domain 43

3.1 Introduction 45
3.2 Results 49
3.3 Discussion 61

Chapter 4 67

The *SAP11* genes of divergent phytoplasmas have a different evolutionary history compared to the rest of the genome 67

4.1 Introduction 69
4.2 Results 74
4.3 Discussion 86

Chapter 5 93

SAP11 effector homologs differentially interact with TCP (sub)classes 93

5.1 Introduction 95

5.2 Results 97
5.3 Discussion 112

Chapter 6 127

SAP effectors from the SAP11$_{AYWB}$ genetic island interact with the TCP family, including Class I TCP (sub)class 127

6.1: Introduction 129
6.2: Results 132
6.3: Discussion 136

Chapter 7: General Discussion 139

7.1 Summary and implication of the key findings 141
7.2 Technical Discussion 143
7.3 The evolution of phytoplasma effectors 144
7.4 The role of the phytoplasma effectors in the promotion of the bacterial virulence and insect vector fitness 149

Appendices 157

Appendix I – Supplemental Tables for Chapter 2 158
Appendix II – Supplemental Figure for Chapter 3 167
Appendix III – Supplemental Figure for Chapter 4 168
Appendix IV – Supplemental Figures for Chapter 5 169
Appendix V – Supplemental Figures for Chapter 6 173
Appendix VI – Published work 175
References 177

List of Figures and Tables

Figure 1.1. The life cycle of aster yellows phytoplasma strain witches' broom (AY-WB)……3

Figure 1.2. The Phylogenetic tree of '*Candidatus* phytoplasma'……7

table 1.1. Comparison of the genetic features between different genus belonging to Mycoplasma, including phytoplasma, and *Escherichia coli*……..10

Figure 1.3: Example of PMUs……12

Figure 1.4: Comparison between the OY-M and AY-WB phytoplasma genomes…….13

Figure 1.5: The different symptoms induced by AYWB phytoplasma strain in *Arabidopsis thaliana*……..14

Figure 1.6: Transmission Electron micrographs showing *Candidatus* Phytoplasma mali in infected *Cuscuta odorata* plant……17

Figure 1.7: AYWB phytoplasma strain gene regulation depending on the host…….19

Figure 1.8: Strategy and main results for the identification of putative phytoplasma effectors……..20

Figure 1.9: Overview of the systemic movement of phytoplasma and effectors in the plant and effector delivery in the phloem……21

Figure 1.10: SAP54 alters floral development by degrading MADS-box transcription factors……..22

Figure 1.11: SAP11 and SAP54 symptoms in *Arabidopsis thaliana*……..23

Figure 1.12: Multiple sequence alignment of the TCP domain of the 24 *Arabidopsis thaliana* TCP proteins……25

Figure 1.13: Origins of the TCPs in different plant families……26

Figure 1.14: Multiple sequence alignment of some SAP11 homologs……..29

Figure 1.15: $SAP11_{AYWB}$ induces leaf crinkling and an increase in stem number in *Arabidopsis*……30

Figure 1.16: SAP11 effector protein is expressed when phytoplasma is delivered in the plant phloem……32

Figure 3.1: Comparison between the $SAP11_{AYWB}$, $SAP11_{MBSP}$ transgenic *A. thaliana* lines and wild type Col-0……..47

Figure 3.2.1: Four SAP11 effector homologs interact with *Arabidopsis thaliana* Class II CYC/TB1 TCPs in yeast……50

Figure 3.2.2: $SAP11_{AYWB}$ interacts with members of Class II CIN-TCPs in yeast……51

Figure 3.2.3: SAP11 binds the TCP domain of TCPs transcription factors in yeast…….53

Figure 3.2.4: Multiple sequence alignment of the TCP domain of *Zea mays* TCP transcription factors……54

Figure 3.2.5: Four SAP11 effector homologs interact with *Zea mays* Class II CYC/TB1 TCPs in yeast……55

Figure 3.2.6: $SAP11_{MBSP}$ do not show interaction with *Zea mays* CII TCP members while three other SAP11 effector homologs do……57

Figure 3.2.7: SAP11 effector homologs interact with the entire helix loop helix motif of the TCP domain of a Class II CYC/TB1-TCP……..58

Figure 3.2.8: The helix-loop-helix motif of the TCP domain does not seem to be sufficient for interaction against SAP11 effector homologs……60

Figure 4.1.1. Phylogenetic tree of '*Candidatus* Phytoplasma', composed of three main clades. ……70

Figure 4.1.2: *SAP11* and *SAP54* effector genes lie within PMU-like regions of the genomes of '*Ca*. P. asteris' phytoplasmas……73

Table 4.2.1: List of the phytoplasma isolates used in my PhD……75-76

Figure 4.2.2: Phytoplasmas with SAP11 effector genes belong to diverse phylogenetic groups……..78

Figure 4.2.3: The SAP11 phylogenetic tree based on the nucleotide sequences is similar to the SAP11 protein-based phylogenetic tree……..80

Figure 4.2.4: Analysis of the Multiple Sequence Alignment of the SAP11 effector homolog sequences……81

Figure 4.2.5: The species-phylogenetic tree of phytoplasma is different from the SAP11-based phylogenetic tree……..84

Figure 4.3.1: The *SAP11* gene appears to form distinct clades across the different phytoplasma isolates……90

Figure 5.2.1: SAP11 homologs interact differentially with the TCP (sub)classes……98

Figure 5.2.2: $SAP11_{SPLL}$ can interact specifically with the TCP domain of Class I TCPs……..101

Figure 5.2.3: the SAP11 effector homologs are sub-localised in the nucleus of *A. thaliana* protoplasts cells……..104

Figure 5.2.4: SAP11 effector homologs differentially destabilise the TCP (sub)classes in *A. thaliana* protoplasts……106

Figure 5.2.5: The intra-region and the TCP binding domain of $SAP11_{AYWB}$ are required for the binding specificity to CIN-TCPs……109

Figure 5.2.6: The intra-region and the TCP binding domain of $SAP11_{SPLL}$ are required for the binding specificity to Class I TCPs……111

Figure 5.3.1: the steps for the Ancestral Sequence Reconstruction……115

Figure 5.3.2: The TCP interaction domain of SAP11. ……119

Figure 5.3.3: SAP11 homolog genes cluster based on their patterns of interaction with the different TCP (sub)classes……122

Figure 6.1.1: The PMU1 and PMU2 of the AY-WB phytoplasma strain……130

Figure 6.2.1: SAP effectors from $SAP11_{AYWB}$ genetic island can interact with TCP, including Class I TCPs……133

Figure 6.2.2: SAP effectors from $SAP11_{AYWB}$ genetic island does not interact with the TCP domain of Class I TCPs, while $SAP11_{SPLL}$ can……135

Figure 7.1: TCP (sub)classes form interaction with multiple transcription factor families……151

Figure 7.2: Phytoplasma promotes plant stem proliferation to increase its niche……152

Figure 7.3: Phytoplasma increases the fitness of leafhopper insect vectors via a set of SAP effectors……155

Appendix I Table 1: Sequence IDs of TCPs from *Zea mays* (Zm), *Arabidopsis thaliana* (At) ……158

Appendix I Table 2: Oligonucleotide sequences (5´ > 3´) for cloning……159

Appendix I Table 3: Synthesised CDS (underlined) flanked by gateway compatible attL1 and attL2 sites……160

Appendix II figure 1: Yeast two-hybrid analysis of phytoplasma $SAP11_{AYWB,}$ $SAP11_{MBSP}$ and *Arabidopsis thaliana* TCP chimaeras of the TCP domain of AtTCP2 and AtTCP18……167

Appendix III figure 1: The phylogenetic tree based on the TCP interaction domains of SAP11 effector homologs is consistent with the SAP11-based phylogenetic tree……168

Appendix IV figure 1: Yeast two-hybrid analysis of phytoplasma SAP11 chimeric versions based on $SAP11_{AYWB}$ and $SAP11_{MBSP}$ and *Arabidopsis thaliana* TCPs……169

Appendix IV figure 2: Further characterisation of the binding specificity region of Class I interactor $SAP11_{SPLL}$……171

Appendix V figure 1: Multiple Sequence Alignment of AY-WB effectors SAP11, SAP56, SAP66, SAP67 and SAP68……173

Appendix V figure 2: SAP effectors from $SAP11_{AYWB}$ genetic island can interact with TCP, including Class I TCPs……174

List of abbreviations

- AAA clade: Asteroplasma, Anaeroplasma and Acholesplasma
- Aux/IAA: AUXIN/INDOLE-3-ACETIC ACID
- bp: base pair
- bHlH: basic helix-loop-helix
- C terminal: carboxy-terminal
- DNA: deoxy-ribonucleic acid
- EAR: Ethylene-responsive element binding factor-associated Amphiphilic Repression
- ETI: effector triggered immunity
- GFP: green fluorescent protein
- HA tag: haemagglutinin tag
- hlh: helix loop helix
- MTF: MADS-box transcription factor
- NCBI: National Centre for Biotechnology Information
- NLS: Nuclear Localisation Signal
- nt: nucleotide
- JAZ: JASMONATE ZIM-DOMAIN PROTEIN
- PAMP: pathogen associated molecular patterns
- PCR: polymerase chain reaction
- PEG: polyethylene glycol
- PTI: PAMP triggered immunity
- PMU: Potential Mobile Unit
- RAD23: RADIATION SENSITIVE23
- RNA: ribonucleic acid
- SAP: Secreted Aster Protein
- SEM clade: Spiroplasma, Entomoplasma and Mycoplasma
- TCP: TEOSINTE BRANCHED1 in maize, CYCLOIDEA in snapdragon and PROLIFERATING CELL FACTORS 1 and 2 in rice
- N terminal: amino terminal

N.B: For phytoplasma isolates abbreviations see **Table 4.2.1** (page 74)

Chapter 1

General Introduction

1.1 Overview

Phytoplasmas are bacterial plant pathogens that reside in the plant phloem (sap stream that transports nutrient to growing tissues) that are transmitted by sap-feeding insects, such as leafhoppers, planthoppers and psyllids (**Figure 1.1**). Phytoplasmas are associated with diseases in more than a thousand plant species from 98 different plant families (Gasparich, 2010). Phytoplasmas are unique pathogens that can perform their invasion and replication in both plants and insects (Sugio et al., 2011b). Phytoplasmas reduce plant growth and yield in both agriculturally essential crops and wild plants and can induce dramatic symptoms, including stunting, yellowing, witches' brooms (proliferation of branches), phyllody (retrograde development of flowers into leaves), virescence (flower tissues that remain green) and hairy roots. With the current global warming issue, crop losses due to phytoplasma may increase in the future, as the phytoplasma insect vectors are expected to expand their geographical ranges.

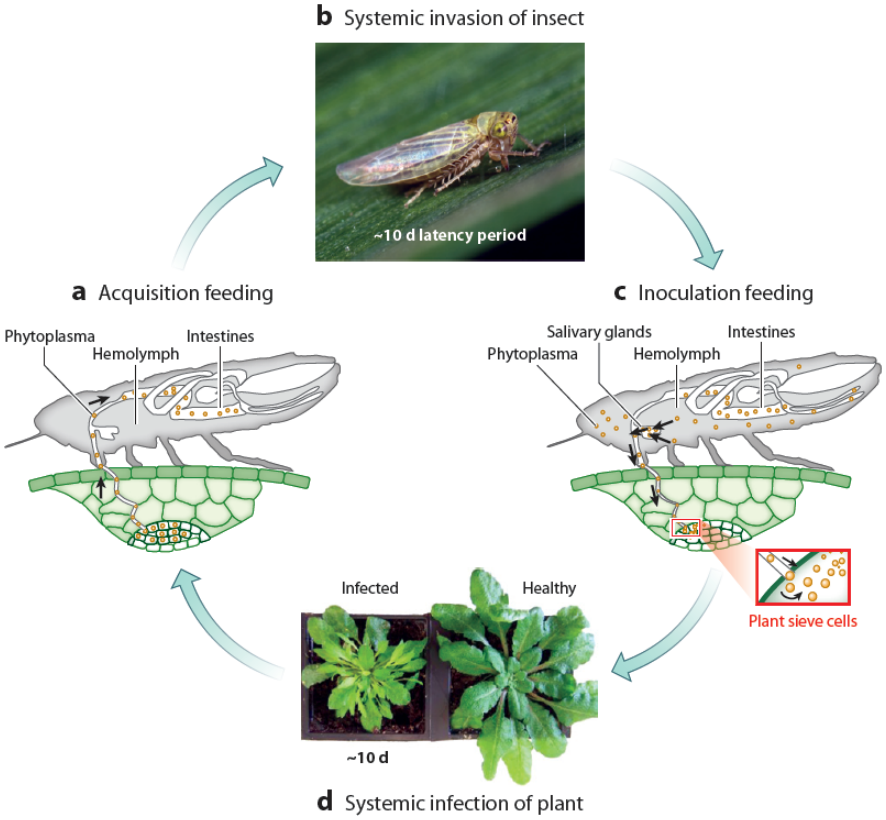


Figure 1.1. The life cycle of aster yellows phytoplasma strain witches' broom (AY-WB). The phytoplasma life cycle comprises two main hosts, the plant and the leafhopper. **A**: the sap-

feeding leafhopper acquires the phytoplasma by feeding from the phloem of an infected plant. **B**: The phytoplasma is colonising the insect. The time it takes to colonise the entire insect is approximately ten days, also referred to as the latency period. **C**: When phytoplasmas reach the salivary glands, the leafhoppers become infectious and can transmit the phytoplasmas to plants when feeding from the phloem. **d:** Phytoplasmas infect plants and can spread systemically throughout the plant via the phloem. Symptoms become apparent at about usually ten days after the first day of exposure to phytoplasma-carrier insect vectors (Figure taken from Sugio et al., 2011).

1.2 A brief history

Since the beginning of the last century, yellows diseases have been reported and described for many plant species (Kunkel, 1926). Plant pathologists and agricultural companies thought phytoplasmas were viruses for half a century. This mistake is due to their small size (they pass through 45-micron sterilisation filters) and the inability to cultivate phytoplasma *in-vitro*. It was only in 1967 that a Japanese team found that plants suffering from yellows disease were infected with Mycoplasma-like bacteria, which are pleomorphic bacteria lacking rigid cell walls that reside exclusively in the phloem (Doi et al., 1967). Since then, the yellows disease agents were referred to as Mycoplasma-Like Organisms (MLOs). MLOs are not cultivable; thus subsequent taxonomic classification of MLOs relied mainly on symptoms induced in affected plant species and insect vector identification (Gasparich, 2010; Errampalli et al., 1991; Wang et al., 2004; Chiykowski, 1990). The further detection, identification and classification of MLOs started along with the development of molecular techniques (Lee et al., 2000). Using classification based on the 16S ribosomal RNA (rRNA) sequence, it became apparent that MLOs are numerous and diverse and belong to a monophyletic group within the Mollicutes in the phylum Tenericutes. They were renamed as phytoplasmas.

In 2004, the International Research Programme for Comparative Mycoplasmology (IRPCM) officially adopted the name '*Candidatus* phytoplasma' (IRPCM, 2004), due to the impossibility to cultivate the bacterium *in-vitro*. The assignment '*Candidatus*' was added because phytoplasmas cannot be cultured, thus

preventing the fulfilment of the Koch's postulates, which are four criteria designed to establish a causative relationship between a microbe and a disease (reviewed in Evans, 1976).

1.3 Phytoplasma phylogeny

Phytoplasmas are Gram-positive derived bacteria that belong to the Firmicutes clade. Phytoplasmas possess a relative low G + C content, along with the rest of the Firmicutes, in comparison with the high G + C content of Gram-positive Actinobacteria (Ventura et al., 2007; Hogenhout and Loria, 2008).

Mollicutes are thought to have diverged from a Clostridium-like ancestor of the class Firmicutes (Gram-positive bacteria) via the loss of their outer cell-wall and a reduction of genome size (Weisburg et al., 1989). The Mollicute group subdivides into two clades: the SEM clade (Spiroplasma, Entomoplasma and Mycoplasma) and the AAA clade (Asteroplasma, Anaeroplasma and Acholesplasma). The phytoplasma genus belongs to the later and is comprised of a large and monophyletic clade paraphyletic to the *Acholeplasma* genus (Gundersen et al., 1994).

1.3.1 16S rRNA classification

The primary marker gene used for phytoplasma phylogeny is the 16S ribosomal RNA (16Sr) gene (IRPCM, 2004; Sugio and Hogenhout, 2012). Carl R. Woese first took advantage of the properties of 16S rRNA and used it as a tool for phylogenetic studies (Woese, 1987). He later argues of the usage of a single marker gene to represent an organism and the parameters to follow for such use. These marker genes need to be involved in key regulatory functions, preferably in the fabric of the cell (Olsen and Woese, 1993). Ribosomes follow such rule as being ubiquitous in every self-replicating cell and being essential in the translation machinery (Green and Noller, 1997). Bacteria comprise three types of ribosomal RNA (5S, 16S and 30S). Phytoplasma genomes contain two rRNA operons which are identical in some phytoplasmas (Schneider and Seemüller, 1994). They are mostly organised in the genome as rRNA operons as follow: 5' – 16S rRNA – spacer region – 23S rRNA – 5S rRNA – trailer region – 3' (Johansson et al., 2002). The 16S rRNA genes fulfil the

conditions established by Carl Woese and remained highly conserved during evolution, making it an excellent tool for phylogenetic analyses. Indeed, the 1500 nucleotide long 16S rRNA molecule is used for the phytoplasma phylogeny as its small size make it easy to sequence and because the presence of universal regions allows confident amplification via PCR (Razin et al., 1998; Johansson et al., 2002). The latter proves very useful, considering the inability to cultivate phytoplasma reliably.

The phytoplasma classification via 16S rRNA sequences comprises two main approaches: the first method relies on phylogenetic analysis of 16S rRNA gene sequences, the second method is based on RFLP analysis of PCR-amplified 16S rRNA gene fragments (Alvarez et al., 2014; Gasparich, 2010). The IRPCM team have established different rules to define a novel phytoplasma species. A phytoplasma is declared as novel species if its 16S rRNA sequence is at least 97.5% different from a previously described phytoplasma species (IRPCM, 2004). However, because of the high number of isolates, this condition can be hard to fulfil; therefore the use of additional molecular markers is necessary to establish the description as a new '*Candidatus* Phytoplasma' species (Seemüller and Schneider, 2004; IRPCM, 2004). Along with the additional molecular marker genes, that will be detailed further, the host range of the phytoplasma strain and the insect species that vector the pathogen also needs to be taken into account before establishing a new phytoplasma species (Seemüller et al., 2002).

Based on the 16S rRNA sequences, the '*Ca.* Phytoplasma' genus is comprised of three main clades, with up to nineteen 16Sr-based groups (**Figure 1.2**) (Chung et al., 2013). The first clade (in blue in **Figure 1.2**) includes the aster yellows (AY) group, belonging to the 16SrI phytoplasma subgroup, that includes *'Ca.* P. asteris' Aster Yellow-Witches' Broom (AYWB) phytoplasma, '*Ca.* P. asteris' Maize Bushy Stunt Phytoplasma (MBSP) (not shown in **Figure 1.2**) and '*Ca.* P. asteris' Onion-Yellow mild-symptom (OY-M) phytoplasma (Lee et al., 2004). The AY group is the largest group of the phytoplasmas, with over 100 isolates described. The 16S rRNA sequences are highly homologous in this group, with the AY group further subdivided into distinct subgroups (16SrIA, 16SrIb and 16SrIC) via RFLP of16S rRNA (Lee et al., 1993, 1998; Jomantiene et al., 1998; Lee et al., 2004).

The second clade comprises mainly the apple proliferation group, or 16SrX, that includes '*Ca.* P. mali' which can infect apple trees (Seemüller and Schneider, 2004). Finally, the third clade is highly diverse, with the likes of the peanut witches'-broom group (16SrII) (Lee et al., 2000; IRPCM, 2004; Chung et al., 2013). The third clade includes '*Ca*. P. aurantifolia' Witches' Broom Disease of Lime (WBDL) and Sweet Potato Little Leaf phytoplasma (SPLL) (Zreik et al., 1995; Gibb et al., 1995).

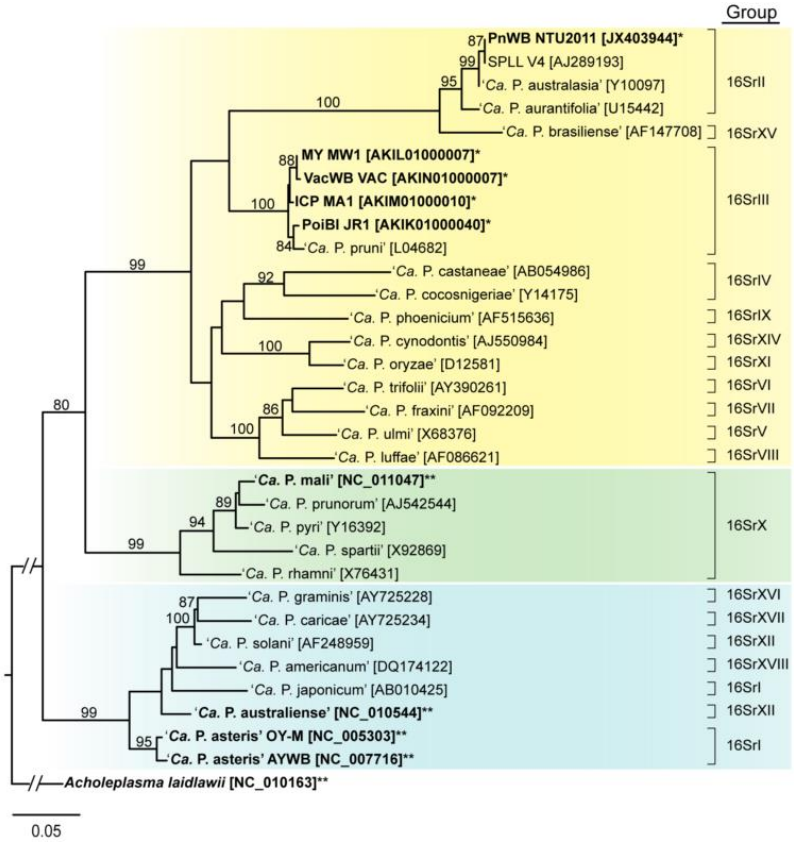


Figure 1.2. The Phylogenetic tree of '*Candidatus* phytoplasma', composed of three main clades. Molecular phylogeny inferred from the 16S ribosomal RNA genes (Figure taken from Chung et al. 2013).

1.3.2 Other marker genes for classification

Other marker genes need to be used to fine-tune the classification of '*Ca.* Phytoplasma' genus. It is especially true when it comes to differentiate closely related strains within the same 16S-based subgroup, as some phytoplasma strains contain two different 16S rRNA sequences within their genome (Liefting et al., 1996; Jomantiene et al., 2002; Davis et al., 2003). This heterogeneity between the two rRNA operons can be high, thus raising the need to find new molecular markers for the

characterisation of the phytoplasma strains (Schneider and Seemüller, 1994; Johansson et al., 2002). The additional marker genes are used to better differentiate phytoplasma strains within the same sub-group (Mitrovic et al., 2011) and/or can be used as comparative phylogenetic analyses via using multiple marker genes (Marcone et al., 2000; Martini et al., 2007; Mitrovic et al., 2011).

For example, ribosomal protein (rp) genes (*rpl22* and *rpS3*) are used to differentiate closely related phytoplasma strains as the *rp* genes mutation rates is relatively higher than the one of the 16S rRNA, allowing a better delineation of the phytoplasma phylogeny (Martini et al., 2007). The *groEL* gene (also known as *cpn60*) encoding for a heat-shock protein and chaperone, is used to differentiate better the AY group 16SrI, as the gene displays lower sequence similarities than 16S rRNA within the three 16SrI subgroups (Mitrovic et al., 2011). The differentiation of some isolates of Maize Bushy Stunt Phytoplasma in Mexico and Brazil has also been achieved via the *cpn60*-based classification (Pérez-López et al., 2016). However, the use of the *groEL* gene is not systematic across the genus as it is sometimes found missing on certain phytoplasma groups like the 16SrIII group (Saccardo et al., 2012). More variable genes, like *secY* (Lee et al., 2004, 2006, 2010; Davis et al., 2013) or *tuf* are also used to fine-tune the classification of phytoplasma groups (Marcone et al., 2000).

1.4 Phytoplasma genomes and genetic features

Eight phytoplasma genomes have been sequenced to completion, namely '*Ca*. P. asteris' OY-M (Oshima et al., 2004), *Ca*. P. asteris AY-WB (Bai et al., 2006), '*Ca*. P. asteris' australiense strain PAa (Tran-Nguyen et al., 2008), '*Ca*. P. mali' strain AT (Kube et al., 2008), '*Ca*. P. asteris' australiense strain SLY (Andersen et al., 2013), '*Ca*. P. asteris' MBSP (Orlovskis et al., 2017), '*Ca*. P. ziziphi' or JWB phytoplasma (jwb-nky) (Wang et al., 2018a) and '*Ca*. P. solani' strain SA-1 (Music et al., 2019). Additionally, multiple teams were unsuccessful in completing the sequencing of phytoplasma genomes, mainly due to the high proportion of multiple-repeat regions. This resulted in the publishing of fifteen phytoplasma genome drafts (Saccardo et al., 2012; Chung et al., 2013; Chen et al., 2014; Kakizawa et al., 2014; Mitrovic et al., 2014; Chang et

al., 2015; Lee et al., 2015; Quaglino et al., 2015; Fischer et al., 2016; Zamorano and Fiore, 2016; Zhu et al., 2017; Sparks et al., 2018). Taken together, the complete and draft genomes provide useful insights into the genetic features of phytoplasma.

1.4.1 Phytoplasma sequencing

Phytoplasma genomes are difficult to sequence for several reasons. First, phytoplasmas are often present in low abundance, providing a challenge to obtain sufficient phytoplasma DNA (Hogenhout and Music, 2010). When the phytoplasma-infected samples are sequenced, the read coverage of the phytoplasma genome is frequently low, making it expensive to obtain sufficient reads. Moreover, many phytoplasma genomes contain large (> 20 kb) repeat-rich regions that are difficult to assemble (Bai *et al.*, 2006; Tran-Nguyen *et al.*, 2008; Chung *et al.*, 2013).

Briefly, the complete sequencing of most of the phytoplasma genomes was performed similarly (Oshima et al., 2004; Bai et al., 2006; Kube et al., 2008; Tran-Nguyen et al., 2008; Hogenhout and Music, 2010). Genomic DNA was extracted from infected plants, either from the phloem (Oshima et al., 2004; Bai et al., 2006; Kube et al., 2008; Andersen et al., 2013; Orlovskis et al., 2017; Music et al., 2019), from the flowers (Tran-Nguyen et al., 2008) or from leafhopper samples (Orlovskis et al., 2017).

1.4.2 Phytoplasma genetic features

First, the phytoplasma genome size is small but considerably varies among the strains, ranging from 600 to 1350 kb (Music et al., 2019). The phytoplasma genome consists of one chromosome with up to several plasmids for the majority of the sequenced phytoplasma (Nishigawa et al., 2002). The phytoplasma genomes display a minimal set of genes involved in the metabolism, similarly to mycoplasmas (Oshima et al., 2004; Bai et al., 2006; Oshima et al., 2013). For example, both mycoplasmas and phytoplasmas have incomplete pathways for the *de novo* synthesis of amino acids and nucleotides and lack genes for synthesis of ATP (ATP synthases) (Oshima et al., 2004; Christensen et al., 2005) (**Table 1.1**). The latter was thought to be a strict requirement of genomes, but its absence in phytoplasma genomes changed our perception of the minimal genes required for an organism (Mushegian

and Koonin, 1996; Oshima et al., 2004). Furthermore, the genome restriction is stronger in phytoplasma than in the rest of the Mollicutes, as they lack components of the phosphotransferase system and the pentose phosphate pathway, required for example for nucleotide synthesis (**Table 1.1**) (Oshima et al., 2004; Christensen et al., 2005).

This restricted number of genes directly involved in the metabolism is consistent with the fact that phytoplasmas are obligate parasites; they rely on the uptake of the nutrients from the host which is consistent with the presence of multiple genes involved in the membrane transport processes, such as ABC transporters for the uptake of peptides and nucleotides. Therefore, the reductive evolution of the phytoplasma is correlated to their nature as obligate parasites (Oshima et al., 2004).

	Phytoplasma	Spiroplasma	Mycoplasma	*E. coli*
Catabolism and production of ATP and NADH/$FADH_2$				
Glycolysis	+[b]	+	+	+
Arginine dihydrolase pathway	−	+	+	+
Oxidative phosphorylation	−	−	−	+
Krebs cycle (tricarboxylic acid cycle)	−	−	−	+
Pentose phosphate pathway	−	−	+	+
ATP synthase	−	+	+	+
Galactose conversion (to make it accessible for glycolysis)	−	−	+	+
Anabolism				
De novo synthesis of purine and pyrimidine nucleotides	−	−	−	+
Salvage pathway for nucleotides	−	+	+	+
Synthesis of amino acids	−	+/−	−	+
Synthesis of fatty acids	−	−	−	+

table 1.1. **Comparison of the genetic features between different genus belonging to Mycoplasma, including phytoplasma, and *Escherichia coli*** (Figure is taken from Christensen et al. 2005).

Comparative genomics between the two '*Ca.* P. asteris' genomes OY-M and AY-WB reveals that both genomes are rich in repeated regions and have a discontinuous GC-skew, which indicates that there is a high degree of recombination and genome instability (Oshima et al., 2004; Bai et al., 2006; Chung et al., 2013; Ku et al., 2013). This degree of recombination can be attributed mainly in the presence of Putative Mobile Units (PMUs) within phytoplasma genomes.

1.4.3 Putative Mobile Units

Among the most striking genetic features in phytoplasma genomes are the Putative Mobile Units (PMUs) (Bai et al., 2006). PMUs are repeat-rich regions that resemble conjugative replicative transposable elements (**Figure 1.3**). The PMUs exhibit characteristic features and can be complete or partial, depending on the presence of their signature genes (Bai et al., 2006; Dickinson, 2010; Toruño et al., 2010). For example, AY-WB phytoplasma carries four PMUs, but only PMU1 is complete because of the flanking of 237-bp inverted repeat regions and the presence of key signature genes (Bai et al., 2006; Dickinson, 2010). Some phytoplasma genomes display one (Kube et al., 2008) to several PMUs (Bai et al., 2006).

PMUs and PMU-like elements contain genes involved in transposition (*tra5*), replication (the helicase *DnaB*, the primase *DnaG* and *tmk* genes), and recombination (single-stranded DNA binding protein (*ssb*), DNA-binding protein HU (*himA*)), regulation (sigma factor *sigF*) and several genes with predicted membrane localization (*hflB* and genes with unknown functions) and putative virulence proteins (effector genes; such as *SAP11* and *SAP54*) (Bai et al., 2006) (**Figure 1.3**). PMU1 is flanked by a 237-bp inverted repeat and is also present as a circular extrachromosomal unit (Toruño et al., 2010) (**Figure 1.3**).

Evidence suggesting that PMUs are responsible for horizontal gene transfer was reported (Chung et al., 2013; Ku et al., 2013; Wang et al., 2018a; Music et al., 2019). Indeed, the characteristic genes that feature PMUs are mostly involved in the replication and the transposition, which prompted the community to hypothesise that these elements are involved in horizontal transfer between phytoplasma strains (Bai et al., 2006; Hogenhout and Music, 2010) (see Introduction of Chapter 4 for more details).

PMUs and PMU-like elements are often clustered together in specific regions of the phytoplasma genomes (**Figure 1.4**). For example, most of the genomes of '*Ca.* P. asteris' AY-WB and '*Ca.* P. asteris' OY-M are syntenic, but some regions that include PMUs show rearrangements (Sugio and Hogenhout, 2012; Chung et al., 2013).

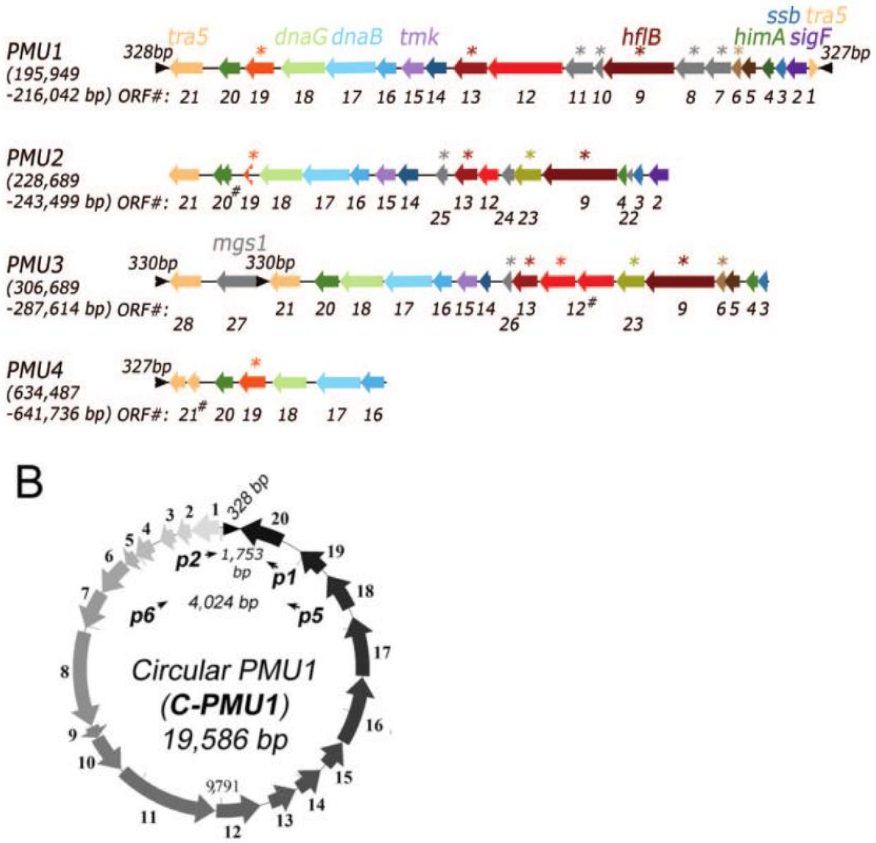


Figure 1.3: Example of PMUs (**A**) Putative Mobile Units of the phytoplasma strain AY-WB. The genome of AY-WB phytoplasma contains four PMUs. Only PMU1 appears fully complete, while the other PMUs are smaller and contain many truncated genes. PMU1 exists as a linear chromosomal form (L-PMU1) and (**B**) an external and circular chromosomal form (C-PMU1). The black line indicates the chromosome, while the block arrows indicate the open reading frame (ORF). Blocks of the same colours indicate paralogous ORFs among the PMUs. The names of the ORFs are displayed when their predicted function are known (Bai et al., 2006). The stars indicate the ORFs of predicted membrane-targeted proteins. The ORF numbers (ORF #) are attributed as based on Bai et al., 2006. (Figure taken from Bai et al., 2006; Toruño et al., 2010)

PMUs are possibly an important factor for generating genome variability and as such, could play a role in the adaptation of the phytoplasma to the environment. Indeed, the PMUs harbour genes encoding for membrane proteins and virulence proteins (effectors). The latter includes SAP11 and SAP54. The majority of AY-WB phytoplasma effector genes (41 out of 56) lies within PMUs (Sugio and Hogenhout, 2012). SAP11 and SAP54 target specific plant processes, as explained further below. Other effectors that are part of PMU coding regions are likely to function in insect

vectors, for example, SAP36, which is a PMU1 gene, and other PMU1 genes are upregulated in phytoplasmas inside insect vectors compared to those inside host plants (Toruño et al., 2010). Phytoplasma genomes also contain several small (4-7 kb) plasmids. For the case of AYWB, 4 circular plasmids have been characterised, containing a total of 22 putative ORFs (Bai et al., 2006).

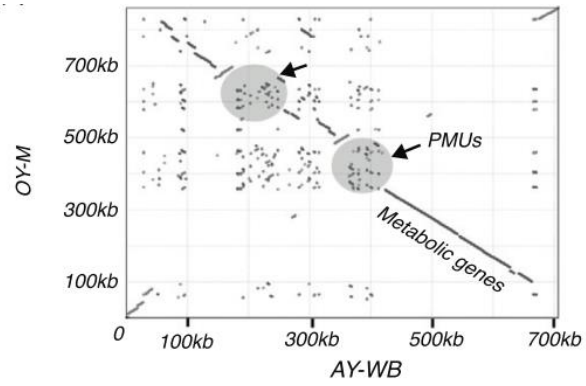


Figure 1.4: Comparison between the OY-M and AY-WB phytoplasma genomes. The majority of the two genomes are syntenic. However, PMU-rich regions show high propensities of rearrangements (Figure taken from Sugio and Hogenhout, 2012).

1.5 The biology of phytoplasma association with their insect vectors

Members of '*Candidatus* Phytoplasma' are pathogens of hundreds of different plant species (Christensen et al., 2005). The geographical distribution of phytoplasmas is determined mainly by the number of insect vector species that can transmit these bacteria (Lee et al., 2000). In addition to its insect vector, phytoplasma can spread via propagative vegetation, cuttings and rhizomes (Gasparich, 2010).

In this present literature review, I will mainly focus on reviewing the literature on AY-WB and maize bushy stunt phytoplasma (MBSP), because these phytoplasmas will be the main subjects of my PhD research among others. AY-WB and MBSP are members of the 16SrI '*Candidatus* Phytoplasma asteris' group (**Figure 1.2**) (Lee et al., 2004). Some of my PhD research will also include work on WBDL (Witches' Broom Disease of Lime) and SPLL (Sweet Potato Little Leaf) phytoplasmas that belong to the 16SrII. The 16SrI and 16SrII groups belong cluster I and II of the phytoplasma phylogeny (**Figure 1.2**) and hence are distantly related to one another.

AY-WB phytoplasma has a broad host range of over 350 plant species, affecting agronomically important field crops, commonly vegetables as well as ornamental plants (Frost et al., 2011). The best examples of host plants are China aster (*Callistephus chinensis Nees*), tomato (*Solanum lycopersicum* L.), lettuce (*Lactuca sativa L.*), *Nicotiana benthamiana*, and *Arabidopsis thaliana*. Twenty four leafhopper species can acquire and deliver aster yellows (AY) phytoplasma, including perhaps AY-WB phytoplasma (Christensen et al., 2005). In the USA, the main insect vector of AY phytoplasmas, including AY-WB, is the aster leafhopper *Macrosteles quadrilineatus* (Forbes). The aster leafhopper is prevalent in Midwest USA (Frost et al., 2011). Consistent with the broad host ranges of AY phytoplasmas, *M. quadrilineatus* is a polyphagous insect species that can feed, lay eggs and shelter on more than 300 plant species. Thus, *M. quadrilineatus* uses these plant species as feeding and/or reproductive hosts (Frost et al., 2011). AY-WB phytoplasmas can replicate and circulate inside *M. quadrilineatus*, which transmits AY-WB to susceptible plants (Bai et al., 2007; Lee et al., 2000). AY-WB infected plants often show witches' brooms, stunting, yellowing and necrosis symptoms (**Figure 1.5**) (Zhang et al., 2004; Sugio et al., 2011b).

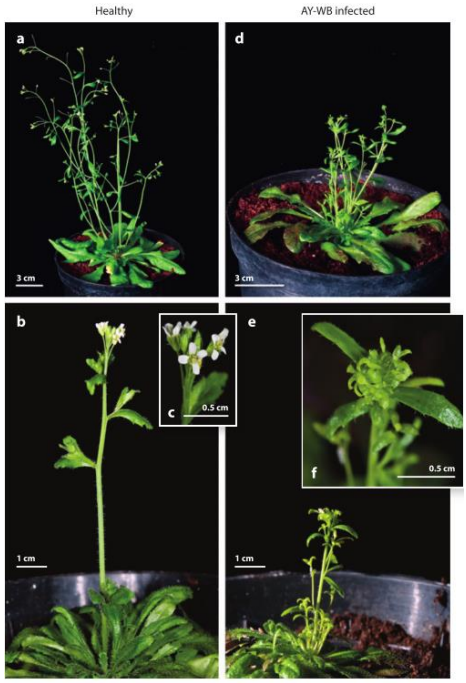


Figure 1.5: The different symptoms induced by AYWB phytoplasma strain in *Arabidopsis thaliana*. (a,b) Healthy plants; (c) Zoom in the flower parts of the (b) picture; (d,e) Infected

plants showing stunting symptoms, increased stem production from the center of the rosette (witches' broom phenotype, (e,f) leafy flower symptoms, (f) zoom of the flower of the (e) picture. Virescence observed on the petals (Figure taken from Sugio, MacLean, et al. 2011)

In contrast to AY-WB phytoplasma, the host range of MBS phytoplasma is restricted to maize (*Zea Maize* L.). MBSP is also a member of the 16SrI group. MBSP was first detected in Mexico in 1955 but is also present throughout south America from south USA, Mexico, Colombia, Brazil, to Peru (Pérez-López et al., 2016). Although the 16S based phylogeny between the isolates from the three countries defines them as Maize Bushy Stunt Phytoplasma (IRPCM, 2004), a recent study has shown that they differ based on *groEL* sequence differences (Pérez-López et al., 2016; Dumonceaux et al., 2014).

Symptoms of MBSP-infected maize plants include chlorotic stripes, stunting, stem proliferation (generating the bushy appearance) and yellowing of the leaves (Pérez-López et al., 2016). The two main insect vectors of MBSP are the corn leafhopper *Dalbulus maidis* and the Mexican corn leafhopper *Dalbulus elimatus*. *D. maidis* is found exclusively in subtropical and tropical areas of America and is a specialist insect pest of maize (Capinera, 2008).

MBSP, *D. maidis* and *D. elimatus* are thought to have co-evolved with maize since its domestication from teosinte. MBSP and the insect vectors are also thought to have co-evolved, because *D. maidis* and *D. elimatus* do not die from MBSP infections. In contrast, other members of the *Dalbulus* genus, such as *D. gelbus* and *D. quinquenotatus*, which live predominantly on gamagrass and generally do not feed from maize (and therefore do not get exposed to MBSP), die from MBSP infection (Nault 1980; Pérez-López *et al.*, 2016).

1.6 Phytoplasma-host interactions

Phytoplasmas are unique pathogens that can invade and replicate in both plants and animals (insects) (Sugio et al., 2011b). Phytoplasmas are directly transmitted to plants by phloem-feeding insects such as planthoppers, psyllids and

leafhoppers. In this review, I will focus on leafhoppers, which are the main vectors of AY-WB and MBSP.

1.6.1 Phytoplasma in leafhoppers

Previous studies have shown that phytoplasmas are located in many leafhopper organs (see next paragraph). Approaches to study the localization of phytoplasmas within leafhopper hosts are (1) fluorescent in-situ hybridization (FISH) techniques, using, for example, biotin-labelled oligonucleotide probes that hybridized to phytoplasma DNA (Webb et al., 1999); (2) immunofluorescence microscopy using a fluorescently labelled antibody to an abundant cell-surface membrane protein of the phytoplasma, such as the Antigenic Membrane Protein (Amp) (Suzuki et al., 2006).

The acquisition of the phytoplasma occurs during the feeding of the insect from the phloem of an infected plant (see **Figure 1.1**). *M. quadrilineatus* can acquire AY-WB from infected plants within 4 days of feeding (Sugio et al., 2011a). When ingested, phytoplasmas move from the mouthparts (stylets) into the lumen of the intestines and attach to the membranes of the midgut epithelial cells and invade this organ. Both AY-WB and MBSP can multiply in the midgut epithelial cells, especially in the muscle fibres that form the outer layers of the midgut (Suzuki et al., 2006). Then, phytoplasmas enter into the haemolymph and make their way to the salivary glands (Webb et al., 1999; Hogenhout et al., 2008). Phytoplasmas colonise the salivary glands of the insect and multiply in the cytoplasm of the salivary gland cells (Christensen et al., 2005; Hogenhout et al., 2008). From these cells, phytoplasmas move into the salivary canals and are delivered into the plant phloem as soon as leafhopper stylets reach the phloem sieve cells (Hogenhout et al., 2008). Once the leafhopper acquires the phytoplasmas it remains carrier during its lifetime (Christensen et al., 2005).

1.6.2 Phytoplasma in plants

When infecting the plant, phytoplasmas are mainly located in the phloem (**Fig. 1.6**). When the leafhopper is feeding on the plant host, phytoplasmas are inoculated via the saliva of the insect into the pierced sieve element. From the sieve,

phytoplasma is spreading in the plant via the sieve tube system and accumulates mainly in the phloem elements (mature sieve tubes as well as immature phloem cells), in the companion cells and in the phloem parenchyma cells (**Fig. 1.6**) (Zimmermann et al., 2015; Pagliari et al., 2016). Furthermore, AY-WB phytoplasma strain can also accumulate in sink areas, like shoots and roots, consistent with the characteristic symptom of witche's broom (Christensen et al., 2005; Hogenhout et al., 2008). As phytoplasma lacks genes coding for flagella or cytoskeleton elements, it is translocated passively along with the assimilate flow (Oshima et al., 2013).

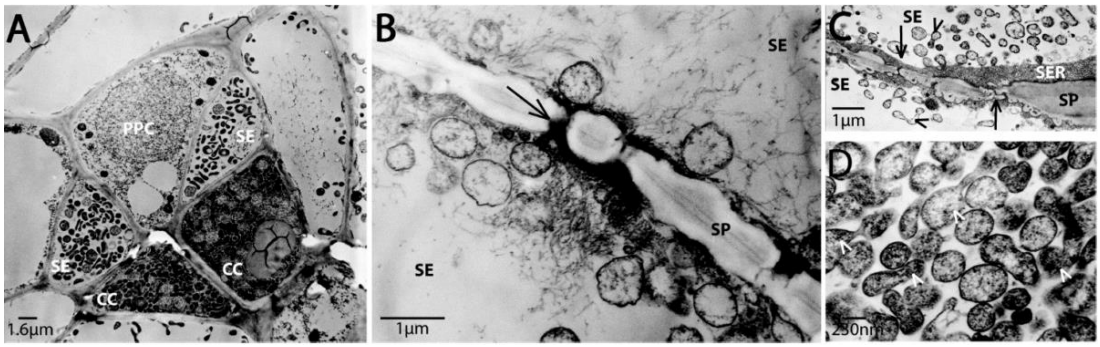


Figure 1.6: **Transmission Electron micrographs showing *Candidatus* Phytoplasma mali in infected *Cuscuta odorata* plant.** (A) Phytoplasma localized in sieve elements (SE), companion cells (CC) and phloem parenchyma cells (PPC) separated by a sieve plate. (b) Two sieve elements & lateral sieve pores. (c) (Figure taken from Zimmermann et al. 2015).

Phloem impairment caused by the phytoplasma accumulation leads to characteristic symptoms such as leaf yellowing, leaf curling, vein necrosis and growth stunting (Lee et al., 2000; Musetti et al., 2013). The phytoplasma-infected phloem shows necrosis, plasmolysis or collapse (Zimmermann et al., 2015; Pagliari et al., 2016). In *Arabidopsis thaliana*, the phytoplasma infection leads to a general alteration in shape and morphology of the phloem components, including phloem hyperplasia and an increase in callose deposition in the phloem cells compared to healthy vascular tissues (Pagliari et al., 2016). Furthermore, the phytoplasma-infected sieve elements are altered, with the sieve pores narrowed or obstructed by callose deposition or agglutination of phloem-protein, the latter often found to envelope phytoplasma cells (Pagliari et al., 2016).

These numerous alterations in the vascular tissues are proof of its importance as a contact structure between the phytoplasma cells and the plant host. In both tomato and *Arabidopsis thaliana*, phytoplasma forms an "adhesion structure" with sieve element components (Buxa et al., 2015; Musetti et al., 2016; Pagliari et al., 2016). In tomato, phytoplasma can adhere to the sieve-element reticulum and to the sieve element plasma membrane (Buxa et al., 2015; Musetti et al., 2016). The "adhesion structure" is a 30-40nm wide tubular connection orientated perpendicularly to the plasma membrane of the sieve element (Pagliari et al., 2016). Upon adhesion to the sieve element plasma membrane, phytoplasma can remobilise the host-actin to facilitate the passage of the pathogen through the sieve pores (Buxa et al., 2015). This is possible via phytoplasma membrane proteins, such as the Immunodominant Membrane Protein (Imp) that can bind to the plant actin (Boonrod et al., 2012). With the lack of phytoplasma genes coding for movement, Pagliari et al. and Boonrod et al. have speculated that the phytoplasma binding to host-actin helps the pathogen for spreading and colonisation through the plant (Boonrod et al., 2012; Pagliari et al., 2016).

The phloem mediates the translocation of metabolites and may, therefore, harbour enough nutrients for the phytoplasma. The sieve elements of phytoplasma-infected phloem are sometimes plugged with callose that is likely a plant defence response to limit systematic movement of phytoplasma (Zimmermann et al., 2015; Pagliari et al., 2016).

1.6.3 Phytoplasma migration between plants and insect vectors

Phytoplasma has a unique relationship with its two hosts. Phytoplasmas often have clear negative impacts on the fitness of the host plants, while it can have a positive impact on the fitness of the insect vector. Indeed, *M. quadrilineatus* leafhoppers survive longer and have increased fertility rates on phytoplasma-infected plants (Beanland et al., 2000; Sugio et al., 2011a). Studies have shown that phytoplasma genes are differentially expressed depending on the host (MacLean et al., 2011; Makarova et al., 2015; Pacifico et al., 2015). For instance, 33% of OY-M phytoplasma genes change during the host switching between the plant and the

insect (Oshima et al., 2011). Approaches using qRT-PCR can be performed to analyse which phytoplasma genes are upregulated or downregulated in the insect vector and plant host (**Figure 1.7**) (Makarova et al., 2015; Pacifico et al., 2015).

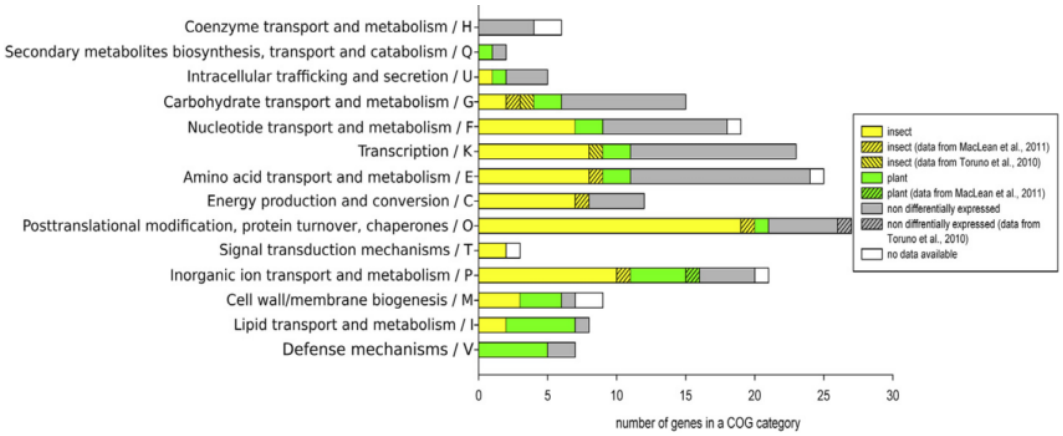


Figure 1.7: AYWB phytoplasma strain gene regulation depending on the host (plant or insect) (Figure taken from Makarova *et al.*, 2015).

The differential gene expression of phytoplasma upon host switching enable adjustments to the extreme environments of insect vectors and plant hosts (Makarova et al., 2015). The environmental conditions trigger a stress-response in phytoplasma. Hence, phytoplasma response involves the upregulation of cytoplasmic co-chaperones such as GroEL or GroES and proteases. Phytoplasmas also modulate processes of plant hosts and insect vectors via the secretion of virulence proteins (effectors) and extracellular membrane-associated proteins (also named SAMPs). The AY-WB phytoplasma genome encodes 20 SAMP. Among these 20, the antigenic AMP protein interacts specifically with proteins of insect vectors, but not with those of non-vectors (Suzuki et al., 2006; Galetto et al. 2011). Nonetheless, the Amp protein expression level of phytoplasmas is three times higher in plant hosts than in insects vectors, suggesting that this protein may also have a role in plants (Makarova et al., 2015). Indeed, Amp protein interacts with the α and β subunits of ATP synthase and actin of both insect vector and plant hosts (Suzuki et al., 2006; Galetto et al., 2011; Makarova et al., 2015).

1.7 Phytoplasma effectors

1.7.1 The Phytoplasma Secretion System

Bacteria possess different types of secretion systems to deliver their virulence proteins into the plant host. In Gram-negative bacteria, the effectors are delivered via the type III secretion system (T3SS), a highly conserved apparatus. The type-III effectors are injected directly into the cytoplasm (intra-cellular effectors) or the extracellular matrix (extra-cellular effectors) via this system (Guttman et al., 2002).

Phytoplasmas have evolved from a Gram-positive bacterium and therefore do not possess type-III secretion systems to inject their virulence proteins in the host. Instead, phytoplasmas rely on two secretion pathways: the YidC system for the integration of surface-exposed membrane proteins; and the Sec-dependent system for the secretion and the delivery of the proteins into the host cell cytoplasm (Oshima et al., 2013; Kakizawa et al., 2004). For example, the Amp protein previously described is secreted by the Sec-dependent pathway (Kakizawa et al., 2004).

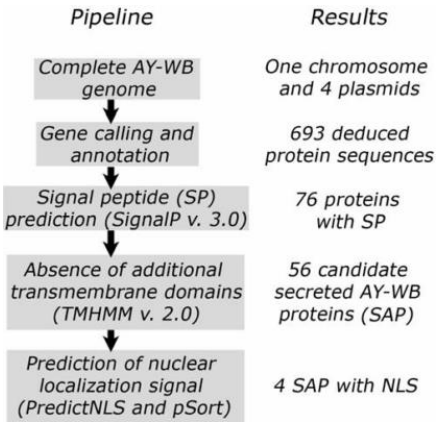


Figure. 1.8: Strategy and main results for the identification of putative phytoplasma effectors (Figure taken from Bai *et al.*, 2009).

1.7.2 Phytoplasma effectors: general description

Effectors are proteins secreted by pathogens and translocate into cells of the host to exert specific functions there (Alfano, 2009). Plant pathogens, such as bacteria, fungi, oomycetes and nematodes can use effectors to suppress host immune responses, such as PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI). Thus, effectors can be considered as pathogen virulence factors,

allowing pathogens to multiply, leading to disease symptoms (Munkvold and Martin, 2009).

Research on phytoplasma effectors began to bloom as soon as phytoplasma genomes became available. Phytoplasmas effectors possess an N-terminal signal peptide (SP), which is required for secretion via the Sec-dependent pathway. Mining of the AY-WB genome identified 56 encoding proteins with signal peptides (Bai et al., 2009) (**Figure 1.8**). These 56 secreted AYWB proteins (SAPs) are candidate virulence proteins or effectors and are delivered directly into the cytoplasm of the plant or the insect cells (Bai et al., 2009; Hoshi et al., 2009). SAPs do not possess transmembrane regions after cleavage of the SP and are predicted to be soluble in the host cellular environment. Among these 56 SAP proteins, the majority appear to be encoded within PMUs (34 out of 56 SAPs) or on plasmids (7 SAPs) (Bai et al., 2009). Both the SAP11 and SAP54 effector genes lie within PMU elements (Bai et al., 2006).

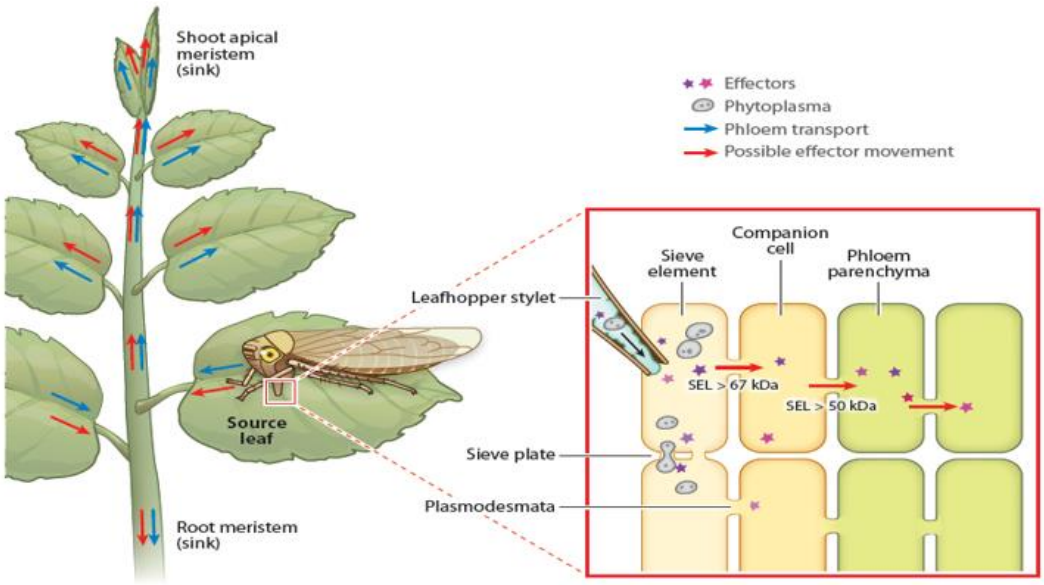


Figure 1.9: Overview of the systemic movement of phytoplasma and effectors in the plant and effector delivery in the phloem. SEL indicates the size exclusion limit of plasmodesmatas (Figure taken from Sugio et al., 2011).

Phytoplasma effectors SAP11, SAP54 and Tengu-su target plant developmental processes, whereas SAP36 most likely plays a role in the insect vector (Sugio and Hogenhout, 2012). Phytoplasma effector genes are differentially expressed upon phytoplasma host switching between the insect vector and plant

host, indicating that different effectors alter plant and insect processes (MacLean et al., 2011; Pacifico et al., 2015; Makarova et al., 2015). Whereas phytoplasmas are restricted to the phloem sieve cells, the effectors can migrate from the phloem sieve cells into adjacent cells and systemically via the sieve elements (**Figure 1.9**).

Tengu-su inducer (TENGU) effector is a small protein (4.5kDa), identified in the OY-M phytoplasma strain (Hoshi et al., 2009). TENGU induces developmental alterations such as witche's broom symptoms and dwarfism. Although phytoplasma is restricted to the phloem, TENGU can be transported from the phloem to other tissues such as apical buds (Hoshi et al., 2009). Another study has shown that the N-terminal domain of TENGU is responsible for inducing symptoms (Sugawara et al., 2013). However, a plant target of TENGU has not been identified.

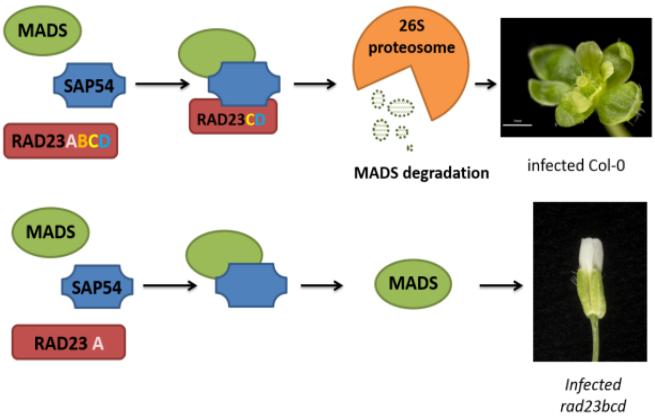


Figure 1.10: **SAP54 alters floral development by degrading MADS-box transcription factors.** (Figure adopted from Zigmunds Orlovskis)

SAP54 is an effector that modulates flower development and floral organ identity in plants (MacLean et al. 2011). Flowers of SAP54 over-expression lines show indeterminate flower phenotypes, leafy sepals and green stamens (**Figure 1.10,** and **Figure 1.11**) (MacLean et al., 2011). SAP54 degrades specific members of the MADS-domain transcription factor (MTF) family, which includes key regulators of floral development. MTFs have four distinct domains. SAP54 interacts specifically with one of these domains, the keratin-like (K) domain of MTFs. MTFs form tetramers and the K domain is involved in the specificity of MTF-MTF interactions. Animals also have MTFs, but these lack the K domain. Therefore, SAP54 may have evolved to selectively target plant MTFs (MacLean et al., 2014). SAP54-mediated degradation of MTFs

requires the 26S proteasome shuttle factor RAD23 (Farmer et al., 2010; MacLean et al., 2014) (**Figure 1.10**). Leafhoppers prefer to colonise phytoplasma-infected and SAP54 transgenic plants. Moreover, they prefer to colonise phytoplasma-infected *rad23cd* plants that show leaf-like flowers to phytoplasma-infected *rad23bcd* plants that have normal flowers (MacLean et al., 2014). Thus, rad23 also plays a role in leafhopper preference.

Recently, it was found that the stable production of SAP54 in *A. thaliana* leads to an increase in the attraction of *M. quadrilineatus*, independently of the presence of leaf-like flowers (Orlovskis and Hogenhout, 2016). This year, the crystal structure of SAP54 has been resolved and consists of two alpha-helices connected by a random loop in a coiled-coil manner (Iwabuchi et al., 2019).

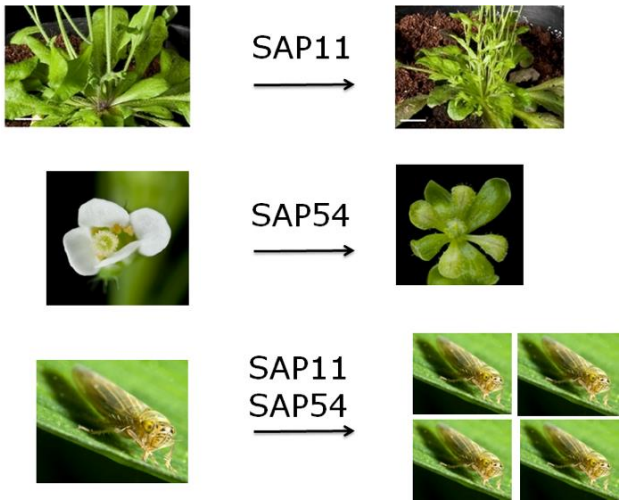


Figure 1.11: SAP11 and SAP54 symptoms in *Arabidopsis thaliana*. Plants and insects shown at left are healthy and those at right are infected with phytoplasma. SAP11 induces witches' brooms and SAP54 leaf-like flowers (phyllody). Both effectors promote leafhopper colonisation (Photos taken from Sugio et al, 2011; Maclean et al, 2011).

SAP11 is an effector that induces witches' brooms symptoms and crinkled leaves in *A. thaliana* (Sugio et al., 2011a) (**Figure 1.11**). SAP11 destabilises specific members of the TEOSINTE BRANCHED1, CYCLOIDEA, PROLIFERATING CELL FACTORS 1 and 2 (TCP) transcription factor family, particularly CIN-TCPs and CYC/TB1-TCPs

(sub)classes, which belong to class II (Sugio et al., 2011a). Before analysing $SAP11_{AYWB}$, an introduction of the plant TCP transcription factors is required.

1.8 The TCP family

1.8.1 Overview of the TCP family

TCP proteins, named after TEOSINTE BRANCHED1 in maize, CYCLOIDEA in snapdragon and PROLIFERATING CELL FACTORS 1 and 2 in rice (Cubas et al., 1999) are a plant-specific transcription factors (TFs) family. The TCP family is characterized and subdivided based on the highly conserved TCP domain. The TCP domain is the functional domain of the TCP protein, involved in its dimerisation and the DNA binding to the target genes (Kosugi and Ohashi, 1997, 2002; Aggarwal et al., 2010). The TCP domain is predicted to form a basic-helix-loop-helix (bhlh) structure, however it is not classified as a canonical bhlh TF as it does not bind to the same DNA elements (Cubas et al., 1999; Aggarwal et al., 2010). Based on the alignment of the TCP domain of multiple species, the TCPs are divided into two classes: the class I PCF-type (PCF clade) and the class II, the later which is subdivided into class II CINCINNATA (CIN clade) and class II CYC/TB1 (known as ECE clade). The main differences between the two classes of TCPs lie in the absence of 4 amino acids in the basic motif of the TCP domain of the class I and the presence of motifs outside the TCP domain, such as the ECE motif and/or the arginine-rich R domain for members of the class II TCPs (Howarth and Donoghue, 2005).

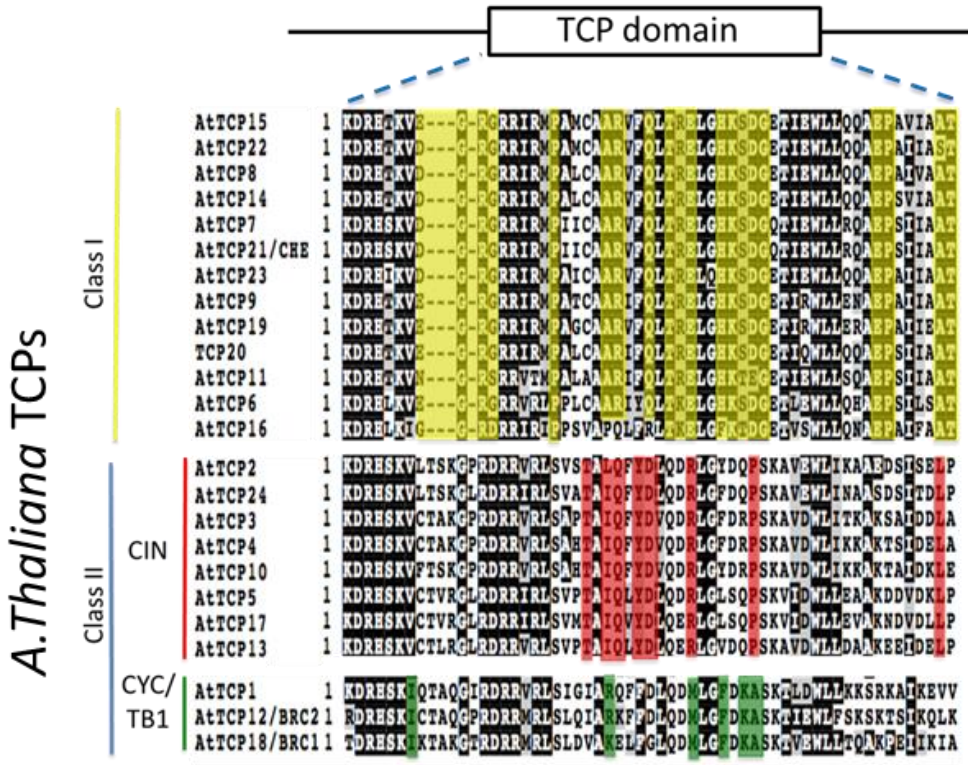


Figure 1.12: Multiple sequence alignment of the TCP domain of the 24 *Arabidopsis thaliana* TCP proteins. The coloured residues indicate the specific conserved residues for each class. The residues highlighted in yellow are specific to class I TCPs. The residues highlighted in red are specific to (sub)class II CIN-TCP, while the residues highlighted in green are specific to (sub)class II CYC/TB1-TCP. A version of this figure is available in our latest paper (Pecher et al., 2019).

1.8.2 Origins of the TCPs

The first written record of the role of the TCP gene family, particularly the CYC/TB1 TCP subclass, dates back centuries ago, in the book of Carl Linnaeus and Daniel Rudberg (Linnaeus and Rudberg, 1744). Carl Linnaeus found a mutant of common toadflax (*Linaria vulgaris*) that did not show the regular zygomorphic floral symmetry but instead showed radial symmetry. Two hundred fifty years later, Enrico Coen and his team isolated the CYCLOIDEA gene responsible for the regulation of zygomorphic flowers (Luo et al., 1996).

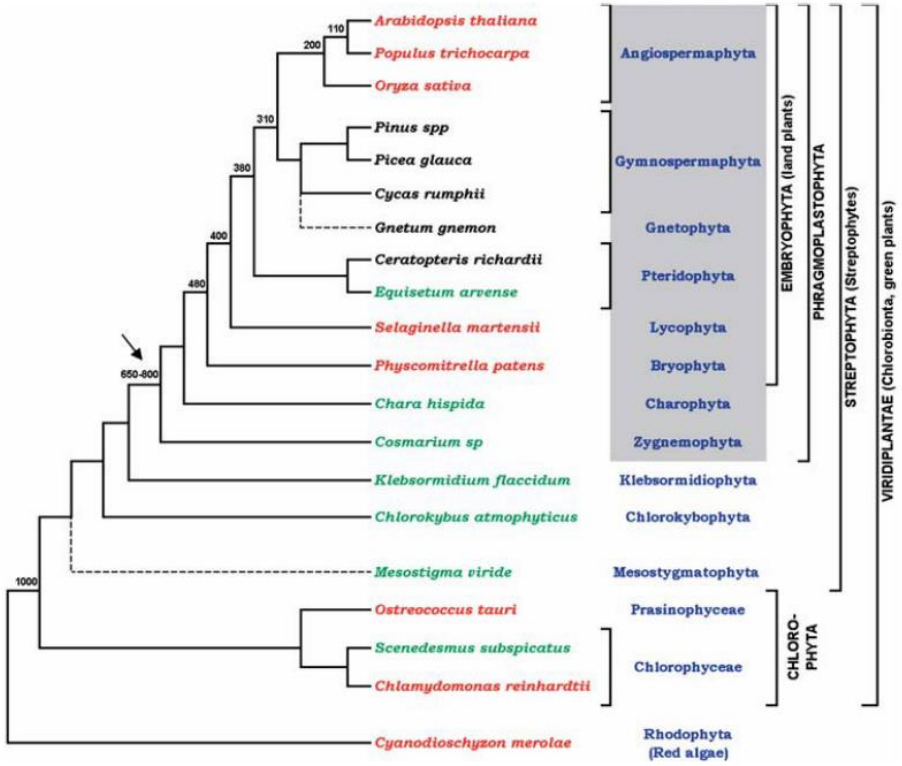


Figure 1.13: Origins of the TCPs in different plant families. The grey box indicates the presence of the TCPs (Navaud et al., 2007). As mentioned in the study, the tree is generated using whole-genome data (red), ESTs (black) or CODEHOP (green). Figure taken from Navaud et al., 2007.

Christine Hervé and her colleagues were the first team to establish the origin of the TCP gene family. Using the two consensus sequences of both Class I (TCP-P) and Class II TCPs (TCP-C), they found that the TCP gene family dates back before the divergence of the Zygnemophyta at a period estimated between 650 and 800 million Years ago (Navaud et al., 2007; Yoon et al., 2004). The TCP gene family existed before the emergence of the land plants, with TCP genes present in freshwater green algae (Charophyta) (Navaud et al., 2007).

The TCP genes are present in many various plant families, both in Angiosperms and Gymnosperms. For example, the *Arabidopsis thaliana* genome contains 24 TCP genes, while the *Zea mays* genome contains 44 TCP genes. In ancient plant lineages, the TCP gene family comprises a few members but along the evolution, the TCP gene family expanded via a series of duplications (Navaud et al., 2007). In *Arabidopsis thaliana*, four TCP genes duplicated early (50Mya) and five TCP

genes duplicated around 200 Mya (Blanc et al., 2000, 2003; Bowers et al., 2003; Navaud et al., 2007; Paterson et al., 2000; Simillion et al., 2002).

1.8.3 The role of TCP (sub)classes

The TCPs are involved in a plethora of key developmental and defence processes (Lopez et al., 2015; Nicolas and Cubas, 2015) such as flower and petal asymmetry (Luo et al., 1996, 1999), plant architecture (Doebley et al., 1997; Aguilar-Martinez et al., 2007), leaf morphogenesis (Nath et al., 2003; Palatnik et al., 2003; Ori et al., 2007; Sarvepalli and Nath, 2011) senescence (Schommer et al., 2008), seed germination (Resentini et al., 2015), embryo growth (Tatematsu et al., 2008), hormone regulation (Danisman et al., 2012; Lopez et al., 2015; Wang et al., 2015b). The TCP family is at the cross-road of the regulation of the plant, and as such is involved in cross-family interaction with other prominent plant transcription factor families (see General Discussion) (Dhaka et al., 2017). The TCP family is also at the heart of the microbe-plant interaction and are targeted by diverse plant pathogens (see General Discussion) (Mukhtar et al., 2011; Weßling et al., 2014). Their role in the plant defence (Kim et al., 2014; Lopez et al., 2015; Wang et al., 2015b; Spears et al., 2019) (see Chapter 5 and General Discussion) contributes significantly to this global targeting by multiple effectors.

Class I TCPs are positive regulators of cell division and are involved in multiple development processes such as leaf and flower development, seed germination and meristem formation (Martín-Trillo and Cubas, 2010; Nicolas and Cubas, 2016; Dhaka et al., 2017). Class II CIN-TCPs are mainly involved in the leaf and petal morphogenesis and maturation (Efroni et al., 2008; Schommer et al., 2008). *A. thaliana* CIN-TCP members are targeted by micro-RNA *miRNA319* (AtTCP2, AtTCP3, AtTCP4, AtTCP10 and TCP24) and are called *jaw*-TCP genes (JAGGED AND WAVY (JAW-D) mutants)) while the rest of the CIN-TCP (sub)class are called TCP5-like TCPs (AtTCP5, AtTCP13, AtTCP17) (Palatnik et al., 2003; Ori et al., 2007; Schommer et al., 2008, 2014; Bresso et al., 2017). Class II CYC/TB1-TCPs are mainly involved in (i) the negative regulation of the axillary branches with the TB1 homologs of *Arabidopsis thaliana* BRANCHED1

(BRC1) and BRC2 (Aguilar-Martinez et al., 2007), and (ii) the control of floral symmetry with the CYCLOIDEA TCPs (Luo et al., 1996, 1999).

Class I TCP and class II CIN-TCPs have antagonistic roles in the plant development (Kosugi and Ohashi, 2002; Li et al., 2005; Martín-Trillo and Cubas, 2010; Danisman et al., 2012). This antagonism is possible because of the distinct but overlapping consensus sequences between the two TCP (sub)classes (Kosugi and Ohashi, 2002; Li et al., 2005; Uberti Manassero et al., 2013). For example, class II CIN-TCPs promote the leaf cell maturation and differentiation, while class I TCPs promote leaf cell proliferation. (Efroni et al., 2008; Schommer et al., 2008; Nicolas and Cubas, 2015). Class I TCPs are genetically redundant, and *Arabidopsis thaliana* single mutant phenotypes do not exhibit substantial changes compared to the wild-type (Aguilar-Martínez and Sinha, 2013). However, when a class I TCP member is fused with a dominant-negative repressor domain EAR, dramatic changes occur in the phenotype (Hervé et al., 2009; Viola et al., 2011). As an example, the dominant-negative forms of AtTCP7 and AtTCP23 show alterations in lateral organ growth, indicating that class I TCPs regulate cell proliferation (Aguilar-Martínez and Sinha, 2013; Hervé et al., 2009).

The antagonistic role of class I and class II CIN TCPs is also reflected in the regulation of the Jasmonic acid hormone (Lopez et al., 2015; Nicolas and Cubas, 2016). In addition to promoting the leaf senescence, JA is an essential hormone in the plant defence against herbivores and necrotrophic pathogens (Reymond and Farmer, 1998; Li et al., 2001). The class II CIN-TCP4 positively regulates the biosynthesis of JA via the regulation of the *LOX2* gene (Bell et al., 1995; Schommer et al., 2008, 2014). On the other hand, class I TCP members TCP9, TCP20 and TCP14 were shown to bind to the *LOX2* promoter, resulting in the downregulation of the JA pathway (Danisman et al., 2012; Yang et al., 2017). The production of JA leads to senescence, which is consistent with the role of class II CIN-TCPs. Indeed, along with the development of the leaf, the abundance of CIN-TCP4 increases over-time, which will promote the maturation of the leaf. In parallel, the JA production will increase, which will aid the transition from cell proliferation to cell expansion, limiting cell proliferation (Palatnik et al., 2003; Li et al., 2005; Efroni et al., 2008; Danisman et al.,

2012). Following this model, the class I TCP genes are expressed at early stage of the leaf development, promoting the cell proliferation and limiting the production of JA during the early stage, until the balance tips to the other side, along with the maturation of the leaf (Li et al., 2005; Danisman et al., 2012).

1.9 Analysis of SAP11$_{AYWB}$

SAP11$_{AYWB}$ is a 14kDa long protein that possesses a bipartite nuclear localization signal (NLS) required for SAP11 targeting of plant cell nuclei (Bai et al., 2009). Other phytoplasmas possess different SAP11 effector homologs (**Figure 1.14**) (see Chapter 4 for more details). SAP11 sequences have NLS and sequences that were previously shown to be involved in TCP-binding (Sugio et al., 2014).

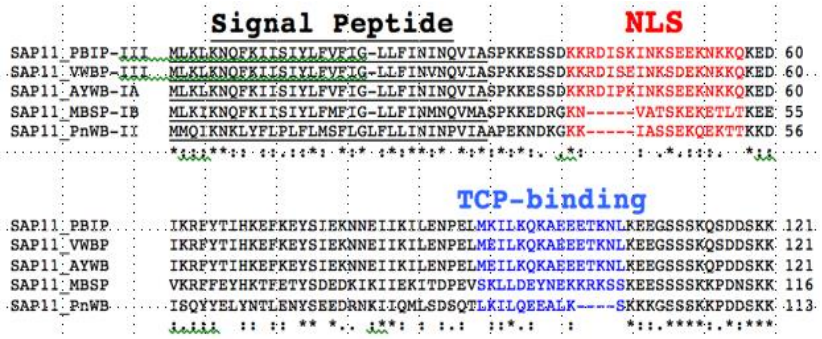


Figure 1.14: Multiple sequence alignment of some SAP11 homologs. SAP11 homologs of poinsettia-branch-inducing phytoplasma (PBIP), vaccinium-witches' broom phytoplasma (VWBP), AY-WB, MBSP and peanut-witches' broom phytoplasma (PnWB). The Nuclear Localisation Signal (NLS) of SAP11$_{AYWB}$ is indicated in red. The TCP-binding domain of SAP11$_{AYWB}$ is indicated in blue (see Chapter 4 for more details). This figure can be found in Sugio et al., 2014.

35S::SAP11$_{AYWB}$ *A. thaliana* lines show: (1) leaf crinkling phenotypes, similarly to 35S::*miR319*a x *miR-3TCP A. thaliana* lines (miR319a and miR-3TCP negatively regulate all 8 CIN-TCPs) (Schommer et al., 2008) and (2) increased stem proliferation, similarly to the *A. thaliana brc1*x*brc2* mutant (BRC1 and BRC2 are CYC/TB1 TCPs) (Aguilar-Martinez et al., 2007), confirming that SAP11$_{AYWB}$ is able to destabilize both CIN-TCPs and CYC/TB1 TCPs (**Figure 1.15**) (for more details, see Chapter 3).

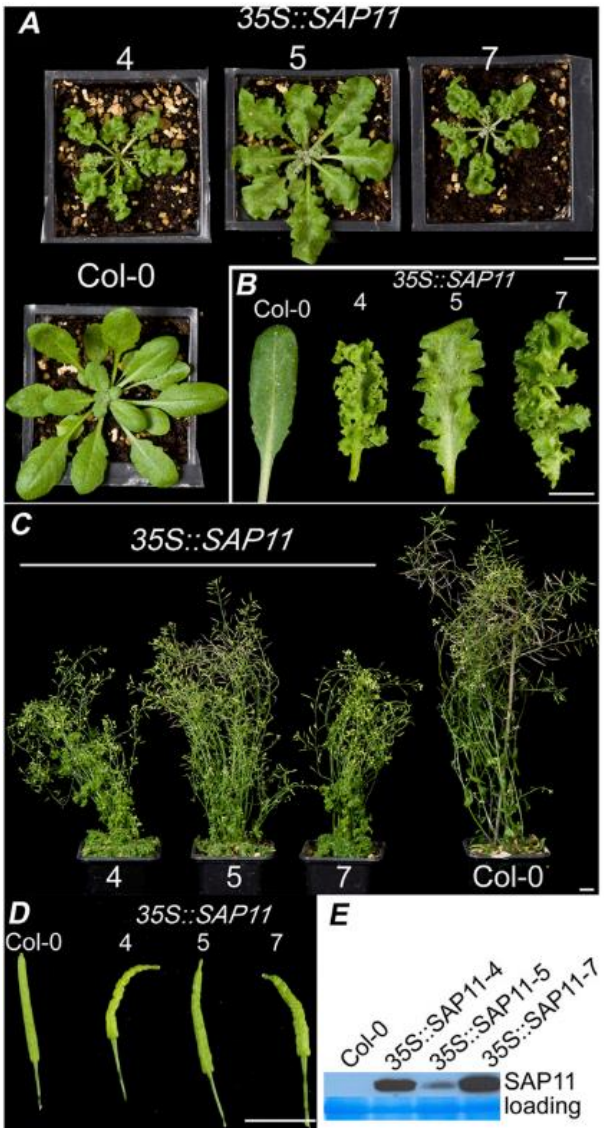


Figure 1.15: **SAP11$_{AYWB}$ induces leaf crinkling and an increase in stem number in *Arabidopsis.*** (**A)** Six-week-old homozygous *35S::SAP11$_{AYWB}$* lines 4, 5 and 7 display crinkled leaves compared to Col-0 ecotype. (**B**) Comparison between ten-week-old Col-0 leaves and transgenic leaves. (**C**) Homozygous ten-week-old *35S::SAP11* lines 4, 5, and 7 display more stems compared to Col-0 ecotype. (**D**) Homozygous ten-week-old *35S::SAP11* lines 4, 5, and 7 display crinkled siliques compared to Col-0 ecotype. (**E**) Western blot of protein extracts of *35S::SAP11* lines 4, 5 and 7 using anti-SAP11. The loading control is the Coomassie-stained Ribulose-1,5-bisphosphate carboxylase oxygenase large subunit. Figure and legend taken from Sugio et al., 2011. *PNAS.*

1.10 The role of SAP11 in the phytoplasma invasion

Consistent with the role of CIN-TCPs and CYC/TB1 TCPs and SAP11-mediated destabilisation of these TCPs, $SAP11_{AYWB}$-transgenic plants show crinkled leaves (cells do not mature, but still proliferate) and increased stem production (resembling witch's broom symptoms of phytoplasma-infected plants (**Figure 1.15**)).

Interestingly, CIN-TCPs also regulate JA signalling via the regulation of the *LIPOXYGENASE2* (LOX2) gene (Schommer et al., 2008, 2014). LOX2 is a lipoxygenase that mediates the first step of the JA synthesis via the conversion of α-linolenic acid (α-LeA; 18:3) into 13-hydroperoxy-9,11,15-octadecatrienoic acid (13-HPOT) (Turner et al., 2002). *LOX2* expression and JA synthesis are downregulated in SAP11 transgenic and AY-WB - infected *A. thaliana*. AY-WB phytoplasma insect vectors (*M. quadrilineatus* leafhoppers) lay more eggs on LOX2-knockdown plants and $SAP11_{AYWB}$ transgenic lines (Sugio et al., 2011a).

Based on these data it was proposed that the SAP11-mediated destabilisation of CIN-TCPs leads to downregulation of LOX2, and the subsequent decrease in JA synthesis, which in turn promotes leafhopper reproduction (**Figure 1.16**). Insects born on phytoplasma-infected plants acquire the phytoplasma and transmit these bacteria to other plants (**Figure 1.16**). Thus, $SAP11_{AYWB}$ is thought to promote phytoplasma spread in the environment by leafhoppers.

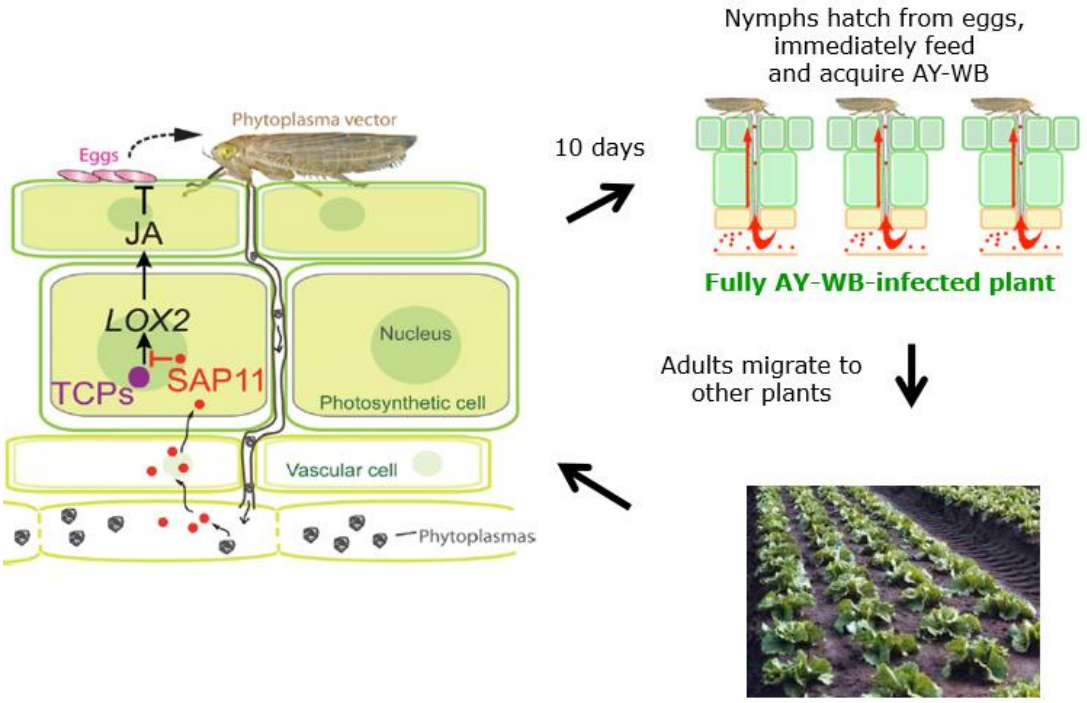


Figure 1.16: SAP11 effector protein is expressed when phytoplasma is delivered in the plant phloem. SAP11 destabilises TCP transcription factors, which lead to the downregulation of the LOX2 gene, jasmonic acid (JA) synthesis and promotion of leafhopper colonisation (Figure taken from Sugio et al., 2011).

SAP11$_{AYWB}$ has also been found to regulate miRNA implicated in phosphate and auxin-signalling in Arabidopsis (Lu et al., 2014). Although it does not change the expression of miR319, the stable expression of SAP11$_{AYWB}$ leads to an accumulation of miR399 and miR827 which positively regulate the phosphate intake and translocation (Lu et al., 2014).

1.12 Aims, outlines and outcomes of the PhD thesis

When I started my PhD, preliminary data suggested that SAP11 effector homologs from divergent phytoplasmas interact with TCP transcription factors in yeast two-hybrid assays. In addition to studies on SAP11 from AYWB, SAP11 from MBSP, a maize specialist, was shown to interact with class II CYC/TB1-TCPs, although it was not clear if the effector bound the other (sub)classes. The lab also cloned SAP11 homologs of WBDL and SPLL and preliminary data showed evidence that these may bind TCPs too.

Beyond these four SAP11 homologs ($SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{SPLL}$ and $SAP11_{WBDL}$), other SAP11 homologs were not yet characterized in the lab. Hence, it was unclear if other SAP11 homologs were similar or different in their TCP interactions compared to $SAP11_{AYWB}$. Considering the diverse roles of TCP (sub)classes in the regulation of both plant development and defence, it was deemed important to assess how the SAP11 effectors evolved to interact differentially with these plant transcription factors.

In summary, the overall **aim of my thesis** was to determine the plethora of interactions the SAP11 effector homologs may have with the members of the three TCP (sub)classes. Based on preliminary data provided to me at the start of my PhD, the **general hypothesis** was that the SAP11 effector homologs have different binding specificities for members of the three TCP (sub)classes. If so, this could lead to distinct outcomes in how phytoplasmas and their insect vectors perform on plant species.

In Chapter 3, I followed on preliminary data to show the binding specificities of the four SAP11 homologs that were already cloned in the lab, i.e. $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{SPLL}$ and $SAP11_{WBDL}$, for selected members of class II CIN-TCPs and TB1/CYC TCPs. I found that all four SAP11 homologs interact with *A. thaliana* and maize TB1/CYC TCPs and that only $SAP11_{AYWB}$ bound CIN-TCPs. I also identified that the SAP11s interact with the conserved TCP domain and that within this domain, the helix-loop-helix region determined binding specificity for the SAP11s. This work was partly included in the Pecher et al., 2019. PLoS Pathogens paper on which I am a co-author and that was put online on Sep 2019.

In Chapter 4, I built on the data of Chapter 3, to identify more SAP11 homologs from divergent clades within the Phytoplasma phylogeny and used them to generate phylogenetic analyses. Comparisons of the phylogenies of the phytoplasma 16S rDNA and SAP11 protein and nucleotide sequences elucidated that different SAP11 effector homologs from various phytoplasmas have a different evolutionary history compared to the rest of the genome. Interestingly, alignments of the SAP11 protein sequences revealed subregions, including those involved in

nuclear targeting and TCP binding, that may have diverged among the SAP11 homologs. In fact, the SAP11 proteins grouped in distinct clades.

In Chapter 5, I then used the SAP11 sequence analyses of Chapter 4 to further dissect how SAP11 homologs interact with TCPs. I found that the regions within SAP11 proteins are involved in mediating differential binding specificities for members of Class I and II TCPs. Intriguingly, the SAP11 clades align with TCP binding specificity. This was further confirmed by generating chimaeras of the SAP11s and of the TCP domains and using these in Y2H assays.

In Chapter 6, I made the surprising discovery that, in fact, the other candidate effectors for which the genes that lie on the SAP11-island in the genome of AYWB also interact with TCPs, but unlike $SAP11_{AYWB}$ these other effectors interact also with class I TCPs.

Altogether, I have been able to further dissect SAP11-TCP binding specificities and showed that the SAP11 effector family has evolved to bind a diversity of members from different TCP (sub)classes. In the general discussion (Chapter 7), I will discuss the implications of my findings. This includes, for example, how my research can lead to solid predictions of what SAP11 homologs may target.

Chapter 2

Material and Methods

2.1 Molecular Cloning of the genes

2.1.1 Molecular Cloning of the genes used in Yeast Two-Hybrid

All the DNA constructs were maintained in *Escherichia coli* as described in Sugio et al., (2014) and Pecher et al., (2019). The cloning of all the TCP genes and SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ were achieved before the start of my PhD by Dr Akiko Sugio, Dr Pascal Pecher and Dr Ali Al-Subhi (Sugio et al., 2014; Pecher et al., 2019). Most of the constructs were already available in expression vectors suitable for Y2H assays. For *SAP56*, *SAP66*, *SAP67* and *SAP68*, the genes were available in pDONR207 (see 2.1.3).

The TCP domain of *A. thaliana* class I TCP members AtTCP6, AtTCP8, AtTCP9, AtTCP14 and AtTCP15 and the helix-loop-helix motif of *A. thaliana* AtTCP2 and AtTCP18 were amplified by PCR from pGADT7 Y2H expression vector carrying the full-sequence of the corresponding genes (see 2.1.3). The amplification of the target sequences was performed by PCR using specific primers, which includes full-length attB adapter primers (see list of primers in **Appendix I – Table 2**). The PCR product was run on EtBr-stained 1% agarose gel to cut out the expected size band and then purified using QIAquick Gel Extraction Kit (QIAGEN). The purified PCR product was then cloned into Gateway-compatible donor vectors, pDONR207 via the BP reaction. After the BP reaction, the constructs were transformed into thermocompetent *Escherichia Coli* (DH5α) and plated on selective LB agar media. Plasmids from positive colonies were purified via the QIAprep Spin Miniprep Kit (QIAGEN) and sequenced.

After the BP reaction, the constructs were transformed into thermocompetent *Escherichia coli* (DH5α) and plated on selective LB agar media. Plasmids from positive colonies were purified via the QIAprep Spin Miniprep Kit (QIAGEN) and sequenced.

2.1.2 Synthesis of the SAP11 and TCP constructs used in yeast two-hybrid

The nine TCP chimaeras and the TCP domain of AtTCP2 and AtTCP18 used in Chapter 3 (**Figure 3.2.7**) were synthesised by Genscript (New Jersey, USA) into the pMS vector by Dr Cristina Canale, as described in Pecher et al., (2019). The SAP11

chimaeras used in Chapter 5 (**Figure 5.2.5**): "SAP11-Chimaera 1", "SAP11-Chimaera 2", "SAP11-Chimaera 3" and "SAP11-Chimaera 4" were synthesised by Genscript (New Jersey, USA) into the pMK vector by Dr Cristina Canale, as described in Pecher et al., (2019). The SAP11 effector homologs used in Chapter 5 (**Figure 5.2.1**), "SAP11-Chimaera 5", "SAP11-Chimaera 6" (**Figure 5.2.5**) and the SAP11 chimaeras "SAP11-Chimaera A" and "SAP11-Chimaera B" (**Figure 5.2.6**) were synthesised by General Biosystems (North Carolina, USA) into the pUC57 vector.

The synthesised sequences were flanked with attachment sites attB1 and attB2, compatible with the LR reaction. The signal peptide of the SAP11 genes and SAP11 chimaera genes were removed. **Appendix I – Table 3** displays the sequences of each synthesised construct detailed here.

2.1.3 Cloning into Y2H expression vectors

The synthesised *SAP11* genes (SAP11 effector homologs and SAP11 chimaeras) (2.1.2)*, SAP56*, *SAP66*, *SAP67* and *SAP68* were cloned via the Gateway Cloning System® using the LR reaction method (Gateway LR clonase Enzyme Mix, Invitrogen, P/N 56484) into pDEST-GBKT7 (binding domain). The synthesised TCP genes and the *TCP* mutant genes in pDONR207 (2.1.1) were cloned similarly into pDEST-GADT7 (activation domain) vectors. The SAP11 effector homologs, the SAP effectors SAP56, SAP66, SAP67 and SAP68 have been amplified without their predicted signal peptide.

2.1.3 Molecular Cloning of the genes used in the protoplast degradation assays

The four *SAP11* effector homolog genes ($SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$) and *TCP* genes were cloned from the Gateway compatible entry clones into the expression vectors with the Gateway LR Clonase II enzyme mix (Invitrogen) before I started my PhD. The cloning was performed as stated in Pecher et al., 2019.

Briefly, the full-length ORFs of *A.thaliana TCPs* were cloned into pUGW15 (Nakagawa et al., 2009) to produce N-terminally HA-tagged proteins. The codon-optimised versions of $SAP11_{AYWB}$ and $SAP11_{MBSP}$ and the non-codon-optimized versions of $SAP11_{SPLL}$ and $SAP11_{WBDL}$ were cloned into pUBN-GFP-DEST without their

signal peptide sequences (Pecher et al., 2019). I also cloned the same four SAP11 effector homologs into pUGW18 to generate N-terminally 4xMyc-tagged SAP11 effector proteins (Nakagawa et al., 2009). The cloning was performed as stated in Pecher et al., (2019).

2.2 Yeast two-hybrid analysis

2.2.1 Yeast transformation

Saccharomyces cerevisiae yeast strain AH109 (Clontech®), auxotrophic for leucine and tryptophan, was streaked out on a 1X YPAD (yeast extract-peptone-adenine-dextrose) agar media and grown for three days at 28°C. A single yeast colony was inoculated in 250 mL of 1X YPAD liquid media and incubated overnight at 28°C. On the next day, the concentration of the yeast cells is measured, with an optimal concentration of 2 x 10^7 cells ml^{-1}. Yeast cells were pelleted via centrifugation at 4000g for 5 minutes. The YPAD media is removed, and the pellet is washed two times with sterile water. After the third centrifugation, the pellet is resuspended with a sterile solution of 1X TE (1M Tris-HCl and 0.1M EDTA; pH 7.5) and 1X LiOAc (10M Lithium Acetate; pH 7.5). Boiled single-strand carrier salmon sperm DNA (Invitrogen) is then added in the yeast solution (100mL of Salmon sperm for 1000mL of yeast solution). The yeast transformation with the different combinations of pDEST_GADT7_TCPx and pDEST_GBKT7_SAP11x is performed in single tubes. Twenty microlitres of the yeast solution is mixed with 0.4µg of the two plasmids. One-hundred and thirty microlitres of PEG solution (50% (w/v) polyethylene glycol/1XTE/1XLiOAc) solution is then added and mixed up and down. The tubes are then put in incubation 28°C with constant mixing. The yeast solutions are then put in 42°C for 15 minutes. The tubes are then centrifuged at 8000g for 30 seconds, then the resulting pellet is resuspended in 500µL sterile water.

2.2.2 Screening for yeast transformants

Each of the yeast solution was plated in single Synthetic Defined (SD) selective medium -LW (leucine and tryptophan dropout) plates and incubated at 28°C for 3 days. SD -LW media stands as the positive control of the yeast two hybrid experiment,

as the growth of the yeast depends on the uptake of the two plasmids. The pGADT7 plasmid will provide the gene involved in the synthesis of leucine while pGBKT7 provides the gene that synthesises tryptophan amino acid. Two colonies of each co-transformant are then re-streaked in a fresh SD-LW plate and incubated overnight at 28°C. A moderate amount of a single yeast streak (around 1mm wide) was used to inoculate 3ml of sterile SD-LW liquid media. The culture was incubated overnight at 28°C.

2.2.3 Screening for protein-protein interactions via yeast-two hybrid

The next day, the O.D of each co-transformant culture was measured and set up to OD=1 via appropriate dilution with sterile water. This way, each yeast solution contains the same quantity of yeast cell, allowing a viable comparison once the solutions are plated on the different selective media. Two-hundred microlitres of diluted solution of co-transformed yeast is prepared and 5µL is dropped on the following selective medium (1) SD-LW media, (2) SD-LWH (leucine, tryptophan and histidine dropout), (3) SD-LWH + 5mM 3AT (3-amino-1,2,4-triazol), (4) SD-LWH + 20mM 3AT, (5) SD-LWAH (leucine, tryptophan, histidine and adenine dropout). The pGADT7 plasmid encodes a fusion protein of GAL4 gene activation domain (AD) and TCP protein, while pGBKT7 plasmid encodes a fusion protein of GAL4 gene binding domain (BD) and SAP11 protein. Upon the interaction between SAP effectors (SAP11, SAP11 chimaeras, SAP56, SAP66, SAP67 or SAP68) and TCP (full TCP, TCP domain, TCP motifs or TCP chimaeras), the AD and BD will activate the expression of the Histidine *His3* gene, allowing the yeast to grow in SD-LWH media. 3AT is a competitive inhibitor of the His3 gene and will inhibit low levels of *His3* expression, thus selecting the stronger interactions between SAP11 protein and TCP protein. A stronger interaction allows the AD and BD to activate the ADE2 gene, that encodes the AIR-carboxylase, which is involved in the purine biosynthetic pathway in yeast (Gedvilaite and Sasnauskas, 1994).

2.3 Transient expression assays in *Arabidopsis thaliana* protoplasts

Isolation and transformation of Arabidopsis protoplasts were performed as described by (Yoo et al., 2007). Protoplasts were generated from 5-week-old Arabidopsis in controlled environmental conditions with a 14h, 22 C°/ 10h, 20°C light / dark period. Six-hundred microlitres of protoplast-suspensions were transformed with the indicated constructs and placed in the dark over-night for gene expression (Pecher et al., 2019). The protoplast preparation and transformation was done as described in our latest paper (Pecher et al., 2019). For detection of Myc-tagged SAP11 effector proteins, Monoclonal Anti-c-Myc antibody (produced in mouse, M5546, Sigma).

The next day, the transformation efficiency of the protoplasts was assessed via fluorescent microscopy (Leica DM6000 microscope). The proportion of transformed fluorescent protoplasts that produce GFP-tagged SAP11 proteins were counted and compared to non-fluorescent protoplasts using the GFP channel of the DM6000 microscope (Pecher et al., 2019).

2.4 Phylogenetic trees

The phylogenetic trees were generated based on either 16SrRNA sequences or SAP11 sequences (nucleotide or amino acid sequences) from the genomes of phytoplasma isolates that have been found to express *SAP11* effector homolog genes. The first is a phylogenetic tree based on the SAP11 sequences of phytoplasmas. The second is a phylogenetic tree based on the 16S ribosomal DNA (rDNA) of the same phytoplasma strains used in the SAP11 tree.

2.4.1 Generation of the 16S rDNA-based phylogenetic tree

The 16S rDNA sequences of all the phytoplasma strains that have SAP11 effector were gathered from the NBCI website, or from collaborators (see Chapter 4 for more details). Then, the 16S rDNA sequences have been aligned using MUSCLE algorithm with default parameters (Edgar, 2004) via the MEGA (Molecular Evolutionary Genetics Analysis) software (version 7) (Kumar et al., 2008). The

resulting Multiple Sequence Alignment (MSA) was used to generate a phylogenetic tree via the Maximum Likelihood algorithm, default parameters (Guindon and Gascuel, 2003). Five thousand bootstrap samples were generated to evaluate the level of support of each branch of the phylogenetic tree. The resulting tree were formatted and annotated in *FigTree* software *v1.4.3* (www.tree.bio.ed.ac.uk/software/figtree/). Then, I coloured and finalised the phylogenetic tree using Inkscape.

2.4.2 Generation of the SAP11-based phylogenetic trees

The SAP11 nucleotide sequences and amino acid sequences were aligned separately using MUSCLE algorithm with default parameters (Edgar, 2004) via the MEGA (Molecular Evolutionary Genetics Analysis) software (version 7) (Kumar et al., 2008). Then, the resulting Multiple Sequence Alignment (MSA) of either the nucleotide sequences or the amino acid sequences were used to generate two phylogenetic trees via the Maximum Likelihood algorithm (Guindon and Gascuel, 2003). One thousand bootstrap samples were generated to evaluate the level of support of each branch of the phylogenetic tree. The resulting tree were formatted and annotated in *FigTree* software*v1.4.3* (www.tree.bio.ed.ac.uk/software/figtree/). Then, I coloured and finalised the phylogenetic trees using Inkscape.

Chapter 3

SAP11 effector binding specificity to plant class II TCPs involves the helix-loop-helix region of the conserved TCP domain

Parts of this chapter were published in:

Pecher, P., Moro, G., Canale, M. C., Capdevielle, S., Singh, A., MacLean, A., … Hogenhout, S. A. (2019). Phytoplasma SAP11 effector destabilization of TCP transcription factors differentially impact development and defence of Arabidopsis versus maize. *PLoS Pathogens*, in press . See **Appendix VI**.

3.1 Introduction

Phytoplasmas have evolved from Gram-positive bacteria via loss of outer cell wall and genome reductions, and rely on sap-feeding insects, such as leafhoppers, planthoppers and psyllids, for transmission (Weintraub and Beanland, 2006). Phytoplasmas are considerably diverse, with over 1000 phytoplasma isolates identified so far (Gasparich, 2010). When the sap-feeding insects deliver phytoplasmas inside the phloem sieve cells of the plant, these bacteria multiply and cause dramatic symptoms, such as virescence (flower organs remaining green), phyllody (production of leaf-like flowers) and witches' brooms (increase of the axillary branches) (Bertaccini, 2007; Hogenhout et al., 2008; Sugio and Hogenhout, 2012; Sugio et al., 2011a; MacLean et al., 2014). Phytoplasmas trigger these considerable changes in plant development through secreting effector proteins directly into the cytoplasm of plant sieve cells in which both the phytoplasma and effectors migrate throughout the phloem and the effectors also unload from the phloem to adjacent tissues (Bai et al., 2009; Hoshi et al., 2009; MacLean et al., 2011; Sugio et al., 2011a; Chang et al., 2018; Pecher et al., 2019).

Aster Yellows strain Witches' Broom (AY-WB; a 16S rDNA group IA (16srIA)) phytoplasma and assigned to '*Candidatus* (*Ca.*) Phytoplasma asteris' (Bai et al., 2006) is one of the most investigated phytoplasmas and several effectors named secreted AY-WB proteins (SAPs), have been described from this bacterium. AY-WB phytoplasma infects a broad range of dicotyledonous plants, including *Arabidopsis thaliana*, and is transmitted mainly by the polyphagous leafhopper species *Macrosteles quadrilineatus* (Hogenhout et al., 2008). Among the AY-WB effectors, $SAP11_{AYWB}$ binds and destabilises TCP transcription factors (see Chapter 1 for more information) (Sugio et al., 2011a, 2014). Previously, the Hogenhout lab found that $SAP11_{AYWB}$ binds and destabilises the class II TCP (sub)classes CIN-TCP and CYC/TB1-TCPs, leading to changes in leaf shape and stem proliferation, and a decrease of jasmonic acid (JA) production; the latter promotes the fertility of *M. quadrilineatus* (Sugio et al., 2011a, 2014; Pecher et al., 2019). Furthermore, we found that $SAP11_{AYWB}$ interacts with the TCP domain of class II CIN-TCPs (Sugio et al., 2014).

The TCP domain is the conserved and functional domain of TCP transcription factors. It is predicted to form a basic-helix-loop-helix (bHLH) structure, similarly to the bHLH transcription factors; however TCP and bHLH transcription factors have distinct specificities and do not bind to the same DNA elements (Cubas et al., 1999; Aggarwal et al., 2010). The TCP domain is required for the dimerisation of TCP transcription factors in such a manner that the helix-loop-helix domains from the TCP-TCP interaction leading to the exposures of the beta-sheets, which then can bind the promoters of target genes (Kosugi and Ohashi, 1997, 2002; Aggarwal et al., 2010). Therefore, the basic motif of the TCP domain is responsible for the TCP binding to the DNA target (Aggarwal et al., 2010), whereas the helix loop helix motif allows the TCP to form homo and heterodimers (Aggarwal et al., 2010).

The Hogenhout lab recently sequenced the genome of Maize Bushy Stunt Phytoplasma (MBSP; a 16S rDNA group 1B (16srIB) phytoplasma and assigned to *Ca.* Phytoplasma asteris (Lee et al., 2004; Orlovskis et al., 2017). Whereas MBSP also belongs to *Ca*. Phytoplasma asteris, like AY-WB, MBSP has a narrow plant host range, being a maize specialist and transmitted by the maize-feeding leafhoppers *Dalbulus maidis* and *D. elimatus* (Nault, 1980). The symptoms of MBSP-infected maize (*Z. mays. L*) include leaf reddening and chlorosis, stunting, the formation of lateral branches and loss of ear development (Orlovskis et al., 2017). When I started my PhD, there was some evidence that $SAP11_{MBSP}$ also interacts with TCP transcription factors, though the specificity of binding was not yet clear. An alignment of the 44 maize TCPs shows that the TCP domain is subdivided into the (sub)classes PCF (Class I) and CIN and CYC/TB1-TCPs (Class II), similarly to *A. thaliana* TCP transcription factors (Martín-Trillo and Cubas, 2010; Pecher et al., 2019). However, some *Z. mays* TCP transcription factor did not clearly belong to either the CIN or CYC/TB1 subclades, as their TCP domains harbour residues of both (sub)classes; these maize TCPs are assigned to a separate group, the CII-TCPs (Pecher et al., 2019).

When I arrived in the lab, there were already phenotype data available for *A. thaliana* plants that ectopically express the $SAP11_{AYWB}$ and $SAP11_{MBSP}$ under control of the 35S promoter; the 35S:$SAP11_{AYWB}$ and 35S:$SAP11_{MBSP}$ *A. thaliana* plants display

distinct phenotypes (**Figure 3.1**) (Pecher et al., 2019). Firstly, the 35S:SAP11$_{AYWB}$ *A. thaliana* plants have crinkled leaves, indicating that leaf cells excessively proliferate, and phenocopying the phenotype of the *35S::miR319a x 35S::miR3TCP* line in which the expression of all 8 *CIN*-TCP genes are knocked down (Efroni et al., 2008; Pecher et al., 2019). However, the leaves of 35S:SAP11$_{MBSP}$ *A. thaliana* plants resemble wild type leaves and do not show crinkling (**Figure 3.1**) (Pecher et al., 2019). Secondly, both *35S:SAP11$_{AYWB}$* and *35S:SAP11$_{MBSP}$ A. thaliana* display increased stem production (resembling witch's broom symptoms of phytoplasma-infected plants), similarly to the *brc1-2 brc2-1* (*brc1 brc2*) (Col-0) *A. thaliana* line, which is a double null mutant for the CYC/TB1 TCP genes *AtTCP18* (brc1) and *AtTCP12* (brc2) (Aguilar-Martinez et al., 2007; Pecher et al., 2019). These data suggested that SAP11$_{AYWB}$ targets all class II (CYC/TB1 and CIN-TCPs) for destabilisation, whereas SAP11$_{MBSP}$ targets only CYC/TB1 TCPs for destabilisation.

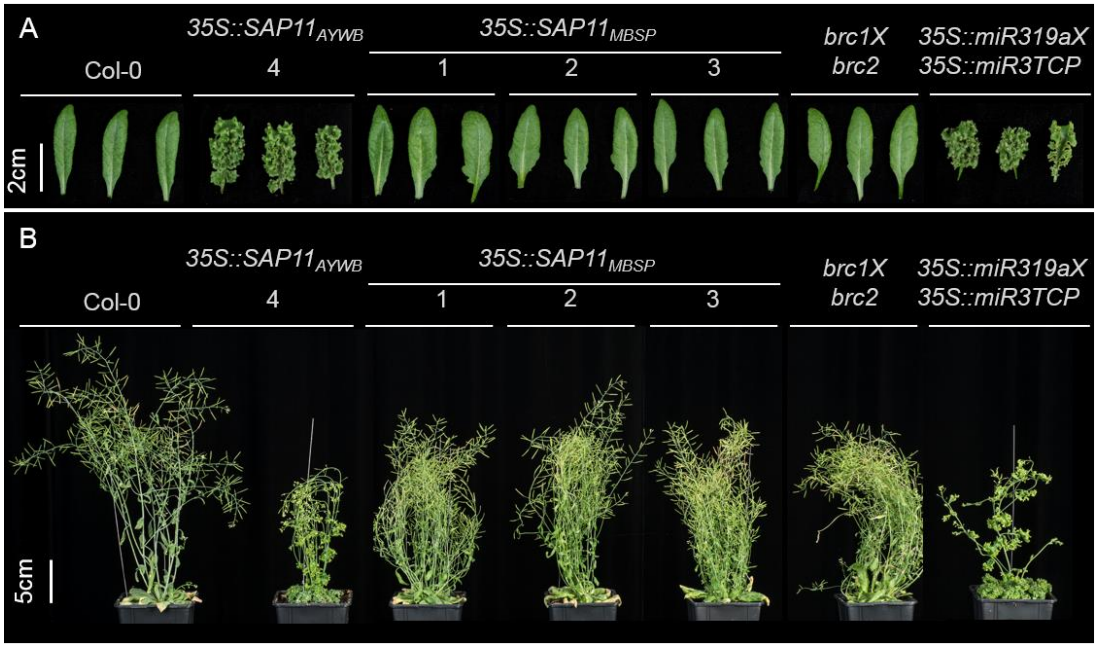


Figure 3.1: Comparison between the *SAP11$_{AYWB}$*, *SAP11$_{MBSP}$* transgenic *A. thaliana* lines and wild type Col-0. Stable expression of SAP11$_{AYWB}$ and SAP11$_{MBSP}$ in *Arabidopsis thaliana* lines show that SAP11$_{AYWB}$ induces lateral shoot branching (witches' broom) and leaf crinkling while SAP11$_{MBSP}$ only induces lateral shoot branching. (**A-B**) *35S::SAP11$_{AYWB}$* stable transgenic *A. thaliana* (Col-0) lines phenocopy both the *A. thaliana brc1-2 brc2-1 (brc1 brc2)* double (Col-0) mutant and *35S::miR319a x 35S::miR3TCP* stable transgenic *A. thaliana* (Col-0) lines

and *35S::SAP11$_{MBSP}$* transgenic lines phenocopy only the *A. thaliana brc1 brc2* mutant . Figure and legends are taken from Pecher et al., 2019. This paper is included in **Appendix VI**.

During the first two years of my PhD, postdoctoral researcher Dr Pascal Pecher in the Hogenhout lab started to generate and characterise *Ubi::FLAG-SAP11$_{AYWB}$* and *Ubi::FLAG-SAP11$_{MBSP}$* transgenic maize lines. He found that *Ubi::FLAG-SAP11$_{AYWB}$* and *Ubi::FLAG-SAP11$_{MBSP}$* transgenic maize lines exhibit a multi-branching phenotype compared to HIIA wild type, phenocopying the *tb1* mutant (Pecher et al., 2019). However, the alteration of the leaf development is only visible in the stable expression line of SAP11$_{AYWB}$ in *Z. mays* (*Ubi::FLAG-SAP11$_{AYWB}$*) but not visible in *Ubi::FLAG-SAP11$_{MBSP}$* compared to the wild type HiIIA maize plant. The differences in the maize phenotype also suggested that SAP11$_{AYWB}$ targets both subgroups of class II TCPs and SAP11$_{MBSP}$ only the CYC/TB1 TCPs.

Therefore, whereas AY-WB and MBSP both belong to Ca. Phytoplasma asteris, albeit in subgroups 16SrIA and 16SrIB, respectively, their SAP11 effector proteins appear to have different specificities for class II TCP transcription factors. Alignment of the two SAP11 proteins show differences in sequences (discussed in the next chapter, Chapter 4). In this chapter, I will investigate the domains with TCPs that determine SAP11$_{AYWB}$ and SAP11$_{MBSP}$ binding specificities. The hypothesis is that SAP11$_{MBSP}$ has a narrower binding range towards the TCP (sub)classes compared to SAP11$_{AYWB}$. Because SAP11$_{AYWB}$ binds the TCP domain of TCP transcription factors (Sugio et al., 2014), I hypothesised that SAP11$_{MBSP}$ also interacts with the TCP domain and that regions within the TCP domains of CYC-TB1 and CIN-TCPs define binding specificities to SAP11$_{MBSP}$ and SAP11$_{AYWB}$.

Here I describe yeast-two hybrid data showing interactions of the SAP11 effector homologs and TCP members from class II CYC/TB1-TCPs and CIN-TCPs. I also added in this study two SAP11 effector homologs from two different phytoplasmas; these are SAP11$_{WBDL}$ of Witches Broom Disease of Lime (WBDL; 16SrII group) phytoplasma that primarily infects lime, and SAP11$_{SPLL}$, of Sweet Potato Little Leaf (SPLL; 16SrII group) phytoplasma that infects a wide range of dicotyledonous plants. The SAP11 proteins of these two phytoplasmas were chosen, because the

phytoplasmas are a specialist and a generalist, like AYWB and MBSP, but belong to a distinct clade within the Phytoplasma phylogeny. Given that $SAP11_{MBSP}$ targets a narrower set of TCPs than $SAP11_{AYWB}$ does, I hypothesised that $SAP11_{WBDL}$ may also target a narrower set of TCPs than $SAP11_{SPLL}$ does.

Here, I show that $SAP11_{AYWB}$ interacts with both class II CIN and CYC/TB1-TCPs, whereas $SAP11_{MBSP}$ only interacts with class II CYC/TB1-TCPs, in agreement with the phenotypes of the transgenic *A. thaliana* and maize plants. I also found that $SAP11_{WBDL}$ and $SAP11_{SPLL}$ interact with CYC/TB1-TCPs, but not CIN-TCPs. Furthermore, the SAP11 effector homologs bind to the TCP domain of *A. thaliana* class II CIN and CYC/TB1. Finally, using chimaeras of the basic region and the helix loop helix motifs of the TCP domain of both class II CIN and class II CYC/TB1 TCP candidates, I show that the SAP11s interact with the full helix-loop-helix motif of TCPs.

3.2 Results

3.2.1 Four SAP11 effector homologs interact with *Arabidopsis thaliana* Class II CYC/TB1-TCPs in yeast

Given previous results demonstrating that both $SAP11_{AYWB}$ and $SAP11_{MBSP}$ interact with class II CYC/TB1-TCPs (Pecher et al., 2019), I wished to investigate if other SAP11 effector homologs from different phytoplasma groups interacted with this TCP (sub)class.

To do this, I tested the interaction between the four SAP11 effector homologs and the *A. thaliana* class II CYC/TB1-TCPs members AtTCP12 (BRC2) and AtTCP18 (BRC1) in the Y2H system (**Figure 3.2.1**). Y2H data show that $SAP11_{AYWB}$ and $SAP11_{MBSP}$ interacted with both CYC/TB1-TCP members, as yeast colonies grew on the selective synthetic dropout medium lacking leucine, tryptophan, adenine and histidine (SD -LWAH) for each combination. Additionally, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ also interacted with Class II CYC/TB1-TCPs. Yeast colonies that contained the GBKT7 empty vector (EV) in the presence of AtTCP12 and AtTCP18 GADT7 plasmids and GADT7 EV in the presence of SAP11 GBKT7 plasmids did not grow, indicating that AtTCP12 and AtTCP18 and the SAP11 homologs do not have autoactivity in yeast. Hence yeast only

grew in the presence of both SAP11s and TCPs, confirming that these proteins interact. Interestingly, yeast colonies with $SAP11_{WBDL}$ and *A. thaliana* TCP12 did not grow, whereas those with the other three SAP11 homologs and *A. thaliana* TCP12 did, indicating that $SAP11_{WBDL}$ apparently does not interact with *A. thaliana* TCP12. However, $SAP11_{WBDL}$ does interact with *A. thaliana* TCP18. To conclude, all SAP11s bind CYC/TB1-TCPs, though $SAP11_{WBDL}$ binds only one of the two CYC/TB1-TCPs.

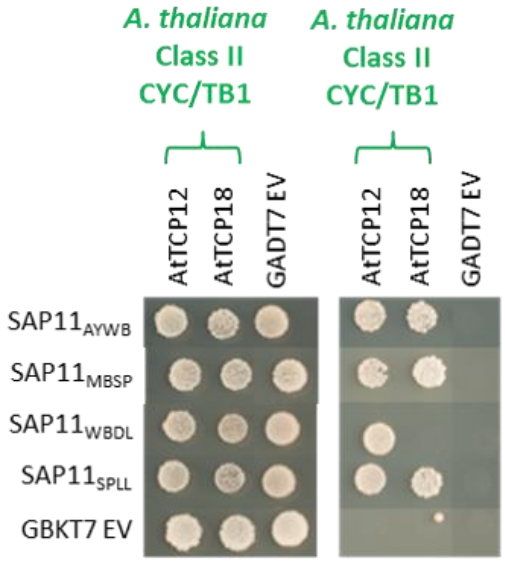


Figure 3.2.1: Four SAP11 effector homologs interact with *Arabidopsis thaliana* Class II CYC/TB1 TCPs in yeast. Yeast two-hybrid analysis of phytoplasma SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ and *Arabidopsis thaliana* Class II CYC/TB1 TCPs. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids, or lacking leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

3.2.2 $SAP11_{AYWB}$ interacts with class II CIN-TCPs in yeast two-hybrid experiments, in contrast to $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$

Given previous results demonstrating that $SAP11_{AYWB}$ interacts with Class II CIN-TCPs, whereas $SAP11_{MBSP}$ does not (Pecher et al., 2019), I wished to investigate if SAP11 homologs from WBDL and SPLL phytoplasmas, which belong to a different group within the phytoplasma phylogeny compared to AYWB and MBSP, interact with these TCPs.

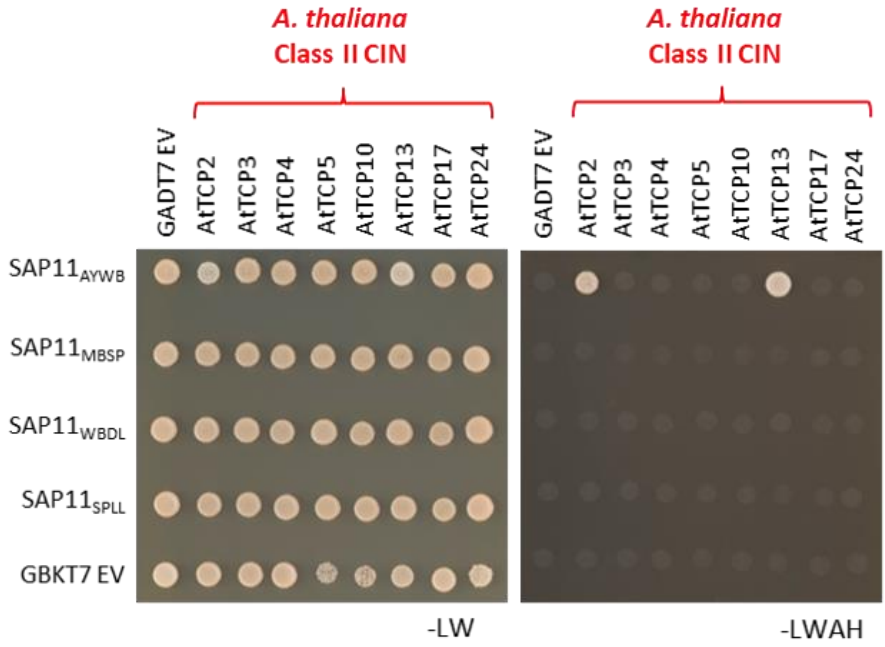


Figure 3.2.2: $SAP11_{AYWB}$ interacts with members of Class II CIN-TCPs in yeast. Yeast two-hybrid (Y2H) analysis of phytoplasma SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ and *Arabidopsis thaliana* Class II CIN-TCPs. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids, or lacking leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

To do this, I performed yeast-two hybrid (Y2H) analyses and tested the interaction of the four SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ and TCPs of the Class II CIN-TCP (sub)class. Y2H data show that $SAP11_{AYWB}$ interacts with *A. thaliana* Class II CIN-TCP2 and TCP13 in yeast, as yeast colonies grew on the selective synthetic dropout medium lacking Leucine, Tryptophan, adenine and Histidine (SD -LWAH), whereas there is no evidence of interactions of the other three SAP11 homologs with these TCPs (**Figure 3.2.2**).

Consistent with previous results, I confirmed that $SAP11_{MBSP}$ did not show interactions with any members of this TCP (sub)class. In addition, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ do not appear to interact with any of the Class II CIN-TCP members in yeast either. $SAP11_{AYWB}$ destabilises all 8 CIN-TCPs of *A. thaliana* (Sugio et al., 2011). However, in the Y2H assays, the SAP11 effector homologs interacted with CIN-TCP2 and TCP13, but not with the other 6 CIN-TCPs (TCP3, TCP4, TCP5, TCP10, TCP17 and TCP24) of *A. thaliana*. Given the absence of positive interactions for these six TCPs for any SAP11, it is most likely that the 6 TCPs did not work in the Y2H assays. Reasons may be that the TCPs did not express in yeast, SAP11-TCP interactions were weak and that the GAL4 tag interfered with TCP activity. For these reasons, only *A. thaliana* TCP2 and *A. thaliana* TCP13 were selected as candidates to investigate SAP11 interactions with class II CIN-TCPs by Y2H analyses in future experiments.

3.2.3 The TCP domain of TCP transcription factors determines the specificity of $SAP11_{MBSP}$ and $SAP11_{AYWB}$ effector binding to CYC/TB1 and CIN-TCPs, respectively.

Preliminary results from our lab showed that $SAP11_{AYWB}$ interacts with the TCP domain of CIN-TCP transcription factor (Pecher, personal communication). However, it is not clear if the TCP domain is also involved in SAP11-binding specificity for CYC/TB1 versus CIN-TCPs. To investigate this, I conducted Y2H assays of $SAP11_{AYWB}$ and $SAP11_{MBSP}$ and the TCP domains of CYC/TB1 TCP18 and CIN-TCP2. This showed that $SAP11_{AYWB}$ interacted with the TCP domains of *A. thaliana* class II CIN-TCP2 and class II CYC/TB1-TCP18 in yeast, as yeast colonies grew on the selective synthetic dropout medium lacking Leucine, tryptophan and histidine supplemented with 20mM of 3-Amino-1,2,4-triazole (3AT) (SD -LWH) (**Figure 3.2.3**). However, $SAP11_{MBSP}$

only binds to the TCP domain of *A. thaliana* TCP18, as no yeast growth was observed for the $SAP11_{MBSP}$ and CIN-TCP2 combination (**Figure 3.2.3**). These results are consistent with SAP11 binding to full-length CYC/TB1 and CIN-TCPs (**Figure 3.2.1** and **Figure 3.2.2**) and show that the SAP11-binding specificity to TCP transcription factors is determined by the TCP domain.

Taken together, these results showed that the TCP domain alone is sufficient for the SAP11 binding in yeast. The fact that two different SAP11 homologs from two phytoplasma strains, having two distinct patterns of interaction with TCP (sub)classes, showed interaction with the TCP domain suggests that this ability is conserved among SAP11 genes from different phytoplasma strains.

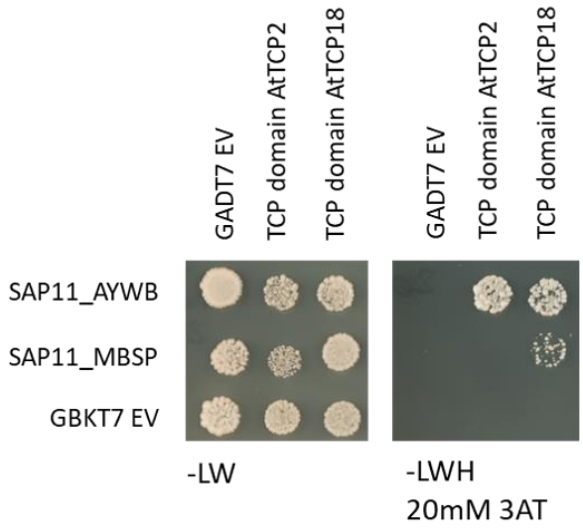


Figure 3.2.3: SAP11 binds the TCP domain of TCPs transcription factors in yeast. Yeast two-hybrid analysis of between SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$ and the TCP domains of *A. thaliana* Class II CIN-TCP2 and Class II CYC/TB1-TCP18. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCP domains and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids, or lacking leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (right); with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP

domain homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

3.2.4 The grouping of TCP domains of TCP transcription factors into distinct clades is conserved among divergent plant species

The previous yeast two-hybrid analysis shows that $SAP11_{MBSP}$ interacts with class II CYC/TB1-TCPs. MBSP is a maize pathogen, and its SAP11 protein should bind maize TCP transcription factors. As TCPs are part of a plant family of transcription factors, they are conserved across the plant species through their TCP domain (Martín-Trillo and Cubas, 2010).

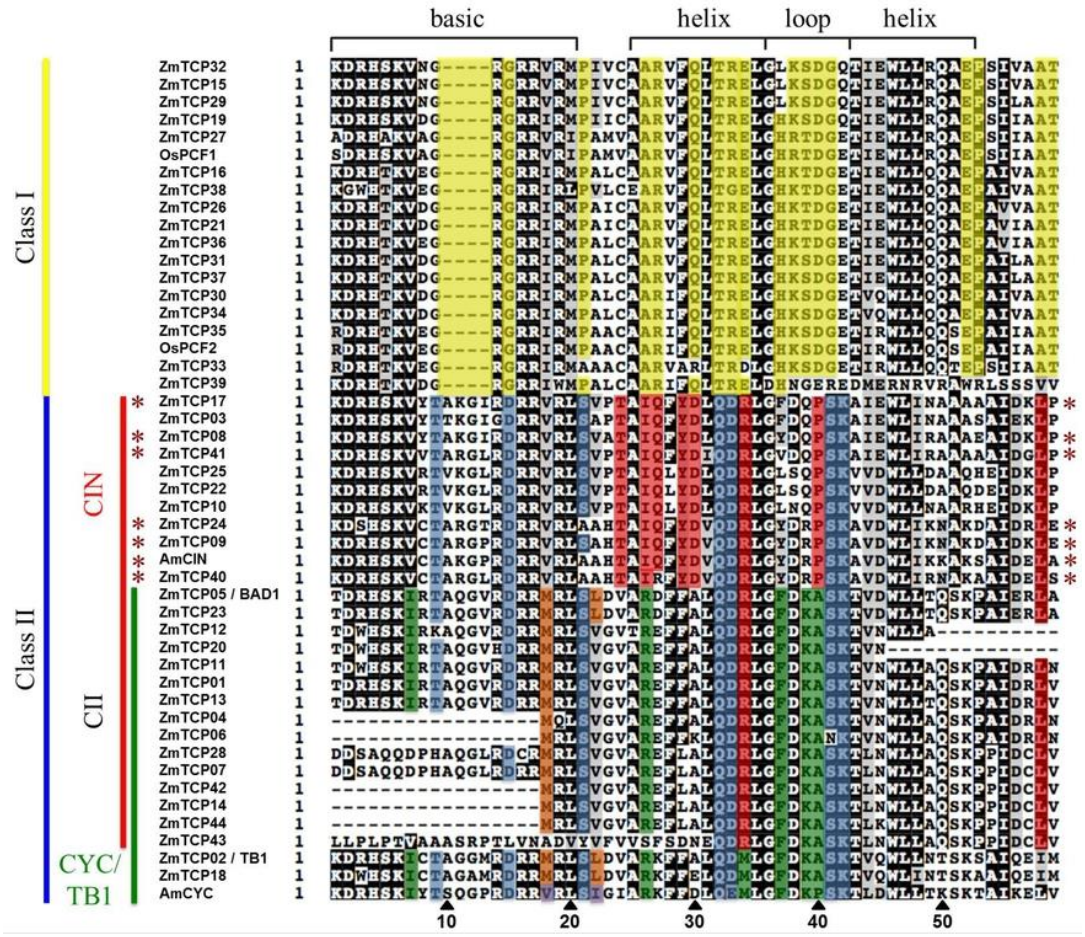


Figure 3.2.4: Multiple sequence alignment of the TCP domain of *Zea mays* TCP transcription factors. The red asterisk indicates the TCP targeted by miR319 micro-RNA. 44 TCP genes are present in the genome of *Zea mays*. Plant TCP transcription factors are divided into two main classes: class I TCP and class II TCP. Additionally, class II TCPs comprise two sub-clades: the CIN clade and the CYC/TB1 clade. *Zea mays* CII TCPs, cannot be categorised in one of the two

sub-clades CIN and CYC/TB1, as they harbour specific residues of both sub-clades. The sequences of the maize TCPs were collected from the Grass Regulatory Information Server (GRASSIUS) (http://grassius.org/grasstfdb.html) (Yilmaz et al., 2009). This figure can be found in our latest paper (Pecher et al., 2019).

The alignment of the TCP domain of the 44 maize TCPs shows that the grouping of TCP domains is conserved in maize (**Figure 3.2.4**). The maize TCPs are divided into class I TCPs and class II TCPs with specific residues that define the two classes (yellow residues for the class I TCPs and blue residues for class II TCPs), similarly to *A. thaliana* TCPs (see Chapter 1). However, the maize genome has an additional group within the class II TCPs: CII TCP(sub)class. The CII TCP domains have amino acids in common with both CYC/TB1 and CIN TCPs (shown in green and red respectively in **Figure 3.2.4**) (Chai et al., 2017; Pecher et al., 2019).

As the TCP domain is conserved among the species, I hypothesise that $SAP11_{AYWB}$ and $SAP11_{MBSP}$ have the same specificities of binding toward both A. thaliana and maize CYC/TB1 and CIN-TCPs. However, it is unclear if $SAP11_{MBSP}$ interacts with members of the CII TCPs.

3.2.5 The four SAP11 effector homologs bind maize (*Zea mays*) Class II CYC/TB1-TCP in yeast

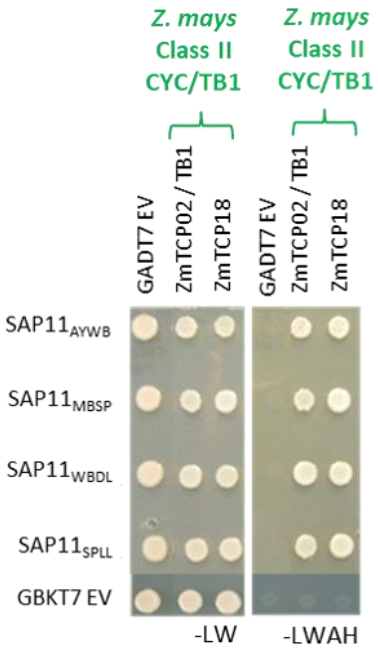


Figure 3.2.5: Four SAP11 effector homologs interact with *Zea mays* Class II CYC/TB1 TCPs in yeast. Yeast two-hybrid analysis of phytoplasma SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ and *Zea mays* Class II CYC/TB1 TCPs. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids, or lacking leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

Given the conservation of TCP domains among the class II TCP subclasses in *A. thaliana* and *Z. mays*, I predicted that $SAP11_{MBSP}$ also interacts with solely maize CYC-TB1 TCPs and not with maize CIN-TCPs, though $SAP11_{MBSP}$ may interact with the CII TCPs. Additionally, as the result of 3.2.1 infers that the four SAP11 effector homologs interact with dicotyledon *A thaliana* Class II CYC/TB1-TCPs, I wanted to confirm that the patterns of interaction were consistent using Class II CYC/TB1-TCPs from a monocotyledon genome (*Zea mays*).

To do so, I tested the interaction between the four SAP11 effector homologs and the *Z. mays* Class II CYC/TB1-TCPs members ZmTCP02 (TB1) and ZmTCP18 in the Yeast-Two Hybrid system (**Figure 3.2.5**). The Yeast-Two Hybrid results showed that the four SAP11 effector homologs interacted with ZmTCP02 (*Z. mays* TB1) and ZmTCP18 (**Figure 3.2.5**), as yeast colonies grew on the selective synthetic dropout medium lacking leucine, tryptophan, adenine and histidine (SD -LWAH). The GBKT7 Empty Vector control confirmed that ZmTCP02 and ZmTCP18 were not auto-active in yeast. Additionally, $SAP11_{WBDL}$ interacts with both members of Class II CYC/TB1-TCPs from *Z. mays*, while it interacted with AtTCP12 but not AtTCP18 (**Figure 3.2.1**).

Taken together, the Y2H analysis of 3.2.2 and 3.2.3 showed that the patterns of interaction between the four SAP11 effector homologs and the CYC/TB1-TCP (sub)classes in both *A. thaliana* (monocotyledon) and *Z. mays* (dicotyledon) are

consistent. These results suggest that the SAP11 effector homologs target conserved regions within the TCP protein across various genomes, namely the TCP domain.

3.2.6 $SAP11_{MBSP}$ do not show interaction with *Zea mays* CII members while three other SAP11 effector homologs do.

Given that there are TCP members that share both the specific residues for CIN and CYC/TB1-TCPs in *Z. mays*, I wanted to establish the patterns of interaction between the four SAP11 homologs and candidates of C II TCPs to gain insights in the targeted residues within the TCP domain. To do so, I performed Yeast-Two Hybrid analysis and tested the interaction between the four SAP11 effector homologs and *Z. mays* TCP3, TCP11 and TCP15.

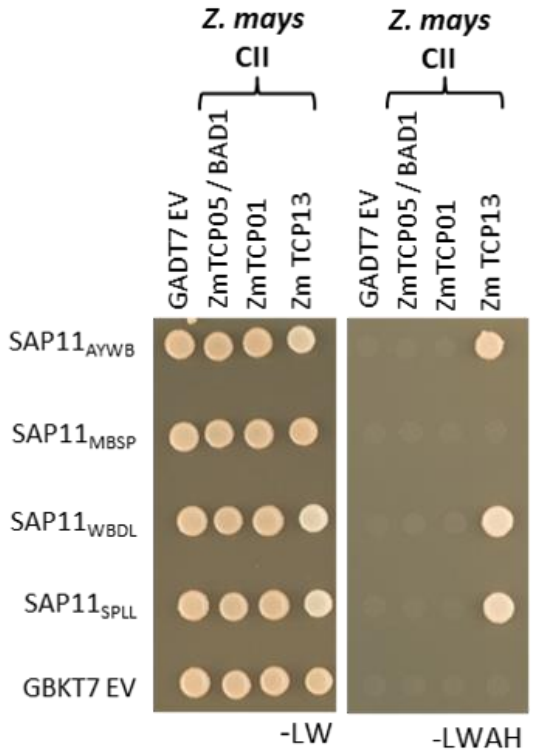


Figure 3.2.6: $SAP11_{MBSP}$ do not show interaction with *Zea mays* CII TCP members while three other SAP11 effector homologs do. Yeast two-hybrid analysis of phytoplasma SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ and with 3 *Zea mays* CII-TCPs. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast

colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids, or lacking leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

Taken together, these results indicate that there are potentially multiple TCP domain residues that are differentially targeted by SAP11 effector homologs.

3.2.7 SAP11 effector homologs interact with the entire helix loop helix motif of the TCP domain of a Class II CYC/TB1-TCP

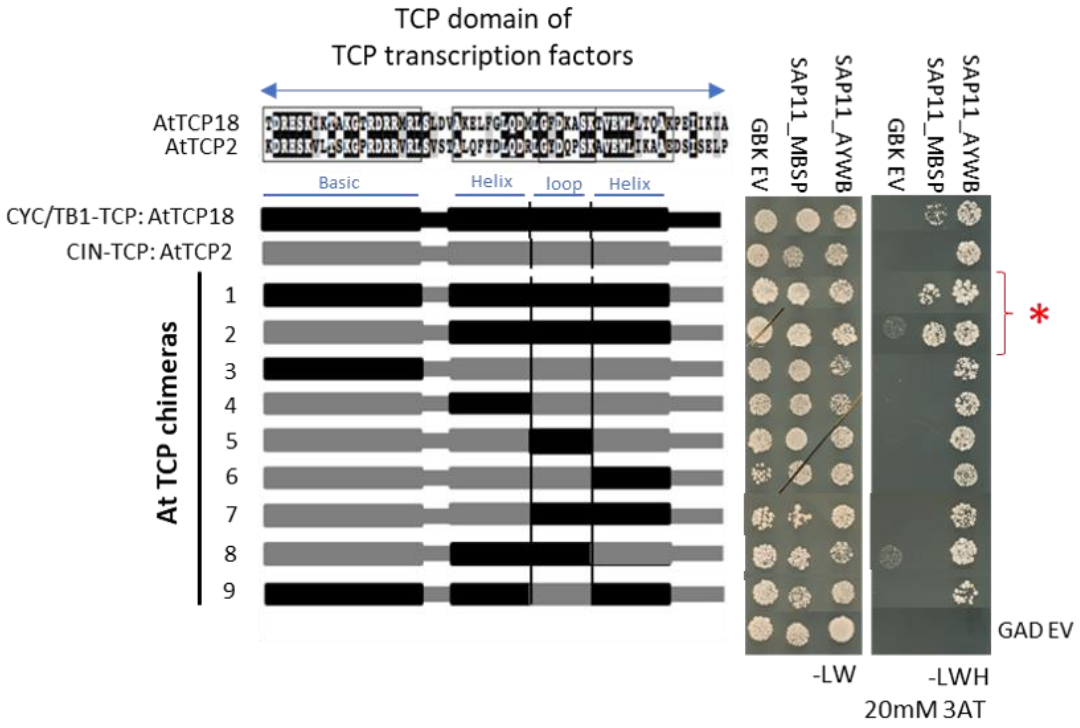


Figure 3.2.7: SAP11 effector homologs interact with the entire helix loop helix motif of the TCP domain of a Class II CYC/TB1-TCP. Yeast two-hybrid (Y2H) analysis of phytoplasma $SAP11_{AYWB}$, $SAP11_{MBSP}$ and chimeric versions based on the TCP domain of *Arabidopsis thaliana* TCP18 (Class II CYC/TB1-TCP) and *At* TCP2 (Class II CIN-TCP). SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while chimeric TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout

(SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids, or lacking leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (right); with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP chimeric versions). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

Given that SAP11 effector homologs target the TCP domains of both Class II TCP (sub)classes, I wanted to establish which region or residues within the TCP domain was specifically targeted by the different SAP11 effector homologs. In order to determine the region or the residues of the TCP domain that provide the SAP11 binding specificity, I generated chimeric versions of the TCP domain of AtTCP18 and AtTCP2, based on the different motifs within the domain: the basic motif and the helix-loop-helix motifs. I tested the interaction using $SAP11_{AYWB}$, which targets both AtTCP18 and AtTCP2 TCP domains and $SAP11_{MBSP}$, which targets only AtTCP18 TCP domain.

Based on these interaction patterns, I used the yeast-two hybrid system to screen for positive interaction between $SAP11_{MBSP}$ and a chimeric version of the TCP domain that harbours the AtTCP18 region sufficient for interaction. First, the controls confirmed that $SAP11_{AYWB}$ targetted both the TCP domain of AtTCP18 and AtTCP2, while $SAP11_{MBSP}$ targetted only AtTCP2 (**Figure 3.2.7**). $SAP11_{AYWB}$ interacted with all the chimaera constructs. However, the results showed that $SAP11_{MBSP}$ binds specifically the full helix-loop-helix motif of AtTCP18 (see the stars in **Figure 3.2.7**). Indeed, the third chimaeras construct, consisting of the basic motif of *A. thaliana* TCP2 and the full helix-loop-helix of *A. thaliana* TCP18, is targeted by $SAP11_{MBSP}$ in yeast, confirming that $SAP11_{MBSP}$ interacts specifically with the complete helix-loop-helix motif of *A. thaliana* Class II CYC/TB1-TCP18.

Furthermore, the results show that the basic motif alone is not specifically targeted by SAP11, as $SAP11_{MBSP}$ does not show any interaction with the fourth chimaera construct that consists in the basic motif of *A.thaliana* TCP18 and the full helix-loop-helix motif of *A. thaliana* TCP2. Taken together, these results suggest that

SAP11 effector is specifically targeting the helix-loop-helix motif, thus possibly disrupting the homo/hetero-dimerisation of the TCP proteins.

3.2.8 The helix-loop-helix motif of the TCP domain interacts with SAP11$_{SPLL}$

As the previous Yeast-Two-Hybrid experiment showed, the helix-loop-helix motif of the TCP domain of AtTCP18 is targeted specifically by SAP11$_{MBSP}$. As a follow-up, I wanted to investigate if the helix-loop-helix motif was sufficient for interaction with SAP11. I thus cloned the helix-loop-helix motif of both *A. thaliana* TCP18 and *A. thaliana* TCP2 and tested their interaction against the four SAP11 effector homologs in yeast. The results suggest that SAP11$_{SPLL}$ interacted weakly with the helix-loop-helix motif of *A. thaliana* TCP18 but did not show any interaction with *A. thaliana* TCP2 (**Figure 3.2.8**). SAP11$_{AYWB}$, SAP11$_{MBSP}$ and SAP11$_{WBDL}$ did not show any interaction with neither of the constructs.

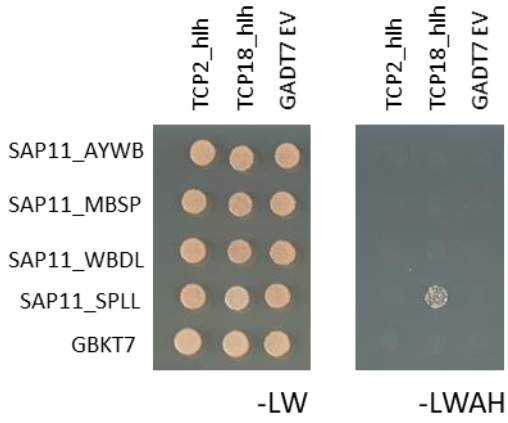


Figure 3.2.8: The helix-loop-helix motif of the TCP domain does not seem to be sufficient for interaction against SAP11 effector homologs. Yeast two-hybrid (Y2H) analysis showing the interaction between phytoplasma SAP11 effector homologs SAP11$_{AYWB}$, SAP11$_{MBSP}$, SAP11$_{WBDL}$ and SAP11$_{SPLL}$ and the helix-loop-helix motifs of the *Arabidopsis thaliana* TCP2 and TCP18. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCP motifs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids, or lacking leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions

of bait (SAP11 homologs) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

Taken together, the results suggest that the helix-loop-helix motif is not sufficient for interaction with the SAP11 effector homologs. The results need to be further confirmed with an additional experiment, such as degradation assays, as performed in our latest published work (Pecher et al., 2019) (see next chapter for more information). Indeed, it would be interesting to clone the helix-loop-helix motif in an expression vector for transient expression in *A. thaliana* protoplasts and test the destabilisation by SAP11 effector homologs.

3.3 Discussion

In this chapter, I showed that SAP11 effector homologs from four different phytoplasma strains have distinct interaction patterns with TCP (sub)classes, both in *A. thaliana* and *Z. mays*, indicating that the SAP11 ability to target TCPs is conserved among divergent plant species, including a dicot and a monocot. Our lab previously showed that $SAP11_{AYWB}$ interacts with class II CIN-TCP and CYC/TB1-TCPs (Sugio et al., 2011a, 2014). I show here that two SAP11 effector homologs from WBDL and SPLL phytoplasma strains also interact with class II CYC/TB1 TCPs while SAP11 effector from MBSP, a maize specialist, does not interact with class II CIN-TCPs. Furthermore, I show that the four SAP11 effector homologs target the class II CYC/TB1-TCPs. Taken together, **I can confirm the first hypothesis stated in the introduction of this chapter that SAP11 effector homologs have specific interactions with members of Class II CYC/TB1 and CIN-TCPs and in maize, this also includes the CII TCPs**.

Furthermore, the SAP11 effector homologs specifically target the TCP domain, which is the conserved region within the TCP proteins (Martín-Trillo and Cubas, 2010; Sugio et al., 2011b, 2014). Based on this finding, I extended the characterisation and found that the full helix-loop-helix motif of the TCP domain is required for interaction with the SAP11 effector protein. Therefore, **I answered the second hypothesis stated in the introduction, confirming that there are specific**

regions within the TCP protein that are required for the binding-specificity of phytoplasma SAP11 effector proteins.

The four SAP11 effector homologs have distinct patterns of interaction with the TCP (sub)classes in yeast. For $SAP11_{AYWB}$ and $SAP11_{MBSP}$, the interaction with the cognate TCP (sub)classes in yeast has been confirmed in *A. thaliana* protoplasts, using degradation assays (Pecher et al., 2019). Indeed, $SAP11_{AYWB}$ binds and destabilises class II CIN and CYC/TB1-TCPs, while $SAP11_{MBSP}$ binds and destabilises class II CYC/TB1 only, in *A. thaliana* protoplasts (Pecher et al., 2019) (for more details about protoplast degradation assays, see Chapter 5). The patterns of interaction and thus destabilisation of $SAP11_{AYWB}$ and $SAP11_{MBSP}$ toward TCP (sub)classes can be correlated with the phenotypes of their stable expression in both *A. thaliana* and *Z. mays* (Pecher et al., 2019; see also in **Appendix VI**).

First, the downregulation of the class II CIN-TCPs via the stable expression of both miR319 and miR3TCP leads to an alteration of the leaf development (Palatnik et al., 2003; Efroni et al., 2008). The leaf alteration of the *35S::miR319a x 35S::miR3TCP* stable transgenic *A. thaliana* (Col-0) line is phenocopied with the phenotype of the stable expression of $SAP11_{AYWB}$ in *A. thaliana*. However, the alteration of leaf development is not visible when $SAP11_{MBSP}$ is stably expressed in *A. thaliana*. These observations are also valid in maize as the alteration of the leaf development is visible in the stable expression line of $SAP11_{AYWB}$ in *Z. mays* (*Ubi::FLAG-*$SAP11_{AYWB}$) compared to the wild type HiIIA maize plant but not visible in *Ubi::FLAG-*$SAP11_{MBSP}$. Therefore, the patterns of interaction of the two SAP11 effector homologs toward Class II CIN-TCPs agrees with the phenotypes SAP11 expression induces in both *A. thaliana* and *Z. mays*.

Secondly, considering that each stable SAP11 transgenic line displays induced axillary branching, phenocopying the *A. thaliana brc1-2 brc2-1* (*brc1 brc2*) double (Col-0) mutant line (Aguilar-Martinez et al., 2007), I also confirmed that the patterns of interaction of $SAP11_{AYWB}$ and $SAP11_{MBSP}$ toward class II CYC/TB1-TCPs are consistent with the cognate phenotypes in *A. thaliana*. This association is also relevant when the two SAP11 effector homologs are stably expressed in maize, as

both *Ubi::FLAG-SAP11$_{AYWB}$* and *Ubi::FLAG-SAP11$_{MBSP}$* transgenic maize lines exhibit a multi-branching phenotype compared to HIIA wild type, phenocopying the *tb1* mutant (Pecher et al., 2019). Furthermore, SAP11$_{WBDL}$ and SAP11$_{SPLL}$ interacted also with class II CYC/TB1. Although there is no evidence that SAP11$_{WBDL}$ or SAP11$_{SPLL}$ induce axillary branching in *A. thaliana* lines, Chang et al. (2018) showed that the stable expression of SAP11$_{PnWB}$, almost identical to SAP11$_{SPLL}$, induces witches' broom phenotypes in *A. thaliana*, phenocopying *brc1 brc2 A. thaliana* mutant (Chang et al., 2018). Thus, the patterns of interaction of SAP11$_{AYWB}$ and SAP11$_{MBSP}$ toward the class II CYC/TB1-TCPs in yeast are aligned with the cognate phenotypes in *A. thaliana* and *Z. mays*. There is strong evidence that transgenic plants that produce SAP11$_{SPLL}$ phenocopy the *brc1 brc2* double mutant *A. thaliana* line based on the comparison with the stable expression of SAP11$_{PnWB}$ (Chang et al., 2018).

It is essential to underline that only two candidates of class II CIN-TCPs, AtTCP2 and AtTCP13 were competent in the Y2H assays, while the rest of the CIN-TCP candidates did not show any interaction with the SAP11 effector homologs. However, previous studies showed that SAP11$_{AYWB}$ interacts *in planta* with all the members of the Class II CIN-TCPs (Sugio et al., 2014). Why do some CIN-TCPs do not interact with SAP11$_{AYWB}$ in yeast, but are still destabilised in protoplasts? One possibility is that SAP11$_{AYWB}$ does interact directly with all the CIN-TCPs, but some CIN-TCPs are not compatible with the Y2H system, either because they are not expressed or because the tag interferes. Another reason would be that SAP11$_{AYWB}$ does not interact with all CIN-TCPs directly, but only indirectly via some CIN-TCPs. Because CIN-TCPs may form heterodimers with each other (Aggarwal et al., 2010; Viola et al., 2012; Danisman et al., 2013), SAP11 may degrade all CIN-TCPs through direct interactions with some key CIN-TCPs only.

Nonetheless, I decided to only select AtTCP2 and AtTCP13 for further studies, because I wished to screen for positive interactions in Y2H assays. Screening for positive interactions is critical as negative interactions, i.e. absence of yeast growth in selective media is less informative and does not allow definite conclusions. For example, the absence of yeast growth in the Y2H assay testing if the helix-loop-helix

motif of the TCP2 and TCP18 is sufficient for interaction with SAP11 effector homologs could be due to either (1) absence of interaction due to protein misfolding in yeast, (2) inserts in plasmids, preventing the expression of proteins in yeast or (3) no interaction. In an ideal situation, one should have positive controls for every interaction. Here, I have established the positive controls for the Y2H of the next chapters. $SAP11_{AYWB}$ is a precise control for interaction with the Class II CIN-TCPs, as Sugio et al., 2011 and Pecher et al., 2019 have shown that it can bind and destabilise the class II CIN-TCPs. CYC/TB1-TCP12 and CYC/TB1-TCP18 are also reliable controls for testing the interaction against class II CYC/TB1.

I show here that SAP11 effector homologs interact with TCP proteins from two different species, either dicotyledonous or monocotyledonous (Pecher et al., 2019). This ability is possible because of the targeting of the TCP domain, which is conserved across *A. thaliana* and *Z. mays*. Additional SAP11 effector homologs could target TCPs from other plant species, considering that SAP11 effector homologs seem to interact with the conserved TCP domain (see Chapter 5 for more details). The alignment of the maize TCPs shows that the TCP domain is subdivided into class I and class II (sub)classes CIN and CYC/TB1-TCPs in a similar way as *A. thaliana* (Martín-Trillo and Cubas, 2010; Pecher et al., 2019). It also reveals that 15 out of the 44 Zm TCPs are called CII TCPs and share conserved residues of both class II CIN and class II CYC/TB1 (Chai et al., 2017; Pecher et al., 2019). When I tested the interaction of three of these ZmTCP candidates (ZmTCP05 / BAD1, ZmTCP01 and ZmTCP13) against the four SAP11 effector homologs, I found that $SAP11_{AYWB}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ interact with these members, suggesting that they might not target the same specific residues in the TCP domain as $SAP11_{MBSP}$. Because $SAP11_{WBDL}$ and $SAP11_{SPLL}$ do not interact with class II CIN-TCPs in *A. thaliana*, it is most likely that the two effector homologs target the specific residues of the class II CYC/TB1 within the TCP domain of the Zm CII-TCPs.

The main finding of the chapter is that SAP11 effector homologs target a specific region within the conserved TCP domain. I showed that $SAP11_{MBSP}$ binds specifically to the full helix-loop-helix motif (hlh) of BRC1 AtTCP18. The hlh motif is

required for the homo/heterodimerisation of the TCPs (Kosugi and Ohashi, 1997; Aggarwal et al., 2010). TCPs can form heterodimers with other TCPs with other TCP (sub)classes, however, they show preference with members of their own (sub)class (Viola et al., 2012; Danisman et al., 2013). Analyses have shown that the TCP domain is highly disordered (Valsecchi et al., 2013), especially the basic motif (Aggarwal et al., 2010). The flexibility of the intrinsically disordered regions at the C-terminal region of the TCP domain is required for the assembly into dimers (Danisman, 2016) and also allows the protein to form multiple partners. Furthermore, the TCPs form dimers even before binding to the DNA target (Aggarwal et al., 2010). Therefore, the SAP11 ability to target the helix-loop-helix motif is a viable strategy for phytoplasma, as it could prevent the dimerization between TCPs from the same (sub)class on the one hand, but also prevents the dimerization between TCPs from different (sub)classes, on the other hand, thus reaching additional layers of control over transcriptional activity (Danisman, 2016; Spears et al., 2019).

Further experiments should be aimed at testing if other partners of the TCP target are affected upon SAP11 destabilisation. For example, $SAP11_{AYWB}$ destabilises class II CIN-TCPs, thus might interfere indirectly with the activity of other partners, such as Class I TCPs. I could assess the binding ability of Class I TCPs, that are not directly destabilised by $SAP11_{AYWB}$, once class II CIN TCPs are destabilised. To this end, I can use EMSA assays to test the binding ability of class I TCPs to the promoter of their target genes upon CIN-TCP destabilisation by $SAP11_{AYWB}$ (Aggarwal et al., 2010; Spears et al., 2019).

We cannot exclude the fact that other SAP11 effector homologs bind specifically to other regions of the TCP domain, such as the basic motif. To test this, I could extend the TCP chimaeras analysis. For example, because the four SAP11 effector homologs target class II CYC/TB1-TCPs, I could not use my chimaera sequences to find out the specific region responsible for the binding specificity toward the class II CIN-TCPs. However, I can assess the binding specificity of the Class II CIN TCPs using the Class I TCP as a backbone. Indeed, $SAP11_{AYWB}$ interacts with class II CIN-TCPs but not class I TCPs. Therefore, I could use chimaeras versions between

class I and class II TCPs and screen for a positive interaction between $SAP11_{AYWB}$ and TCP chimaeras that includes a Class I TCP domain backbone.

The stable expression of $SAP11_{AYWB}$ and $SAP11_{MBSP}$ in maize also triggers additional effects compared to their counterparts in *A. thaliana*. Indeed, we found that while *Ubi::FLAG-SAP11$_{MBSP}$* maize plants were similar to the *tb1* mutant concerning the sex determination of the organs, *Ubi::FLAG-SAP11$_{AYWB}$* triggers ear formation only instead of tassel (male) and ear (female) (Doebley et al., 1995; Hubbard et al., 2002) (Pecher et al., 2019). Several hypotheses could explain this: SAP11 could bind to additional plant targets in maize. Alternatively, the maize TCPs could have additional functions, as there is more redundancy within the maize genome.

To conclude, I found that the SAP11 effector homologs target the TCP domain of *A. thaliana* class II CIN and CYC/TB1. Finally, using chimaeras of the basic region and the helix loop helix motifs of the TCP domain of both class II CIN and class II CYC/TB1 TCP candidates, I found that $SAP11_{MBSP}$ interacts specifically with the full helix-loop-helix motif of class II CYC/TB1-TCP18. Therefore, this chapter confirms the hypothesis stated the introduction: SAP11 have distinct interaction patterns, **SAP11 effectors interact specifically with the hlh motif of the conserved TCP domains of the plant TCP transcription factors**.

Based on this result, if there are binding specificities within the TCP protein, then the hypothesis is that specific region within the SAP11 protein may be required for the binding specificity to the TCPs. Thus, the hypothesis I will address in the next chapter is that the **SAP11 has/have (a) specific region(s) that provide(s) the binding specificity toward the TCP(sub)classes and that *SAP11* gene may have evolved to target the TCP (sub)classes differentially**.

Chapter 4

The *SAP11* genes of divergent phytoplasmas have a different evolutionary history compared to the rest of the genome

4.1 Introduction

In the previous Chapter I established that the SAP11 effector homologs target specific regions within the TCP domain, of TCP transcription factors, and target all class II TCPs or only the CYC/TB1 subclass of this class. These results raise the hypothesis that SAP11 proteins also have specific regions that enable binding specificity to the TCP sub-classes and that *SAP11* genes may have evolved to target the TCP (sub)classes differentially. To address this hypothesis, it is essential to compare SAP11 protein sequences across the phytoplasma phylogeny. Therefore, I took a more comprehensive look and extended the analysis of the four SAP11 effector homologs of the previous chapter to include a broader set of SAP11 effector protein sequences. In this chapter I will investigate the SAP11 phylogeny, and in the next chapter (Chapter 5) I will use my findings to study which domains within the SAP11 proteins determine specificity for binding TCPs.

The classification of the phytoplasma genus relies on the phylogeny of phytoplasma 16S rDNA sequences (Lee et al., 2000; IRPCM, 2004). Based on this phylogeny, the '*Ca.* Phytoplasma' genus is roughly divided into nineteen 16Sr-based groups that form three major clades (**Figure 4.1.1**) (Chung et al., 2013). The first clade (highlighted in blue in **Figure 4.1.1**) the 16SrI phytoplasma group that includes the aster yellows (AY) phytoplasmas some of which were assigned to *Candidatus* (*Ca.*) Phytoplasma asteris (Lee et al., 2004). This group is further divided into the subgroups 16SrIA, 16SrIB and 16SrIC based on the 16S rDNA sequences and comprises the two previously studied 16SrIA '*Ca.* P. asteris' AY-WB phytoplasma and the 16SrIB Onion Yellows (OY) and MBSP phytoplasmas (Lee et al., 1993, 1998; Jomantiene et al., 1998). Sequences of additional genes, such as *groEL,* are used to further differentiate subgroups (Mitrovic et al., 2011; Pérez-López et al., 2016). Secondly, phytoplasma WBDL ('*Ca*. P. aurantifolia') and SPLL, which were studied in my previous chapter, belong to Clade III in the phytoplasma phylogenetic tree (highlighted in yellow in **Figure 4.1.1**), a clade that contains most of the 16Sr groups (Lee et al., 2000; IRPCM, 2004; Chung et al., 2013). Finally, Clade II (highlighted in green in **Figure 4.1.1**) includes the apple proliferation (AP) phytoplasma group,

comprising phytoplasma species such as '*Ca*. P mali' strain AT (Seemüller and Schneider, 2004; Kube et al., 2008).

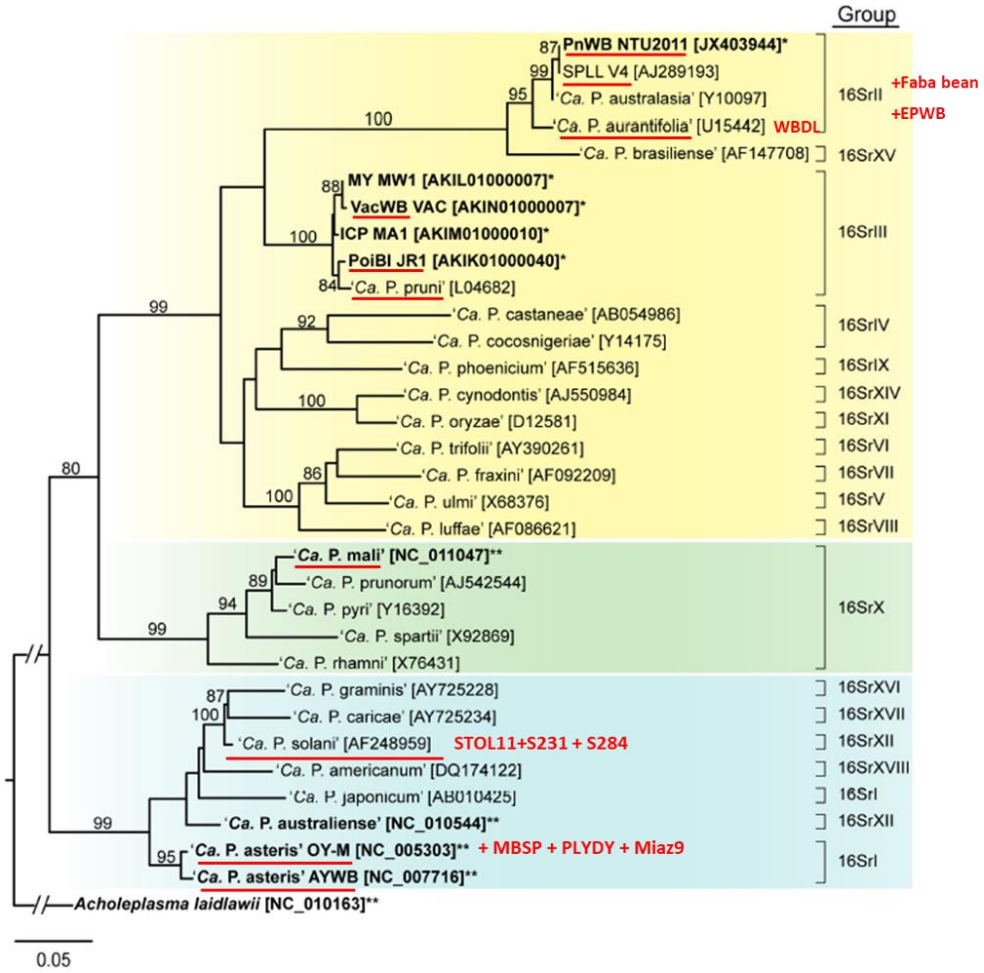


Figure 4.1.1. Phylogenetic tree of '*Candidatus* Phytoplasma', composed of three main clades. Molecular phylogeny inferred from the 16S ribosomal (16Sr) DNA sequences (Figure has been modified from Chung et al. 2013). Phytoplasmas that were found to have a SAP11 homolog are underlined in red and were further studied in this chapter. For abbreviations, see table 4.2.1.

Genes for SAP11 homologs were found in several phytoplasmas across the 16S rDNA phylogenetic tree (underlined in **Figure 4.1.1.**) and shows that phytoplasmas in all three clades have *SAP11* genes. Furthermore, the list of SAP11 sequences grew substantially with the help of collaborators that sequenced additional phytoplasma genomes and provided 16S rDNA and SAP11 sequences (see **table 4.2.1**).

Among the most striking features of phytoplasma genomes are the Putative Mobile Units, or PMU (Bai et al., 2009). PMUs are genomic regions that are repeat-rich and resemble conjugative replicative transposable elements (Bai et al., 2006; Toruño et al., 2010) (See General Introduction). Some PMUs appear to be active composite transposons; for example, PMU-1 of AY-WB has full-length sequences of all key genes and was found to also exist as a circular extrachromosomal plasmid in AY-WB, whereas other PMUs appear to carry one or more truncated genes and may therefore not be functional transposons or are dependent on other PMUs for activity (Bai et al., 2006; Dickinson, 2010; Toruño et al., 2010). Genes characteristic of PMUs are involved in DNA recombination and replication (e.g. *ssb*, *dnaB* and *dnaG*) and transposition (e.g. *tra5*) (Bai et al., 2006). PMU1 is also flanked by large inverted repeat regions, and the sequences of these repeats are also present in some other PMUs though mostly as a single sequence (Bai et al., 2006). Based on these characteristics, it was hypothesised that PMUs can replicate, horizontally transfer between phytoplasmas and integrate into genomes (Bai et al., 2006, 2009; Dickinson, 2010; Hogenhout et al., 2008; Hogenhout and Music, 2010).

Furthermore, PMUs may also have a role in phytoplasma pathogenicity, because in the AY-WB genome the majority of candidate effector genes (41 out of 56) lie on PMUs (Bai et al., 2006, 2009; Dickinson, 2010; Hogenhout and Music, 2010). AY-WB PMU1 encodes the candidate effector SAP36, which is upregulated in insects (Toruno et al., 2010) and hence may have a role in modulating insect processes during phytoplasma infection, whereas AY-WB PMU2 encodes SAP54, which modulates plant processes (MacLean et al., 2011; MacLean et al., 2014). Another PMU-like region in the AY-WB genome includes SAP11, SAP56, SAP66, SAP67 and SAP68 in an operon-like configuration (Bai et al., 2006; Toruño et al., 2010; Sugio and Hogenhout, 2012). Additionally, the expression level of PMU genes, as well as the ratio between linear and circular forms of AY-WB PMU1, vary in AY-WB-infected plants and AY-WB carrier insect vectors. The circular PMU1 is found at higher proportions in insects compared to the plant host, suggesting a role of this PMU in phytoplasma invasion of insects or that insect host cells may be a good environment for horizontal transfer of PMUs or other DNA between phytoplasmas (Dickinson,

2010; Toruño et al., 2010; MacLean et al., 2011; Sugio and Hogenhout, 2012). PMUs also encode sigma factors, which have a role in regulating gene expression, including gene regulations involved in phytoplasma switching between plant and insect hosts (Ishii et al., 2013), corroborating a role of PMUs in regulating phytoplasma virulence.

Evidence that phytoplasmas horizontally exchange DNA have since been reported (Chung et al., 2013; Ku et al., 2013; Wang et al., 2018a; Music et al., 2019). Comparative phylogeny of PMU signature genes among divergent phytoplasmas revealed that the PMUs of the 16SrII PnWB (peanut witches' broom) and the 16SrV JWB (jujube witches'-broom) phytoplasmas appear to originate from a 16SrI group phytoplasma, because the phylogenetic trees of PMU genes group these phytoplasmas together (Chung et al., 2013; Ku et al., 2013; Wang et al., 2018a). Furthermore, the recent genome sequence of '*Ca*. P. solani' SA-1 strain also revealed that PMU-associated DnaG sequences were most closely related to either '*Ca*. P. asteris' or '*Ca*. P. mali' depending on the PMU (Music et al., 2019).

Comparative analyses of phytoplasma genomes also suggest that there is a positive correlation between the number of PMU-like sequences and the number of plant species infected by phytoplasmas (Sugio and Hogenhout, 2012; Music et al., 2019). For example, both *Ca*. P. solani' SA-1 and '*Ca*. P. asteris' AY-WB have broad host ranges and between 10 and 20% of their genomes consist of PMU-like regions, whereas '*Ca*. P. asteris' MBSP and '*Ca*. P. mali' strain AT are specialists of maize and apple, respectively, and have far fewer PMU sequences (Bai et al., 2006; Kube et al., 2008; Orlovskis et al., 2017; Music et al., 2019).

The *SAP11* genes of at least three different phytoplasmas lie within PMU-like genomic regions (**Figure 4.1.2**) (Sugio and Hogenhout, 2012). The *SAP11* gene lies within PMU-like regions in the genomes of '*Ca*. P. asteris' AY-W, MIaz9 and MBSP phytoplasmas, though no PMU-like sequences were found near the SAP11 gene in the genome of OY-M phytoplasma, which is also a '*Ca*. P. asteris' member (**Figure 4.1.2**). *SAP54* genes also lie within PMU-like regions in the genomes of multiple phytoplasmas (**Figure 4.1.2**) (Sugio and Hogenhout, 2012; Orlovskis et al., 2017).

Therefore, SAP11 and SAP54 genes associate with PMU-like elements in the genomes of most phytoplasmas.

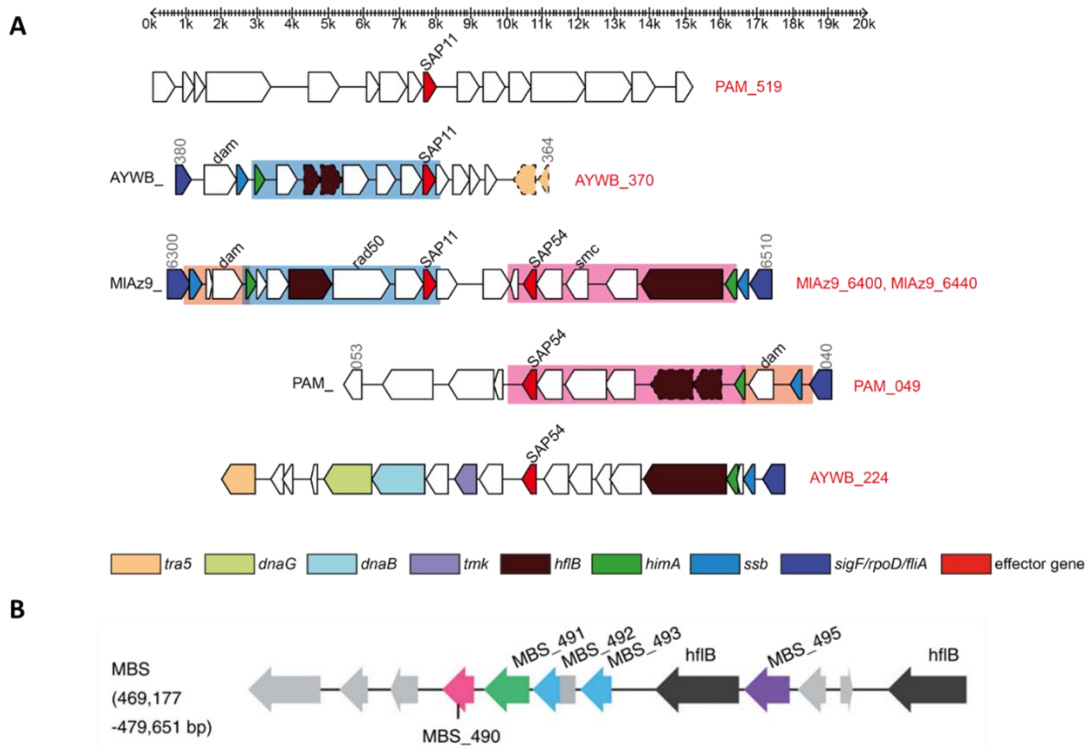


Figure 4.1.2: *SAP11* and *SAP54* effector genes lie within PMU-like regions of the genomes of '*Ca*. P. asteris' phytoplasmas. (**A**) PAM_519, AYWB_370, MiAz9_6400 correspond to the locus tags of the *SAP11* gene from '*Ca.* P. asteris' OY-M, AY-WB and MiAzp, respectively. PAM_049, AYWB_224, MiAz9_6440 correspond to the locus tags of *SAP54* gene homologs of '*Ca.* P. asteris' OY-M, AY-WB and MiAzp phytoplasmas, respectively. (**B**) The *SAP11* gene of MBSP (MBS_490) also lies in a PMU region in the MBSP chromosome (Sugio and Hogenhout, 2012). Figure A was made by Dr. Chih-Horng Kuo. The black line indicates the chromosome, while the block arrows indicate the open reading frame (ORF). Blocks of the same colours indicate paralogous ORFs among the PMUs. PMU regions that are paraloguous between different phytoplasma genomes are indicated with blue (SAP11) and pink boxes (SAP54). The names of the ORFs are displayed when their predicted function are known (Bai et al., 2006; Chung, personal communication). The stars indicate the ORFs of predicted membrane-targeted proteins. Figure B was taken from Sugio et al., 2012. *Current Opinion in Microbiology*.

Given the locations of *SAP11* genes within PMU-like regions, divergent phytoplasmas may have acquired *SAP11* homologs via horizontal gene transfer.

Therefore, phytoplasmas may carry highly similar SAP11 sequences that all target the same TCPs. It is also possible that *SAP11* sequences have evolved to target different TCPs. In any case, it is highly likely that the evolutionary history of SAP11 does not match that of the phytoplasma 16S rDNA sequences. In this chapter, I will investigate if **the *SAP11* gene has evolved differently from the rest of the genome.**

As phytoplasma phylogenies are based on 16S rDNA sequences (IRPCM, 2004), I collected the 16S rDNA sequences of all phytoplasmas that have a SAP11 homolog and conducted phylogeny analyses with these to assess the evolutionary relationships of the phytoplasmas. Subsequently, I performed phylogenetic analyses on the *SAP11* gene of these phytoplasmas. I found that the SAP11 and 16S rDNA phylogenies do not match, suggesting that *SAP11* genes were exchanged among phytoplasmas via horizontal DNA transfer. Nonetheless, the SAP11 genes cluster into distinct clades that suggest that the functions of this effector may have changed.

4.2 Results

4.2.1 Characterisation of phytoplasmas that have SAP11 homologs

To investigate the main hypothesis of this chapter, which is that the *SAP11* gene may have evolved to target the TCP (sub)classes differentially, I wished to extend my analysis from the four SAP11 effector homologs studied in chapter 3 to a larger set of SAP11 effector homologs. To do so, I collected SAP11 sequences from various phytoplasma isolates from the GenBank protein database. Additionally, I included SAP11 sequences from isolates collected in Wisconsin (USA) by Prof. Russell Groves and collaborators (University of Wisconsin, Madison, USA), the Poznań area in Poland collected by Miss Agnieszka Zwolińska and collaborators (National Research Institute (NRI), Poznan, Poland) and Dr Chih-Horng Kuo (Academia Sinica, Taipei, Taiwan). Moreover, I obtained the 16S rRNA sequences of these phytoplasmas from NCBI or collaborators. **Table 4.2.1** shows an extended list of the phytoplasma isolates used in this study, including the GenBank accessions numbers and additional information. Phytoplasmas for which there were no or incomplete SAP11 and 16S rDNA sequences were excluded from this study.

phytoplasma strain	abbreviation used in the thesis	GenBank accession 16S ribosomal DNA sequence	GenBank accession: SAP11 amino acid sequence (protein id)	16Sr (sub)group	notes	Geographic Origin	Host Range	reference
Ca. P. asteris sample p42.IB	**p42.IB**	unpublished data	unpublished data	16SrI-B	sequenced by collaborators, sampled from leafhopper. Sample p42 contains a mix of phytoplasma strain from the 16SrIA and 16srIB groups.	USA - Wisconsin. HARS Farm	-	-
Ca. P. asteris sample p44.IB	**p44.IB**	unpublished data	unpublished data	16SrI-B	sequenced by collaborators, sampled from leafhopper. Sample p44 contains a mix of phytoplasma strain from the 16SrIA and 16srIB groups.	USA - Wisconsin. Bager Farm	-	-
Ca. P. asteris Onion Yellows mild-symptom (OY-M) phytoplasma	**OY-M**	AP006628.2	WP_011160870.1	16SrI-B	This strain (OY-M) is derived from OY-W (onions yellow disease strain). Complete genome available.	Japan	Generalist	Oshima et al., 2001
Ca. P. asteris Periwinkle-leaf yellowing phytoplasma (PLYDY) strain DY2014	**PLYDY**	FJ437568.1	TKA87952.1 & TKA88140.1	16SrI-B	There are two SAP11 transcripts within the genome named SAP11 PLYDY.1 (TKA87952.1)and SAP11 PLYDY.2 (TKA88140.1).	Taiwan	Generalist	Chen et al., 2011
Ca. P. asteris MiAz9	**MiAz9**	unpublished data	unpublished data	16SrI-B	sequenced by collaborators, sampled from leafhopper (*Macrosteles leavis*) in Poland.	Poland	-	-
Ca. P. asteris Maize Bushy Stunt Phytoplasma (MBSP) strain M3	**MBSP**	CP015149.1	WP_069028078.1	16SrI-B	Complete genome available.	Brazil	Maize Specialist	Orlovskis et al., 2017
Ca. P. asteris sample p42.IA	**p42.IA**	unpublished data	unpublished data	16SrI-A	sequenced by collaborators, sampled from leafhopper. Sample phyto 42 contains a mix of phytoplasma strains from the 16SrI-A and 16srI-B (sub)groups.	USA - Wisconsin. HARS Farm	-	-
Ca. P. asteris sample p44.IA	**p44.IA**	AY389828.2	unpublished data	16SrI-A	sequenced by collaborators, sampled from leafhopper. 16S sequence identical to AYWB. Sample p44 contains a mix of phytoplasma strain from the 16SrIA and 16srIB groups.	USA - Wisconsin. Bager Farm	-	-
Ca. P. asteris Aster yellows witches' broom phytoplasma strain (AYWB)	**AYWB**	AY389828.2	WP_011412651.1	16SrI-A	Complete genome available.	USA	Generalist	Bai et al., 2006
Ca. P. asteris sample p45.IC	**p45.IC**	unpublished data	unpublished data	16SrI-C	sequenced by collaborators, sampled from leafhopper.	Poland	-	-

phytoplasma strain	abbreviation used in the thesis	GenBank accession 16S ribosomal DNA sequence	GenBank accession: SAP11 amino acid sequence (protein id)	16Sr (sub)group	notes	Geographic origin	Host Range	reference
Ca . P. solani strain STOLBUR (STOL11)	**STOL11**	AF248959.1	CCP88060.1	16SrXII-A	100% SAP11 sequence identity for the three P. solani strains. Initially isolated from *Capsicum annuum* in Serbia.	Serbia	Generalist	Davis et al., 2001
Ca . P. solani (S231)	**S231**	JQ730741.1	CCP88602.1	16SrXII	100% SAP11 sequence identity for the three P. solani strains.	Serbia	Generalist	Mitrovic et al., 2014
Ca . P. solani (S284/09)	**S284**	JQ730740.1	CCP88060.1	16SrXII	100% SAP11 sequence identity for the three P. solani strains.	Serbia	Generalist	Mitrovic et al., 2014
Ca . P. mali strain AT	**ATP**	NC_011047	WP_012504341.1	16SrX-A	Complete genome available.	Germany	Rosaceous Fruit Trees Specialist	Kube et al., 2008
Poinsettia branch-inducing phytoplasma (PBIP) strain PoiBI	**PBIP**	AF190223.1	WP_017191934.1	16SrIII-H	isolated from *Euphorbia pulcherrima (poinsettia)* in the USA by Dr Ing-Ming Lee.	USA	-	Saccardo et al., 2012
Vaccinium witches'-broom phytoplasma (VWBP) strain VAC	**VWBP**	NZ_AKIN01000007.1	WP_017193717.1	16SrIII-F	Isolated from *Vaccinium myrtillus* .	Germany	-	Saccardo et al., 2012
Faba bean phyllody (FBP) (Faba bean)	**Faba bean**	EF193354.1	KY214419.1	16SrII-C	-	Sudan	-	Martini et al., 2007
Ca . P. aurantifolia Witches' broom disease of Lime (WBDL)	**WBDL**	U15442.1	WP_078123193.1	16SrII-B	-	Oman	Lime Specialist	Zreik et al., 1995
'Echinacea purpurea' witches'-broom phytoplasma (EPWB) strain NCHU2014	**EPWB**	NZ_LKAC01000001.1	WP_054598480.1	16SrII-A	-	Taiwan	-	Chang et al., 2015
Peanut witches'-broom phytoplasma (PnWB) strain NTU2011	**PnWB**	JX403944.1	WP_004994795.1	16SrII-A	-	Taiwan	-	Chung et al., 2013
Sweet Potato little leaf (SPLL) phytoplasma, strain V4	**SPLL**	AJ289193.2	unpublished data	16SrII-A	unpublished genome. SAP11 sequence retrieved from Dr Chih-Horng Kuo.	Australia	Generalist	Gibb et al., 1995

Table 4.2.1: List of the phytoplasma isolates used in my PhD. Details of the phytoplasmas studied, including their abbreviations as they appear in the thesis, the Genbank accession numbers for the 16S rRNA sequences and the SAP11 amino-acid sequences. The geographic

origins of the phytoplasma isolates are also indicated if known, as well as the plant host range ability for each phytoplasma.

Based on **table 4.2.1**, the retrieved sequences are from various phytoplasma isolates from different geographical origins. For example, sequenced phytoplasma isolates were initially sampled from East-Asia (*Ca.* P. asteris OY-M strain sampled in Japan (Oshima et al., 2001), or *Ca*. P. asteris Periwinkle-leaf yellowing phytoplasma (PLYDY) from Taiwan (Chen et al., 2011)), Middle-East (*Ca.* P. aurantifolia WBDL from Oman (Zreik et al., 1995), Africa (Faba bean phyllody FBP from Sudan (Martini et al., 2007)), Europe (*Ca*. P. mali strain AT from Germany (Kube et al., 2008) or *Ca.* P. solani strains from Serbia (Davis and Dally, 2001; Mitrovic et al., 2014)), the American continent (Poinsettia branch-inducing phytoplasma strain PoiBI from USA (Saccardo et al., 2012)) or the Oceania continent (Sweet Potato Little-Leaf phytoplasma was first observed in Northern Australia (Gibb et al., 1995)).

The phytoplasma isolates were collected from different plant species and included specialist phytoplasmas, such as MBSP, or generalists, such as *Ca.* P. asteris AY-WB. For example, '*Ca.* P. aurantifolia' WBDL primarily infects lime ('Citrus sp.) and is first characterised in Oman but is also found in Iran and other countries of the Middle East (Zreik et al., 1995; Faghihi, 2007). The phytoplasmas I included in my study are also transmitted by a diverse insect vector species (Weintraub and Beanland, 2006); the relationship between phytoplasma and their insect vectors also defines the phylogeny of phytoplasma (Seemüller et al., 2002; IRPCM, 2004). For example, *'Ca.* Phytoplasma asteris' MBSP is transmitted by multiple species of the genus *Dalbulus*, though only *D. maidis* and *D. elimatus* are deemed viable insect vectors (Ebbert et al., 2001).

The Hogenhout lab sequenced AY-WB-infected plant and leafhopper samples from the John Innes Centre insectary. When compared with the published AY-WB genome (Bai et al., 2006), the genomes were almost identical, containing a few SNPs. More precisely, both the 16S rDNA and the SAP11 sequences were identical; thus, these sequences were not included. Samples p42 and p44 were sampled from leafhoppers from two different farms in the USA. Both sequenced samples contain

phytoplasmas from the 16SrI-A and I-B groups, indicating that the leafhoppers samples were mix-infected with phytoplasmas groups. Interestingly, only the SAP11 sequence of the 16SrI-A was identified, and no other SAP11 homologs were found, suggesting that 16SrI-B in Wisconsin, USA may not have a *SAP11* gene. Nevertheless, for each sample, only one *SAP11* gene copy was sequenced and included in this study (**Table 4.2.1**). Although the 16S rRNA sequences are identical between p42, p44 and AY-WB, the SAP11 sequences between p42 and p44 are slightly different, thus were kept in the analysis. Concerning sample p45, sequenced from a leafhopper in Poland, two distinct copies of *SAP11* genes were found and included in the analysis.

To conclude, the phytoplasma isolates gathered for this study have the advantage to be diverse, from different geographical origins, plant species and insect vectors.

4.2.2 Phytoplasma belonging to different 16Sr groups have SAP11 effector genes

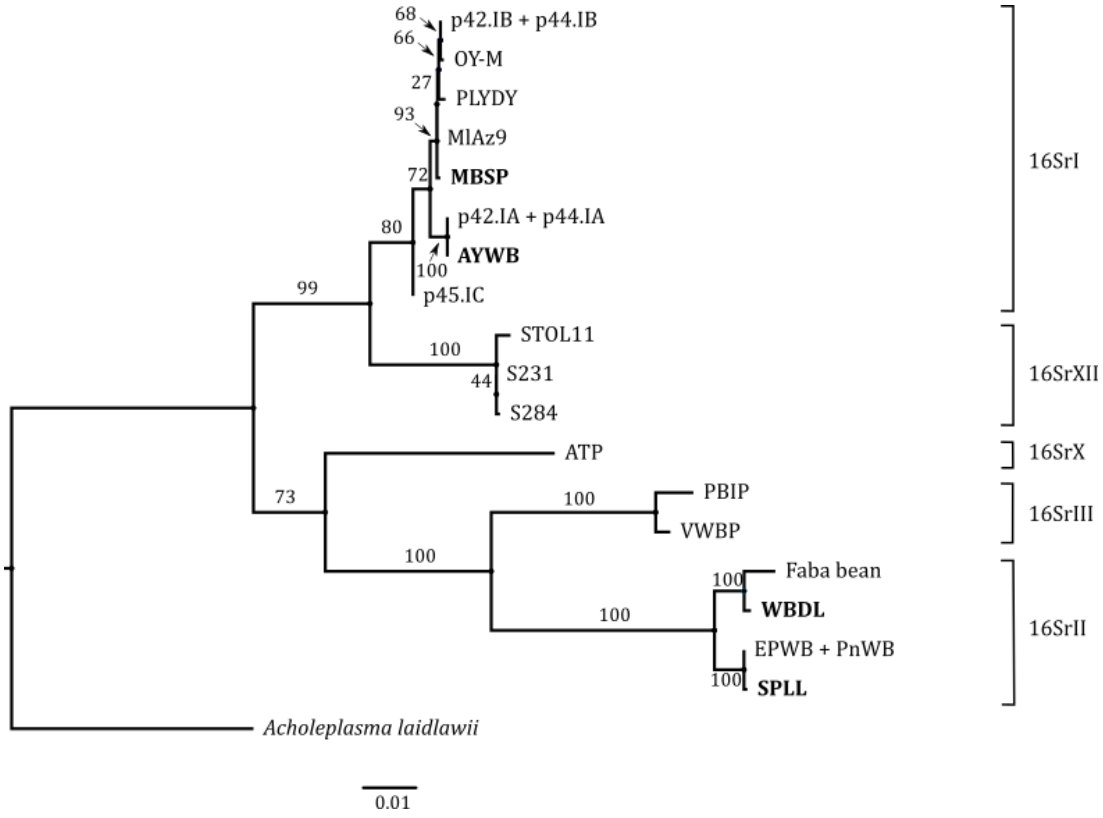


Figure 4.2.2: Phytoplasmas with SAP11 effector genes belong to diverse phylogenetic groups. Phylogenetic tree generated via the MEGA software using the Maximum Likelihood algorithm based on the alignment of 16S rDNA sequences. Phytoplasmas and sequence IDs

are described in **Table 4.2.1**. Numbers indicated on top of the branches indicate the percentage value of bootstrap support, calculated from 5000 re-samplings. I indicated the four phytoplasmas studied in Chapter 3 in bold as points of reference.

Before studying the phylogeny of SAP11, I first investigated the evolutionary relationships of phytoplasmas with SAP11 based on their 16S rDNA sequences. To do so, I collected the 16S rDNA sequences from the phytoplasma genomes (**Table 4.2.1**) and generated a Multiple-Sequence Alignment (see Appendix) using the Muscle alignment method (default parameters) (Edgar, 2004). Based on this Multiple Sequence Alignment, I then generated a phylogenetic tree using the Maximum-Likelihood method (see Material and Method Chapter for more details). The 16S rRNA sequence of *Acholeplasma laidlawii*, a Mollicute organism closely related to phytoplasma (Lim and Sears, 1992; Namba et al., 1993; Oshima and Nishida, 2007), was used as an outgroup (Lim and Sears, 1989, 1992).

The result shows that the species-tree generated via the Multiple-Alignment of the 16S rRNA sequences is consistent with published phylogenetic trees (**Figure 4.2.2**) (Hogenhout et al., 2008; Chung et al., 2013). For example, AY-WB phytoplasma and MBSP both belong to the 16srI group *Ca.* Phytoplasma asteris (in bold in **Figure 4.2.2**). The phylogenetic tree shows that the 16SrI group is subdivided into three subgroups, as shown in the phylogenetic tree. Firstly, the subgroup 16SrI-B, which includes the well-studied OY-M phytoplasma strain (Oshima et al., 2001, 2004; Tran-Nguyen et al., 2008), the MiAz9 phytoplasma isolate sampled from Poland (Zwolińska, unpublished) and MBSP phytoplasma strain M3, sampled from Brazil (Orlovskis et al., 2017). Secondly, AY-WB and phytoplasma isolates from samples p42 and p44 collected in Wisconsin belong to the 16S rIA subgroup. Finally, the sample p45 from Poland belongs to subgroup 16SrI-C. Additionally, '*Ca.* P. mali' strain AT (ATP) forms a monophyletic group, consistent with previously established trees (Hogenhout et al., 2008; Chung et al., 2013). Finally, WBDL and SPLL, studied in the previous chapter, belong to a different phytoplasma group, the 16SrII group (in bold in **Figure 4.1.1**). Therefore, I found SAP11 genes from divergent phytoplasmas.

4.2.3 Phylogenetic trees based on SAP11 nucleotide and protein sequences are similar

Next, to follow up with the investigation of the evolution of the *SAP11* gene, I wished to investigate the phylogeny of the *SAP11* genes and compare it with the species tree presented earlier. First of all, I needed to have a reliable SAP11 phylogenetic tree. For this, I decided to compare the phylogenetic trees between the SAP11 nucleotide sequences and the amino-acid based SAP11 sequences. This allows me to have more confidence in the quality of the SAP11 phylogeny that is important for future analyses.

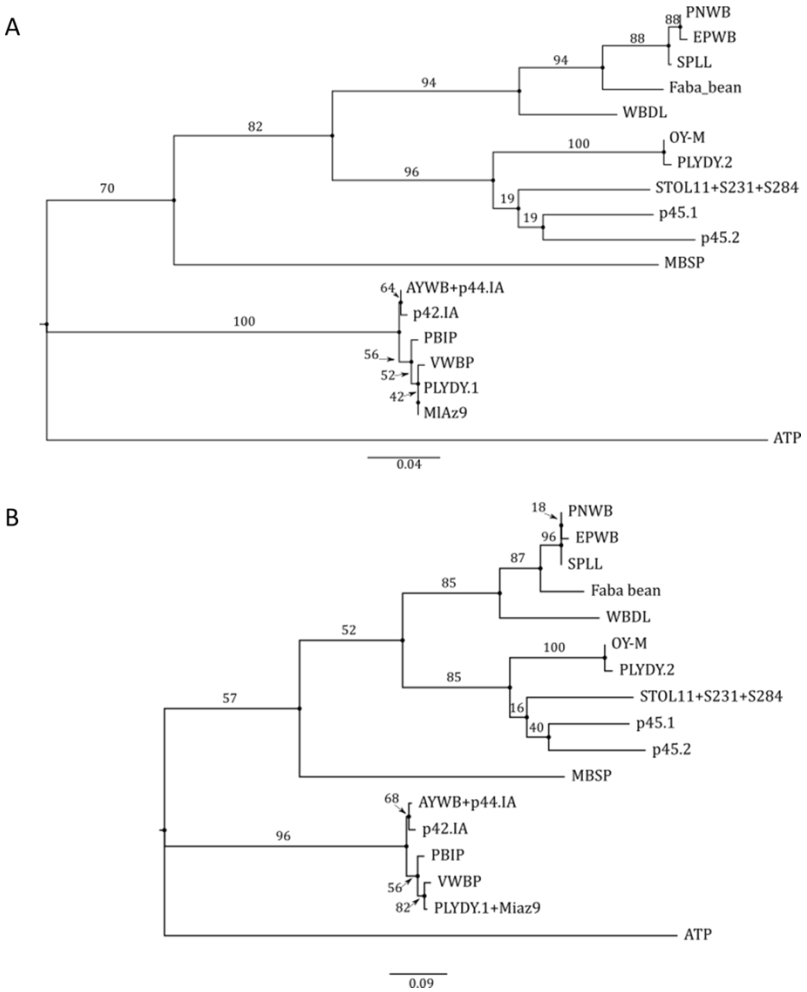


Figure 4.2.3: The SAP11 phylogenetic tree based on the nucleotide sequences is similar to the SAP11 protein-based phylogenetic tree. Phylogenetic tree generated via the MEGA software using the Maximum Likelihood algorithm. Phylogenetic trees based on Multiple Sequence Alignments of either SAP11 nucleotide sequences (**A**) or SAP11 amino acid sequences (**B**). The different nucleotide and amino-acid SAP11 sequences were included

without their Signal Peptide. The numbers indicated on top of the branches indicate the percentage value of bootstrap support, calculated from 1000 re-sampling.

To do so, Multiple Sequence Alignments were generated from SAP11 nucleotide and protein sequences of all the phytoplasmas studied and used for generating phylogenetic trees via the MEGA software (see **Figure 4.2.4 that displays the Multiple Sequence Alignment of SAP11 proteins**). Results showed that the two phylogenetic trees were similar. As expected, some sequences were identical in protein sequence, but differed in nucleotide sequences, because of the redundancy of the genetic code. The SAP11 effector homologs clustered in five different clades. The first cluster consists of $SAP11_{PnWB}$ and other 16SrII group phytoplasmas. The second clade consists of $SAP11_{OY\text{-}M}$, $SAP11_{PLYDY.2}$ and $SAP11_{solani}$ (the three effectors from the solani strains) $SAP11_{p45.1}$ $SAP11_{p45.2}$. The third clade consists of $SAP11_{MBSP}$ which forms a monophyletic group. The fourth clade includes $SAP11_{AYWB}$ and $SAP11_{p44.1}$, $SAP11_{PBIP}$, $SAP11_{PLYDY.}1$, $SAP11_{VWBP}$ and $SAP11_{MiAz9}$. Finally, $SAP11_{ATP}$ forms a monophyletic clade.

Given that the SAP11 genes are grouped into several separate clades, it is likely that the amino acid sequences of these effectors show profound differences that may affect their ability to bind various TCPs. Therefore, I studied the SAP11 protein alignment further.

4.2.4 The SAP11 protein sequence alignment

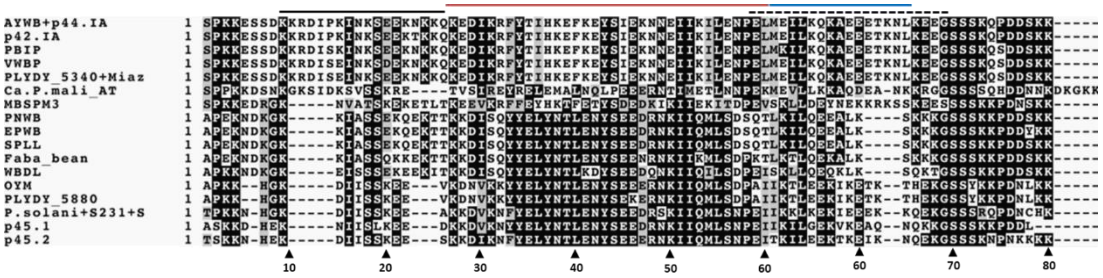


Figure 4.2.4: Analysis of the Multiple Sequence Alignment of the SAP11 effector homolog sequences. Multiple Sequence Alignment of SAP11 effector protein homologs generated using Muscle (default parameters) algorithm. The Nuclear Localization Signal of $SAP11_{AYWB}$ (NLS) is indicated as a black line, the $SAP11_{AYWB}$ region sufficient for binding to the TCP protein, in my thesis called "TCP binding domain" is indicated as a blue line (Sugio et al.,

2014) and the region between the NLS domain and the TCP binding domain, in my thesis called "intra-region" is indicated as a red line. The dashed region represents

Before generating the SAP11 phylogenetic tree based on the amino acid sequences, I started with producing the Multiple Sequence Alignment of SAP11 effector homologs, without the signal peptide. First, the Multiple Sequence Alignment of SAP11 shows that the C-terminal sequences are conserved, whereas the central region has more sequence variations between the proteins. Within the central region, there are two domains.

First, the 'KEEGSSSKQPDDSKK' sequence located at the C-terminus of $SAP11_{AYWB}$ did not affect the binding of this protein to TCP when deleted, but the deletion of the additional 'MEILKQKAEEETKNL', predicted to form a coiled-coil structure (dashed line **Figure 4.2.4**), did (Sugio et al., 2014; Pecher et al., 2019). Thus, the region 'MEILKQKAEEETKNL' is required for TCP-binding and is called in my thesis "TCP binding domain. There are multiple variations within the TCP-binding domain, suggesting that this region might also be part of the binding specificity toward TCP targets.

Secondly, the $SAP11_{AYWB}$ NLS sequence is not conserved among other SAP11 effector homologs (**Figure 4.2.4**). Indeed, the NLS prediction software NLStradamus (Nguyen Ba et al., 2009) predicts that $SAP11_{SPLL}$, $SAP11_{MBSP}$, $SAP11_{PnWB}$, $SAP11_{Faba\_bean}$, $SAP11_{EPWB}$ and $SAP11_{p45.2}$ exhibit a predicted NLS at the C-terminal part of their sequence. For example, 'KKRKSSKEESSSSKKPDNSKK' located at the C-terminus of the $SAP11_{MBSP}$ sequence is highly likely an NLS. This is in contrast to $SAP11_{AYWB}$, in which NLStradamus predict the NLS to be 'KKRDIPKINKSEEKNKKQ', which lies at the N-terminus of the mature protein, in agreement with experimental data that demonstrated that deletion of this sequence affects the $SAP11_{AYWB}$ localisation to plant cell nuclei, whereas deletion of the C-terminal 15 amino acid does not (Sugio et al., 2014). Therefore, NLSs have most likely evolved independently in $SAP11_{AYWB}$ and $SAP11_{MBSP}$, suggesting that nuclear localisation of SAP11 is essential. Some SAP11 sequences, including $SAP11_{WBDL}$, $SAP11_{OYM}$ have not been predicted to have an NLS domain. However, I show in my thesis that $SAP11_{WBDL}$

localises in the cell nuclei of *A. thaliana* protoplasts and another study also showed that $SAP11_{OY\text{-}M}$ localises in the nucleus (Chang et al., 2018).

Within these two domains, the central region between the NLS domain of $SAP11_{AYWB}$ and the TCP binding domain is called the intra-region, which is also highly diverse within the different SAP11 effector homologs.

4.2.5 The species-phylogenetic tree of phytoplasma is different from the SAP11-based phylogenetic tree

Now that the SAP11-based phylogenetic tree has been established (4.2.3), I need to compare it with the phylogenetic species-tree. I thus gathered the phylogenetic tree based on the 16S rRNA (**Figure 4.2.2**) and the protein-based SAP11 phylogenetic tree (**Figure 4.2.3**) and I colour-coded the branches of the two trees according to the 16S rRNA subgroups. This distinction would allow me to discern any difference between the trees and find indications that the *SAP11* gene has diverged from the rest of the genome, represented here by the conserved 16S rRNA gene.

As shown in **Figure 4.2.5** (**A** and **B**), the SAP11-based tree is different from the species tree. For example, as we can see in the SAP11-based tree, the SAP11 sequences belonging to the 16S r-I group (red branches) do not cluster anymore. Although the SAP11 sequences of the 16SrI-A phytoplasmas are still clustering together, as expected considering the high level of sequence similarity, the SAP11 sequences from 16SrI-B and 16SrI-C have diverged. Noticeably, SAP11 members of the 16SrI-B have diverged from the rest of the genome. For example, $SAP11_{MBSP}$ has drastically diverged and do not cluster with any other SAP11 sequence. Other members of the 16SrI-B SAP11 sequences, such as $SAP11_{OY\text{-}M}$, now form a clade with SAP11 effector homologs belonging to the 16SrXII group (solani strains). $SAP11_{Miaz9}$ is clustered with $SAP11_{AYWB}$, even though the former belongs to the 16SrI-B subgroup and the latter belongs to the 16SrI-A subgroup.

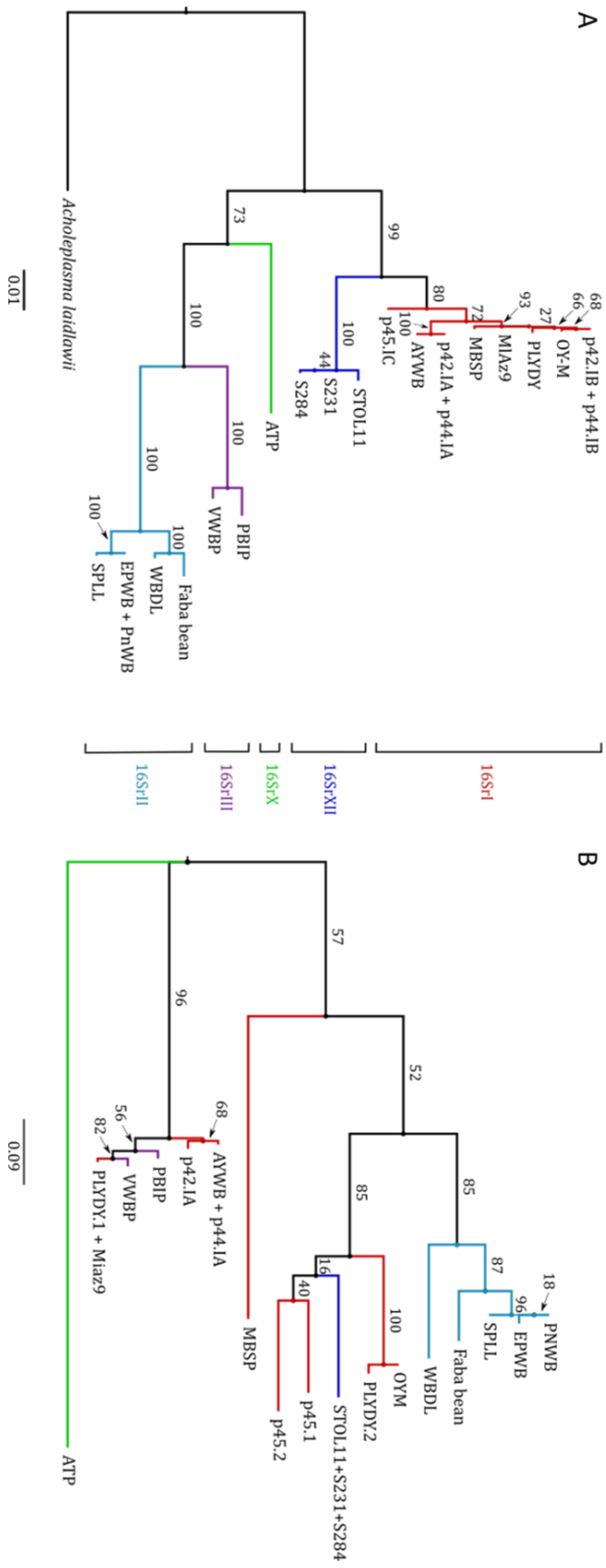


Figure 4.2.5: The species-phylogenetic tree of phytoplasma is different from the SAP11-based phylogenetic tree. Comparison between the phylogenetic tree based on the 16S rRNA

sequences of the phytoplasmas that express SAP11 and the phylogenetic tree of the phytoplasma SAP11 effector proteins based on the multiple sequence alignment of phytoplasma SAP11 effector homologs (**Figure 4.2.4**). Phylogenetic trees generated via the MEGA software using the Maximum Likelihood algorithm. **A**: Phylogenetic tree based on the Multiple Sequence Alignment of the 16S rRNA sequences of all the phytoplasma species that express *SAP11*. **B**: Phylogenetic tree based on the SAP11 homolog genes from different phytoplasma isolates. As an example, the AY-WB phytoplasma and MBSP phytoplasma (both in bold) belong to the same clade 16SrI (in red) (A) but $SAP11_{AYWB}$ and $SAP11_{MBSP}$ have divergent sequences in the SAP11-based phylogenetic tree (**B**). The numbers indicated on top of the branches indicate the percentage value of bootstrap support, calculated from 5000 re-sampling (**A**) or 1000 (**B**).

The two transcripts $SAP11_{PLYDY.1}$ and $SAP11_{PLYDY.2}$ are different, and cluster either with $SAP11_{OY\text{-}M}$ ($SAP11_{PLYDY.2}$) or with $SAP11_{Miaz9}$ ($SAP11_{PLYDY.1}$). The two transcripts $SAP11_{p45.1}$ and $SAP11_{p45.2}$, however, cluster in the same clade. This suggests that the *SAP11* gene underwent duplication for some phytoplasmas or horizontal transfer occurred.

The divergence of the SAP11 sequences can be explained via the Multiple Sequence Alignment of the SAP11 protein sequences (**Figure 4.2.4**). Indeed, if we focus solely on the 16SrI phytoplasmas, $SAP11_{AYWB}$ harbours a full NLS domain (black line in **Figure 4.2.4**) while $SAP11_{OY\text{-}M}$ and $SAP11_{MBSP}$ lack an NLS domain at the N-terminal side of their sequence (Nguyen Ba et al., 2009). More precisely, the presence of the NLS domain at the N-terminal part of $SAP11_{AYWB}$, $SAP11_{PBIP}$, $SAP11_{p42.IA}$ type explains the clade delimitation within the tree, as the $SAP11_{AYWB}$ clade and $SAP11_{ATP}$ are clustered together based on their predicted N terminal NLS domain while the rest of the effector homologs cluster based on their predicted C-terminal NLS domain. This suggests that the NLS domain may have evolved independently in the SAP11 proteins.

Taken together, these observations suggest that the *SAP11* gene has been exchanged among phytoplasma via horizontal transfers.

4.3 Discussion

This chapter shows that the *SAP11* genes from phytoplasma isolates have a different evolutionary history from the rest of the genome because the SAP11 protein-based phylogenetic tree is distinctly different from the species-tree based on the 16S rDNA sequences.

4.3.1 Details about the phytoplasma isolates of this study

This conclusion is supported via the extensive list of phytoplasma isolates that express *SAP11*. The fact that the 21 phytoplasma isolates belong to the three main clades of phytoplasma is a clear indication that the ability to express *SAP11* is widely extended across the genus (IRPCM, 2004). Some samples were removed from the study, as I did not have reliable 16S rRNA sequences available or reliable SAP11 sequences.

In addition to belonging to each of the phytoplasma clades, the phytoplasmas are also very different based on their geographical localisation, host range and their affinity with insect vectors. The combination of all these parameters strengthens the phylogenetic analyses performed in this chapter and confirms the importance of SAP11 in the pathogenicity of phytoplasma. On a technical point of view, these parameters can also be used to fine-tune the phylogeny of phytoplasma isolates. Indeed, although the 16S ribosomal RNA is an excellent primary marker gene for the phytoplasma phylogeny (IRPCM, 2004; Sugio and Hogenhout, 2012), its high heterogeneity among phytoplasmas from identical 16S sub-groups proves to be problematic to identify confidently different isolates (Liefting et al., 1996; Jomantiene et al., 2002; Davis et al., 2003). Additional marker genes such as *groEL* could be used to differentiate further the16SrI phytoplasma group (Mitrovic et al., 2011; Pérez-López et al., 2016). However, I did not use these additional markers, as the 16S rRNA phylogeny is enough to show the incongruency between the species phylogeny and the SAP11-based phylogeny.

With the help of collaborators, the addition of new phytoplasma isolates in the study was essential for a robust phylogenetic analysis. The number of SAP11

sequences is important to build a reliable phylogenetic tree and establish monophyletic clades, such as $SAP11_{MBSP}$ and $SAP11_{ATP}$. At the start of my PhD, I did not have enough SAP11 sequences. Therefore, it was difficult to correlate the patterns of interaction of SAP11 effector homologs and the SAP11 phylogeny. Having over 20 phytoplasma isolates allowed me to establish reliable hypotheses for future work (see 4.3.3 in this section for more details).

4.3.2 The *SAP11* gene resides within PMUs in phytoplasmas

The evolution of the *SAP11* gene could be attributed mainly to the most characteristic and singular feature of the phytoplasma genome: Putative Mobile Units (Bai et al., 2006; Dickinson, 2010; Toruño et al., 2010). The PMUs encode for genes involved in duplication and transposition, thus potentially allowing genetic material to be shared within PMU structures and across phytoplasmas. Indeed, studies have shown that PMU signature genes are shared within phytoplasma isolates, confirming the horizontal gene transfer across isolates (Chung et al., 2013; Wang et al., 2018a; Music et al., 2019). As the majority of the AY-WB effectors reside in PMU structures, this trend could be similar in other phytoplasmas (Bai et al., 2006, 2009; Toruño et al., 2010), although some phytoplasmas do not exhibit their *SAP11* genes in PMU structures (see Introduction of this Chapter). Therefore, *SAP11* genes are also shared between phytoplasmas as the gene resides in PMU structures across most of the phytoplasma isolates.

The rate of evolution of the PMU-related genes such as *SAP11* is likely to be faster than the rate of genes that reside outside PMU structures. Pathogens often have compartmentalised genomes with different rates of evolution (Croll and McDonald, 2012; Dong et al., 2015). For example, bacteria often display genes such as toxins and antibiotic resistance factor genes within plasmids, which are compartmentalised (Robicsek et al., 2006; Croll and McDonald, 2012). Viruses also rely on satellite RNAs for compartmentalisation (Hu et al., 2009). In filamentous pathogens, the notion of two-speed genomes has emerged to describe this concept (Raffaele et al., 2010; Dong et al., 2014, 2015). Filamentous pathogens such as *Phytophtora infestans* have a significant portion of repeat-rich regions within their

genome, similarly to phytoplasma genomes (Raffaele et al., 2010; Dong et al., 2015). In their review, the authors elegantly explain that effectors reside in repeat-rich regions mobile elements. These repeat regions were found to have a different rate of evolution than the rest of the core genome, hence the term "two-speed" genome (Dong et al., 2015). The concept of the two-speed genome relies on different parameters. First, the genome tends to be large and possess repeated regions. Secondly, the pathogens must rely on an extensive repertoire of effectors and thirdly these effectors must reside on compartmentalised mobile elements (Croll and McDonald, 2012). Aside from the size of the genome, these features are very relevant to the genetic features of phytoplasma, thus strongly suggest that SAP11 effector is evolving at a faster rate than the core genome.

The phylogenetic approach that I used here could also be useful to study the phylogeny of other SAP effectors, as multiple effectors reside in PMU structures (Bai et al., 2006; Chung et al., 2013; Ku et al., 2013) (see General discussion for more details).

4.3.3 The *SAP11* gene appears to form distinct clades across the phytoplasma isolates based on the interaction patterns with TCP(sub)classes

If the *SAP11* gene has a different evolutionary history from the rest of the genome, is it possible that the gene has evolved based on other characteristics? For example, does *SAP11* evolve based on the host range of the phytoplasma isolate and/or based on its ability to interact with plant targets? The first hypothesis is that *SAP11* gene has evolved based on its host range while the second hypothesis is that *SAP11* gene has evolved based on its ability to interact with plant targets. Alternatively, both hypothesis can be true.

The first hypothesis is supported by the phylogeny of $SAP11_{MBSP}$ and $SAP11_{ATP}$ compared to the rest of the SAP11 phylogeny. Indeed, both effectors form distinct monophyletic groups in the SAP11 phylogeny, and they are both expressed from strictly specialist of maize or apple respectively (Nault, 1980; Seemüller and Schneider, 2004). Both specialists have a lower proportion of repeated regions and PMU numbers compared to generalist strains such as '*Ca*. P. asteris' AY-WB

phytoplasma or '*Ca*. P. solani' (Bai et al., 2006; Kube et al., 2008; Orlovskis et al., 2017). For example, '*Ca.* P. mali' strain AT displays incomplete PMUs, with only one annotated transposase gene compared to 27 transposase genes in AY-WB (Bai et al., 2006; Kube et al., 2008). The positive correlation between the number of repeated regions or PMU number and the host range capacity of phytoplasmas strongly supports this first hypothesis (Sugio and Hogenhout, 2012; Music et al., 2019). This correlation makes sense as a phytoplasma isolate displaying a broad array of plant hosts is more likely to encounter different phytoplasma isolates within the same niches, thus share genetic material through the transposition of mobile elements. The likelihood of acquiring new effectors through horizontal transfer is thus high. I did not test this hypothesis during my thesis, however, comparative analyses of PMU signature genes of the phytoplasma isolates expressing *SAP11* could help to confirm the positive correlation between the number of repeated regions or PMU number within a genome and the host range of the phytoplasma (Sugio and Hogenhout, 2012; Music et al., 2019).

The second hypothesis is that *SAP11* gene has evolved based on its ability to interact with plant targets. To support this hypothesis, I combined the Y2H results of the four SAP11 effector homologs studied in chapter 3 (in bold in **Figure 3.2.6**) with the radial disposition of the amino-acid SAP11-based tree (**Figure 3.2.6**). For example, the patterns of interaction between $SAP11_{AYWB}$ and $SAP11_{MBSP}$ against Class II TCP (sub)classes are different, and it is clearly shown as $SAP11_{AYWB}$ belongs to a different clade than $SAP11_{MBSP}$. $SAP11_{WBDL}$ and $SAP11_{SPLL}$ also cluster differentially. The difference in binding specificity is an indication that the *SAP11* gene forms distinct clades across the phytoplasma isolates.

The current overview suggests that the SAP11 ability to bind to class II CYC/TB1-TCPs is generalised across the phytoplasmas. Indeed, the four SAP11 effector homologs belong to three distinct clades across the SAP11 phylogeny (red clade for $SAP11_{AYWB}$, green clade for $SAP11_{MBSP}$ and light yellow for both $SAP11_{WBDL}$ and $SAP11_{SPLL}$) but all interact with this (sub)class (**Figure 4.3.1**). Furthermore, a study has shown that $SAP11_{ATP}$, $SAP11_{PnWB}$ and $SAP11_{OY\text{-}M}$ bind and destabilise *A. thaliana*

class II CYC/TB1-TCPs (Chang et al., 2018). This suggests that the ancestral state of SAP11 was also able to interact with class II CYC/TB1-TCP.

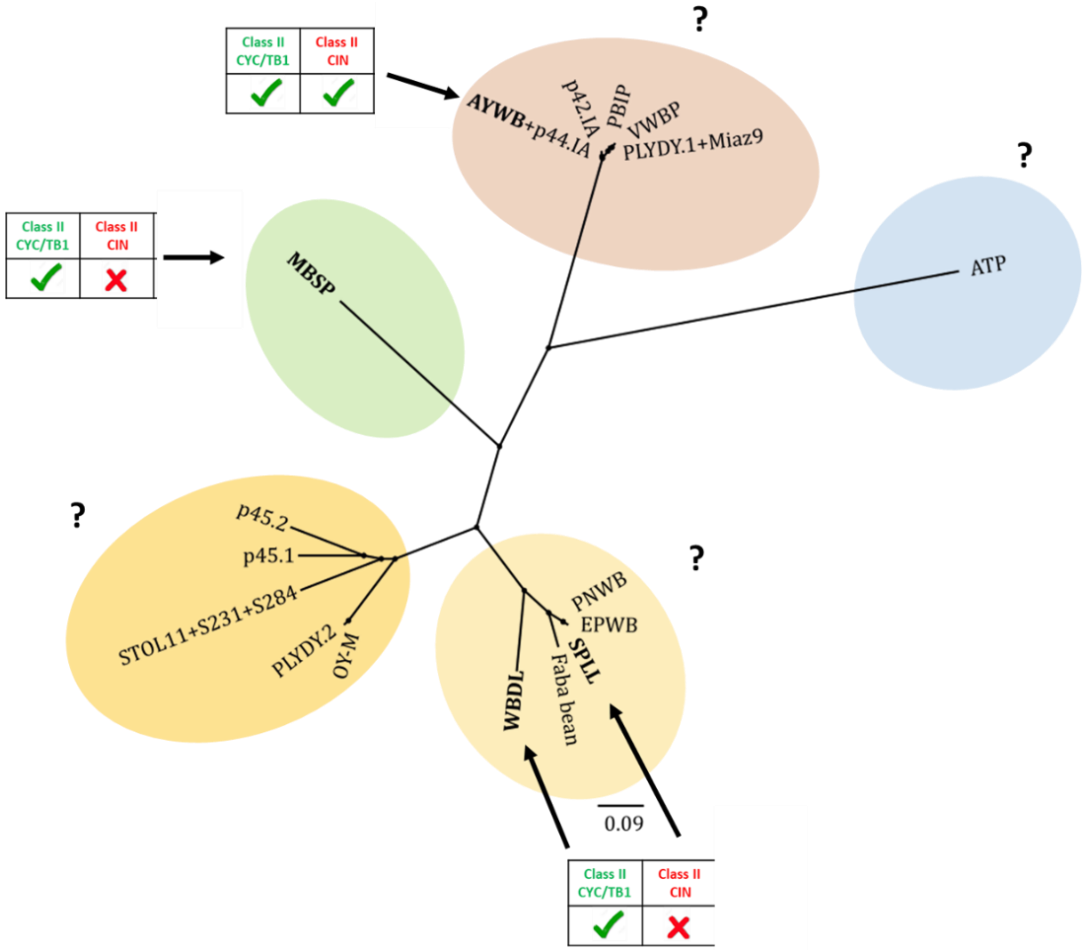


Figure 4.3.1: The *SAP11* gene appears to form distinct clades across the different phytoplasma isolates. Radial phylogenetic tree based on the Multiple Sequence Alignment of the SAP11 effector homolog proteins from different phytoplasma isolates. The phylogenetic tree has been generated via the MEGA software using the Maximum Likelihood algorithm. The results of the Y2H analysis between $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ (in bold) and the Class II TCP (sub)classes are indicated in the boxes. Based on the Y2H results and the phylogeny, I observe that the SAP11 gene forms different clades (coloured circles) across the phytoplasma isolates.

$SAP11_{AYWB}$ interacted with class II CIN-TCP members in yeast, while the three other SAP11 effector homologs did not (Chapter 3). If the *SAP11* gene has evolved based on its ability to interact with different TCP (sub)classes, then the SAP11 effector homologs that belong to $SAP11_{AYWB}$ could interact with class II CIN-TCPs.

Furthermore, Chang et al. showed that the stable expression of *SAP11$_{ATP}$* displays an alteration of the leaf development in *A. thaliana* (Chang et al., 2018). The phenotype suggests that SAP11$_{ATP}$ could interact with class II CIN-TCPs. It is important to note that the author could not find a clear destabilisation of class II CIN-TCPs by SAP11$_{ATP}$.

SAP11$_{ATP}$ behaves as an outgroup in this tree and forms a monophyletic group such as SAP11$_{MBSP}$. In this case, does this mean that his pattern of interaction with TCP (sub)classes is unique and different from the other SAP11 clades? Is SAP11$_{ATP}$ able to interact with multiple TCP (sub)classes as the transgenic lines suggest or do its TCP binding range is reduced like SAP11$_{MBSP}$? It would be exciting to assess if SAP11 binds and destabilises class I TCPs. No study has shown the interaction nor the destabilisation of Class I TCPs by any SAP11 effector homolog. The next chapter will address this question.

PLYDY phytoplasma and sample p45 display two versions of distinct SAP11 effector homologs. Gene duplication is an important mechanism that leads to the functional diversification of an effector (Hahn, 2009). This could indicate that some phytoplasma isolates rely on duplicated effector genes. SAP11$_{PLYDY.1}$ and SAP11$_{PLYDY.2}$ are part of two distinct clades in the SAP11 phylogeny. In this case, one version of *SAP11* could have been acquired via horizontal transfer with another phytoplasma strain (see Chapter 7 General Discussion). The next chapter will investigate if the patterns of interaction with the TCP (sub)classes between the two SAP11 transcripts are distinct.

The Nuclear Localization Signal of SAP11, located at the N-termini of the protein, is not conserved among the SAP11 effector homologs. The NLS prediction software NLStradamus (Nguyen Ba et al., 2009) predicted that SAP11$_{SPLL}$, SAP11$_{MBSP}$, SAP11$_{PnWB}$, SAP11$_{Faba_bean}$, SAP11$_{EPWB}$ and SAP11$_{p45.2}$ exhibit an NLS at the C-terminal part of their sequence. The distinction of the clades observed in the radial phylogenetic tree can be explained partly by the NLS sequences of the effector homologs as some SAP11 effector homologs including SAP11$_{AYWB}$ display a bipartite NLS domain and are part of the same clade.

The next step is to investigate if the other SAP11 effector homologs target the TCPs similarly to the four previously studied SAP11 effector homologs. If yes, what are the patterns of interaction for the other clades? Are the patterns of interaction the same within the same clades?

To conclude, the evolution of SAP11 is indeed different from the core genome of phytoplasmas. The radial disposition of the amino-acid SAP11-based (**Figure 4.3.1**), at first a practical way to represent the SAP11 phylogeny, is above all a clear indication that the *SAP11* gene forms distinct clades across the phytoplasma isolates. Coupled with the initial results presented in Chapter 3, I establish here the hypothesis that the ***SAP11* gene may have evolved to target the TCP (sub)classes differentially and that SAP11 has (a) specific region(s) that provides the binding specificity toward the TCP(sub)classes**.

The next chapter will address this hypothesis, with the objective to test if (1) other SAP11 effector homologs target TCPs; (2) SAP11 effector homologs have different patterns of interaction with the TCP (sub)classes; (3) characterize the interaction between SAP11 effector homologs and the TCP proteins, including the degradation ability of SAP11 toward TCP targets; (4) *SAP11* genes across the phytoplasma isolates cluster based on their interaction patterns with its plant target.

Chapter 5

SAP11 effector homologs differentially interact with TCP (sub)classes

Parts of this chapter were published in:

Pecher, P., Moro, G., Canale, M. C., Capdevielle, S., Singh, A., MacLean, A., … Hogenhout, S. A. (2019). Phytoplasma SAP11 effector destabilization of TCP transcription factors differentially impact development and defence of Arabidopsis versus maize. *PLoS Pathogens*, in press . See **Appendix VI.**

5.1 Introduction

The two previous chapters showed that the *SAP11* gene has a different evolutionary history and that some SAP11 effector homologs bind to specific regions within the TCP protein, leading to distinct binding specificities toward the class II TCP (sub)classes. The SAP11 radial phylogenetic tree of Chapter 4 shows that *SAP11* genes from different phytoplasma isolates form clades. When I coupled the Y2H analysis of Chapter 3, I observed that the *SAP11* genes tend to cluster based on their ability to target specific TCP (sub)classes. The differences in the multiple alignments of SAP11 protein sequences strongly suggest that (a) specific region(s) within SAP11 determine(s) the binding specificity to the different TCP (sub)classes. The findings of the previous chapters lead to further questions. Can the rest of the SAP11 effector homologs from various phytoplasma isolates interact with the TCPs? If yes, do they have distinct patterns of interaction, similarly to the four previously tested SAP11 effector homologs?

To verify if the ***SAP11* gene may have evolved to target the TCP (sub)classes differentially and that SAP11 has (a) specific region(s) that provides the binding specificity toward the TCP(sub)classes**, I needed to extend my analysis of the interaction between SAP11 effector homologs and TCP (sub)classes candidates. The comparative phylogenetic analyses of the phytoplasmas in Chapter 4 were complemented with functional analyses of the SAP11 effector proteins of these phytoplasmas. For this, I cloned multiple SAP11 effector homolog candidates from each of the clades of the SAP11 radial phylogenetic tree (**Figure 4.3.3**) and tested their interaction with TCP candidates of each (sub)class, including class I TCPs. I first used the yeast two-hybrid analysis, which relies on strong Y2H positive controls that were established in Chapter 3. Then, the Y2H analysis was complemented by degradation assays in *A. thaliana* protoplasts (Yoo et al., 2007). This assay is critical, as it provides *in planta* insights into the function of the SAP11 effector.

Transient protoplast assays are commonly used as a cell-based system to study signal transduction pathways (He et al., 2007). Mesophyll protoplasts from *A. thaliana* are cell-autonomous and respond to various signals such as light, plant

hormones, sugar in a similar way than tissues and whole plant (Lius et al., 1997; Gonsalves, 1998; He et al., 2007). It is therefore convenient to induce the protoplasts using these signals and assess transcriptional changes over-time (He et al., 2006, 2007). For example, protoplasts from species such as tobacco exhibit auxin-mediated responses (Koshiba et al., 1995; Ulmasov et al., 1997; Worley et al., 2000). They were used to assess over-time the degradation of Aux/IAA proteins, which are auxin transcriptional repressors (Worley et al., 2000; Dos Santos Maraschin et al., 2009).

Transient protoplast assays are often used to study the plant immunity, as the cells can be transfected with effector genes with a constitutive or inducible promoter (He et al., 2007). Protoplast assays gave insights into cell death mechanisms (Asai et al., 2000; Wu et al., 2003) or protein degradation upon effector activity (He et al., 2007; Pecher et al., 2019) as it would occur in natural conditions. For example, the recognition mechanisms of the AvrAC effector were elucidated in *Arabidopsis* protoplasts (Wang et al., 2015a). Another study used protoplasts as a cell-based genetic screen to study the MAP kinase signalling upon expression of *AvrPto* and *AvrPtoB* effectors (He et al., 2006).

In my case, transient protoplast assays were used to test the destabilisation of TCPs upon interaction with SAP11 effector homologs. Besides providing complementation of the yeast assays, they show the function of the effector *in planta*. Transient protoplast assays are also very useful to investigate indirect interactions between SAP11 homologs and TCP candidates. I also used protoplasts to observe the subcellular localisation of the effector, that are predicted to localise in the nucleus of the plants (Bai et al., 2009; Chang et al., 2018).

Based on recent studies published during my PhD, additional SAP11 effector homologs, such as $SAP11_{ATP}$ or $SAP11_{OYM}$, bind and destabilise class II TCP (sub)classes (Janik et al., 2017; Chang et al., 2018). There is therefore strong evidence that the ability to interact with TCP (sub)classes is extended across the phytoplasma isolates. It is however unclear what are the patterns of interaction of the different SAP11 homologs and if different SAP11 clades have distinct binding range toward the TCP (sub)classes. It is also unknown if SAP11 binds to class I TCPs.

Similarly to Chapter 3, a Y2H chimaera study was performed to elucidate regions within proteins that provide binding specificity. In that case, I used chimaera versions of three different SAP11 effector homologs with distinct patterns of interaction with the three TCP (sub)classes.

In this chapter, I analysed the patterns of interaction between SAP11 effector homologs belonging to the different *SAP11* clades and TCP candidates of each (sub)class. I found that the ability to target TCPs is conserved across SAP11 from different phytoplasma strains. Furthermore, the SAP11 effector homologs have distinct patterns of interaction with the plant TCP (sub)classes and some homologs bind to class I TCPs. The patterns of interaction of the four SAP11 effector homologs tested in chapter 3 were mainly confirmed via degradation assays in *A. thaliana* protoplasts (see discussion part in this chapter). When combining the *SAP11* phylogeny with the patterns of interaction (via both Y2H and degradation assays), I confirmed the hypothesis that the *SAP11* gene has evolved based on its ability to interact distinctly with the TCP(sub)classes.

Furthermore, some SAP11 effector homologs tend to bind to either class I or class II CIN. This led me to characterise the region within SAP11 that provides the binding specificity to either class II CIN or class I TCP. Using, SAP11 chimaeras, I found that the same region within SAP11 provides the binding specificity to either class I or class II CIN-TCP.

5.2 Results

5.2.1 SAP11 homologs interact differentially with the TCP (sub)classes

In order to investigate if additional SAP11 effector homologs interact with TCPs and if they can differentially interact with TCP (sub)classes, I tested the interaction between SAP11 from different clades against TCP candidates from the three (sub)classes (Class I; Class II-CIN and Class II-CYC/TB1) via the Yeast-Two Hybrid system. To do so, I synthesised seven SAP11 candidates that belong to each clade from **Figure 4.3.1** (circles) and tested their interaction with key candidates of TCP

(sub)classes. I used the four SAP11 effector homologs studied in Chapter 3 as controls.

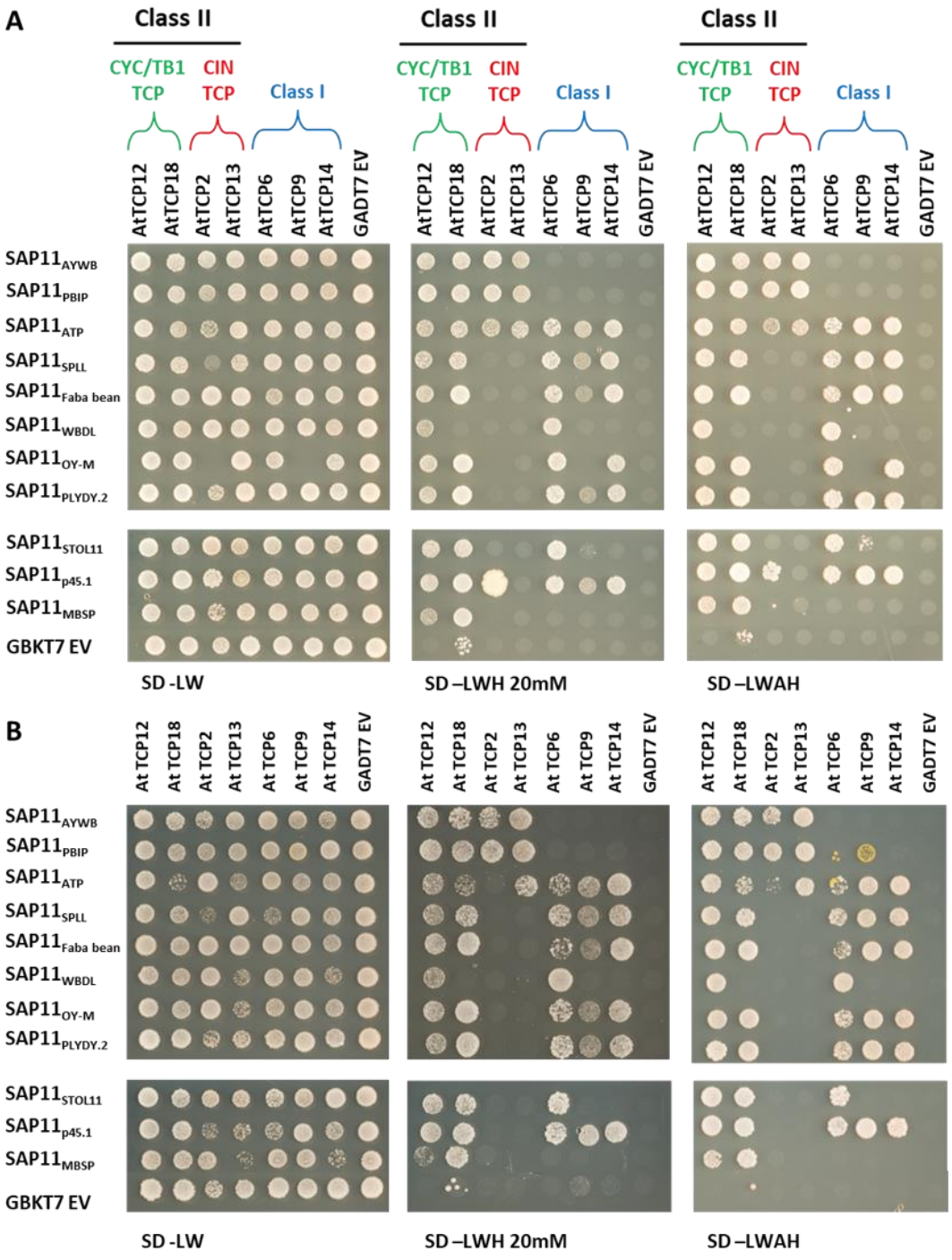


Figure 5.2.1: SAP11 homologs interact differentially with the TCP (sub)classes. Yeast two-hybrid analysis of phytoplasma SAP11 effector homologs from different phytoplasmas and *Arabidopsis thaliana* TCP (sub)classes. Two repetitions are displayed here, the first repetition (**A**) and the second (**B**). The repetitions are identical, albeit some contaminations (growth of

$SAP11_{p45.1}$+AtTCP2 for the first repetition (**A**)) and growth of $SAP11_{PBIP}$ + AtTCP9 and $SAP11_{PBIP}$ + AtTCP14 for the second repetition (**B**). TCP candidates were selected from the different (sub)classes of TCPs, Class I, Class II CIN and Class II CYC/TB1-TCP. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (middle) or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP homologs). The experiment was repeated two times with different batches of transformed yeasts and these showed identical results as shown, albeit contaminations.

The Yeast-Two Hybrid results show that the different SAP11 homologs have clear distinctive patterns of interaction with the TCP (sub)classes (**Figure 5.2.1**). The four SAP11 effector homologs tested in Chapter 3 serve as reliable controls. As expected, $SAP11_{AYWB}$ interacted with both Class II CIN-TCPs (AtTCP2 and AtTCP13) and Class II CYC/TB1 (AtTCP12 and AtTCP18) while $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ interacted only with the Class II CYC/TB1 TCPs. Neither $SAP11_{AYWB}$ nor $SAP11_{MBSP}$ showed interaction with Class I-TCPs, consistent with previous yeast-two hybrids and protoplast degradation assays (Pecher et al., 2019).

All the SAP11 effector homologs interacted with both members of Class II CYC/TB1 AtTCP12 and AtTCP18 in yeast (first two columns of **Figure 5.2.1**), except for $SAP11_{WBDL}$ that did not show interaction with AtTCP12. The negative control GBKT7 showed no interaction with AtTCP12, as colonies did not grow in the selective media SD –LWAH. The results presented here showed growth of yeast colonies carrying GBKT7 and AtTCP18 in the selective media SD-LWAH. This result is due to technical contamination, as I had tested multiple times (well over three times) these interactions.

$SAP11_{PBIP}$, closely related to $SAP11_{AYWB}$ and belonging to the same clade (red circle **Figure 4.3.1**) interacted with Class II CIN AtTP2 and AtTCP13, as expected considering its high sequence similarity with $SAP11_{AYWB}$. $SAP11_{ATP}$ from *Ca*. P. mali, behaving as an outgroup in the SAP11 phylogeny also interacted with Class II CIN-TCPs. The negative control GBKT7 showed no interaction as expected with neither AtTCP2 nor AtTCP13. Yeast colonies co-expressing $SAP11_{p45.1}$ and *AtTCP2* grew on both selective media SD –LWH (20mM 3AT) and SD –LWAH, due to some technical contamination; this was confirmed in the additional Y2H repetition displayed in **Figure 5.2.1.B**.

The TCP candidates AtTCP6, AtTCP9 and AtTCP14 were selected among the Class I TCP members of *A. thaliana* based on preliminary Y2H assays. The results show that, surprisingly, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ interacted with Class I TCPs, which was not shown in previous studies. $SAP11_{WBDL}$ interacted with AtTCP6 only, while $SAP11_{SPLL}$ interacted with the three class I TCP candidates. The negative control ensured that the Class I TCPs did not interact unspecifically with the Gal4 binding domain. Along with $SAP11_{WBDL}$ and $SAP11_{SPLL}$, $SAP11_{Faba\ bean}$, $SAP11_{OY\text{-}M}$, $SAP11_{PLYDY.2}$, $SAP11_{STOL11}$ and $SAP11_{p45.1}$ interacted with class I TCP members. $SAP11_{STOL11}$ interacted with Class I AtTCP6 but interacted weakly with Class I AtTCP9. All the SAP11 effector homologs able to interact with Class I TCPs interacted with Class I AtTCP6. However, there are SAP11 effector homologs that did not interact with Class I TCP and thus serve as reliable controls. In addition to its interaction with both (sub)classes of Class II-TCPs, $SAP11_{ATP}$ also interacted with Class I TCPs. This result shows that it is the only SAP11 effector homolog that can interact with all the TCP (sub)classes. This feature seemed to be correlated with its particular place in the SAP11 phylogeny, as it behaves as an outgroup.

To conclude, all the SAP11 effector homologs interacted with TCP candidates of Class II CYC/TB1. With the exception of $SAP11_{ATP}$ and $SAP11_{MBSP}$, the SAP11 homologs either interacted with Class I TCP or with Class II CIN-TCPs. The SAP11 phylogeny could explain this trend. Indeed, the *SAP11* genes cluster based on their pattern of interaction with the TCP targets. $SAP11_{ATP}$ interacted with every member of the (sub)classes of TCPs. A recent study showed that $SAP11_{ATP}$, from *Ca.* P. mali,

strain AT can destabilise Class II CIN-TCPs and Class II CYC/TB1-TCPs but does not destabilise the selected Class I candidates (Chang et al., 2018). Additional assays in planta are required to confirm the pattern of interaction between $SAP11_{ATP}$ and the TCP (sub)classes.

5.2.2 $SAP11_{SPLL}$ interacts specifically with the TCP domain of Class I TCP6.

The results in Chapter 3 showed that the TCP domain is involved in the SAP11-binding specificity for both class II TCP (sub)classes. However, it is unclear if the TCP domain of class I TCPs is specifically targeted by Class I interactor SAP11 effector homologs. To test this, I conducted Y2H assays of $SAP11_{SPLL}$ and $SAP11_{WBDL}$ and the TCP domains of class I TCP candidates.

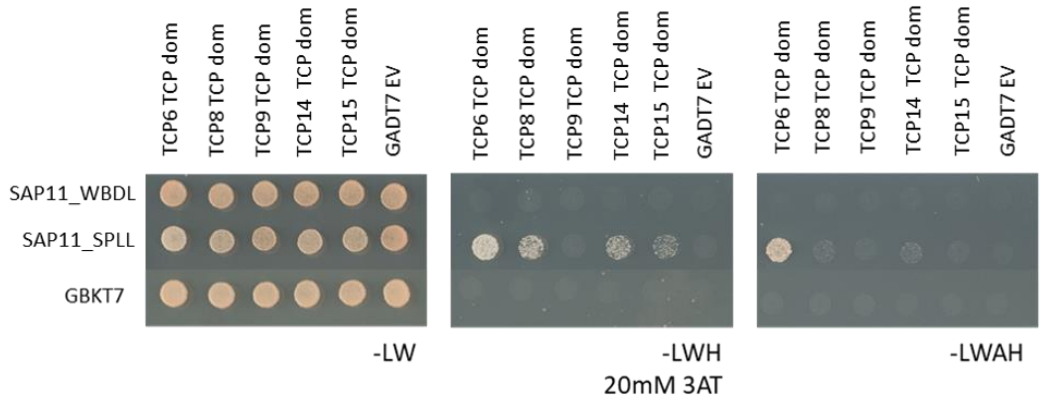


Figure 5.2.2: $SAP11_{SPLL}$ can interact specifically with the TCP domain of Class I TCPs. Yeast two-hybrid analysis of phytoplasma SAP11 effector homologs $SAP11_{SPLL}$, $SAP11_{WBDL}$ and the TCP domains (TCP dom) of *A. thaliana* Class I TCP6, TCP8, TCP9, TCP14 and TCP15. SAP11 effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (middle) or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP homologs). The

experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

The Yeast Two-Hybrid results showed that $SAP11_{WBDL}$ did not show interaction with either the TCP domain of AtTCP6, AtTCP8, AtTCP9, AtTCP14 and AtTCP15. However, $SAP11_{SPLL}$ interacted with the TCP domain of AtTCP6, AtTCP8, AtTCP14 and AtTCP15, as colonies grew in the selective media SD lacking leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation). However, yeast colonies producing $SAP11_{SPLL}$ and TCP domain of AtTCP15 grew on the selective media SD lacking leucine, tryptophan, adenine and histidine. The negative control confirmed that the TCP domain of Class I TCPs did not interact unspecifically.

To conclude, these results show that the TCP domain of class I is also targeted by SAP11. Taken together, the interaction between SAP11 effector homologs and TCP domain show that the effector bind to the TCP domains of the three (sub)classes (Class I, Class II-CIN and Class II-CYC/TB1-TCPs). Thus, the targeting of the TCP domain is extended across the phytoplasmas able to express *SAP11*.

5.2.3 SAP11 effector homologs are sub-localised in the nucleus of *A. thaliana* protoplasts cells

Given that the results show that SAP11 effector homologs differentially interacted with TCP (sub)classes in yeast, I wished to confirm that the effector proteins could interact with TCPs *in planta*, thereby confirming other results so far. Our lab previously established that $SAP11_{AYWB}$ and $SAP11_{MBSP}$ destabilise TCP targets in *A. thaliana* protoplasts (Pecher et al., 2019). However, it was not established that $SAP11_{WBDL}$ and $SAP11_{SPLL}$ destabilise plant TCPs. For this, the four *SAP11* effector homolog genes and *TCP* genes of the three (sub)classes were cloned into expression vectors that transiently express the genes in 4-weeks-old *A. thaliana* mesophyll protoplasts. Each combination of *SAP11* effector homolog genes and *TCP* genes were co-expressed in *A. thaliana* protoplast solutions. I then performed degradation assays to confirm that SAP11 effector destabilised TCP proteins *in-planta* (Pecher et al., 2019; Chang et al., 2018).

The quality of the protoplast cells is critical for the reliable production of SAP11 and TCP proteins (Yoo et al., 2007). The proportion of transformed protoplasts that produce GFP-tagged SAP11 effector proteins is assessed via counting protoplast cells producing GFP-tagged SAP11 effector proteins using the Leica DM6000 microscope (Fluorescence channel) (**Figure 5.2.3.B)**. **Figure 5.2.3.A** showed that *GFP-SAP11* effector homolog genes were indeed expressed in *A. thaliana* protoplasts. Via the GFP channel, the GFP control showed that for each combination, GFP protein is produced in the cytoplasm, the membrane and the nucleus of the cell. However, GFP-tagged SAP11 effectors seemed to be localised in the nucleus of the mesophyll cells (Bai et al., 2009).

Figure 5.2.3.B show that most of the protoplasts produce GFP-SAP11$_{AYWB}$ or GFP-SAP11$_{MBSP}$. However, there was a reduced number of transformed protoplasts producing GFP-SAP11$_{WBDL}$ or GFP-SAP11$_{SPLL}$. This might be because GFP-SAP11$_{AYWB}$ and GFP-SAP11$_{MBSP}$ are codon-expressed, while GFP-SAP11$_{WBDL}$ and GFP-SAP11$_{SPLL}$ are not.

To conclude, the microscope data shows that the four SAP11 effector homologs seem to localise in the nucleus. I did not include a reliable control for nucleus localisation, such as a plasmid expressing a yellow fluorescent protein fused with Nuclear Localization Signal, or DAPI staining, as the primary objective here was to control the transformation efficiency of the protoplasts, before the degradation assay. However, our lab showed that SAP11$_{AYWB}$ is localised in the nucleus (Bai et al., 2009; Sugio et al., 2014). The significant difference in transformation efficiency between SAP11$_{AYWB}$, SAP11$_{MBSP}$ and SAP11$_{WBDL}$ and SAP11$_{SPLL}$ is essential to assess before analysing the western blots of protoplast degradation assays.

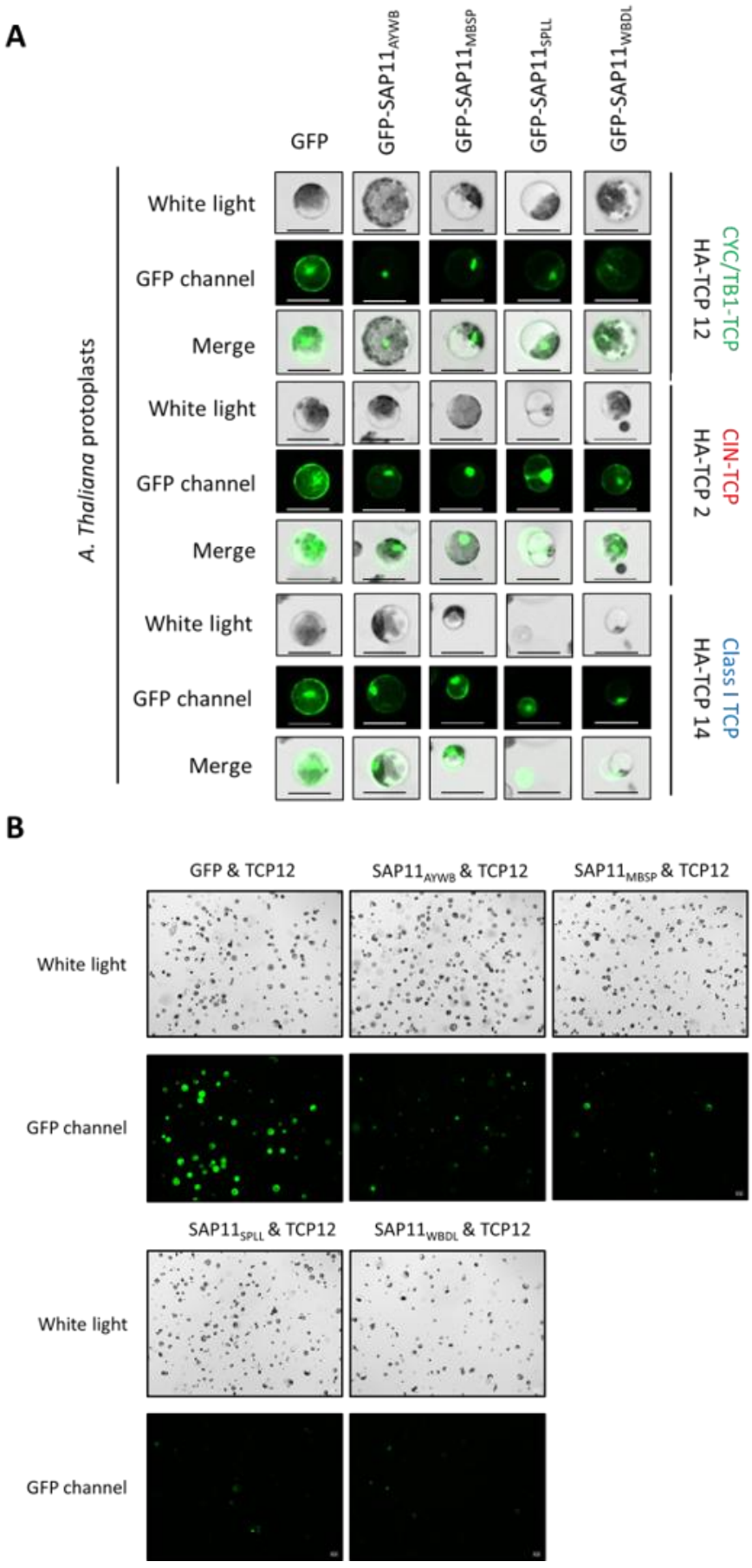

A
GFP
GFP-SAP11AYWB
GFP-SAP11MBSP
GFP-SAP11SPLL
GFP-SAP11WBDL
A. Thaliana protoplasts
White light
GFP channel
Merge
White light
GFP channel
Merge
White light
GFP channel
Merge
CYC/TB1-TCP
HA-TCP 12
CIN-TCP
HA-TCP 2
Class I TCP
HA-TCP 14
B
GFP & TCP12
SAP11AYWB & TCP12
SAP11MBSP & TCP12
White light
GFP channel
SAP11SPLL & TCP12
SAP11WBDL & TCP12
White light
GFP channel

Figure 5.2.3: the SAP11 effector homologs are sub-localised in the nucleus of *A. thaliana* protoplasts cells. (A) Confocal laser-scanning microscopy images show that SAP11 effector homologs were produced in the presence of candidates of TCP (sub)classes in *A. thaliana* mesophyll protoplasts. The fluorescence signal of GFP-tagged SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{SPLL}$ and $SAP11_{WBDL}$ is assessed via the Fluorescence channel (GFP channel) of the Leica DM6000 microscope. (B) The transformation efficiency of the different SAP11 effector homologs can be controlled via the measure of the fluorescence of the GFP tag. The fluorescence signal of GFP-tagged SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{SPLL}$ and $SAP11_{WBDL}$. $SAP11_{SPLL}$ and $SAP11_{WBDL}$ are not codon-optimised; thus, the fluorescence is reduced in these samples.

5.2.4 SAP11 effector homologs differentially destabilise the TCP (sub)classes in *A. thaliana* protoplasts

After showing that the GFP-tagged SAP11 effector homologs were produced in the presence of each TCP candidate, I performed protoplast degradation assays using GFP-antibody to detect GFP-SAP11 and HA antibody to detect HA-TCP (see Material and Methods). The protoplast degradation assays were used over co-immunoprecipitation assays to molecularly characterise the effector *in-planta* and assess if additional SAP effectors, especially Class I interactors can also destabilise their TCP targets.

The western blots show that SAP11 effector homologs differentially destabilised TCP (sub)classes (**Figure 5.2.4**). For example, GFP-$SAP11_{AYWB}$ destabilised Class II CYC/TB1 AtTCP12, as shown in the anti-HA western blot. Indeed, in the presence of GFP-$SAP11_{AYWB}$, HA-tagged TCP12 was barely visible compared to HA-TCP12 when in presence with the negative control GFP. GFP-$SAP11_{AYWB}$ was detected by western blot with GFP antibodies confirming that GFP-$SAP11_{AYWB}$ is present. The band size is expected to be around 38 kDa (9.5kDa for $SAP11_{AYWB}$ + 28kDa of GFP).

$SAP11_{AYWB}$ and $SAP11_{MBSP}$ destabilised HA-tagged TCP12 (expected size of around 41kDa), as shown in the anti-HA western blot, consistent with the previous Yeast-Two Hybrid studies. $SAP11_{WBDL}$ and $SAP11_{SPLL}$ also seemed able to destabilise HA-TCP12, consistent with the previous yeast results. However, the two GFP-tagged SAP11 effector homologs were not detected with the anti-GFP. As shown in the

confocal images (**Figure 5.2.3**), this could be since the two constructs expressing $SAP11_{SPLL}$ and $SAP11_{WBDL}$ had not been codon-optimised for *A. thaliana*, unlike $SAP11_{AYWB}$ and $SAP11_{MBSP}$.

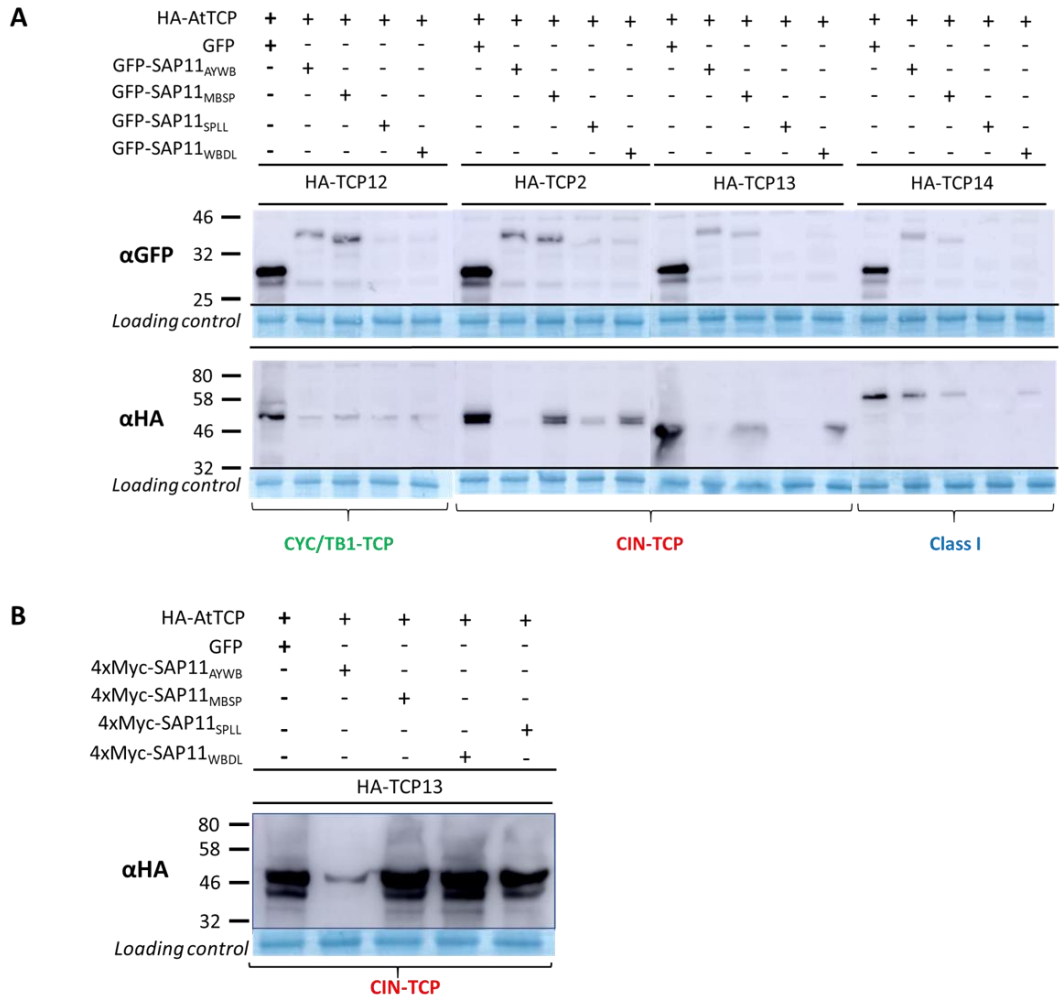


Figure 5.2.4: SAP11 effector homologs differentially destabilise the TCP (sub)classes in *A. thaliana* protoplasts. Western blots of *A. thaliana* protoplast destabilisation assays. A: *SAP11* effector homolog genes were co-expressed with candidates of TCP (sub)classes Class I, Class II CIN and Class II CYC/TB1 genes. $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ differentially destabilise Class II CYC/TB1-TCP AtTCP12, Class II CIN-TCPs AtTCP2 and AtTCP13 and Class I AtTCP14. GFP-tagged SAP11 effector homologs or GFP alone were detected using anti-GFP antibody while HA-tagged TCPs were detected with a specific anti-HA antibody. Loading controls: Amidoblack-stained large RUBISCO subunit. **B:** 4xMyc-$SAP11_{AYWB}$, 4xMyc-$SAP11_{MBSP}$, 4xMyc-$SAP11_{WBDL}$ and 4xMyc-$SAP11_{SPLL}$ differentially destabilise Class II CIN-TCPs HA-AtTCP13. HA-tagged TCPs were detected with a specific anti-HA antibody. I could not

detect the four SAP11 effector homologs using a specific antibody against 4xMyc. Loading controls: Amidoblack-stained large RUBISCO subunit.

$SAP11_{AYWB}$ destabilised both HA-tagged CIN-TCP2 and HA-TCP13 while $SAP11_{MBSP}$ did not, confirming the patterns of interaction shown in previous yeast-two hybrid assays. It is important to note that when present with GFP-$SAP11_{MBSP}$, HA-TCP13 seemed to be less abundant compared to the negative control. However, previous protoplast degradation assays confirmed that $SAP11_{MBSP}$ could not destabilise Class II CIN-TCPs and AtTCP13 specifically (Pecher et al., 2019). Furthermore, using other constructs over-expressing 4xMyc-tagged *SAP11* effector homolog genes (**Figure 5.2.4.B**), $SAP11_{MBSP}$ did not destabilise AtTCP13, consistent with the destabilisation assays presented in **Figure 5.2.4.A.** Surprisingly, $SAP11_{SPLL}$ seems to destabilise both HA-TCP2 and HA-TCP13, as shown in the anti-HA western blot, whereas $SAP11_{SPLL}$ did not show interaction with the two Class II CIN-TCP members in yeast. This result could be either a direct destabilisation of Class II CIN TCPs or an indirect destabilisation resulting from the destabilisation of a negative regulator of class II CIN-TCPs, such as Class I TCPs. I used other constructs that produce 4xMyc-tagged SAP11 effector homologs (**Figure 5.2.4.B**) and found that $SAP11_{SPLL}$ did not destabilise AtTCP13. I need to repeat the protoplast degradation assay for $SAP11_{SPLL}$. On the other hand, $SAP11_{WBDL}$ did not destabilise Class II CIN-TCP2 nor TCP13, consistent with the previous Yeast-Two Hybrid assays.

Finally, the degradation assay shows that $SAP11_{SPLL}$ destabilised Class I TCP14, consistent with the Yeast-Two Hybrid assays. However, I was not able to detect $SAP11_{SPLL}$ using the anti-GFP antibody. $SAP11_{AYWB}$, $SAP11_{MBSP}$ did not destabilise TCP14. $SAP11_{WBDL}$ appears to destabilise weakly TCP14 compared to the control. In yeast, Class I interactor $SAP11_{WBDL}$ interacted with only Class I TCP6 but not Class I TCP14. The degradation assay of TCP14 suggests that $SAP11_{WBDL}$ might bind weakly Class I TCP14. It would have been interesting to test the degradation of Class I TCP6 by $SAP11_{WBDL}$.

To conclude, the patterns of TCP destabilisation by $SAP11_{AYWB}$ and $SAP11_{MBSP}$ are consistent with the patterns of interaction in yeast, confirming my results.

$SAP11_{AYWB}$ bind and destabilised both (sub)classes of Class II TCPs but not Class I TCPs, while $SAP11_{MBSP}$ bind and destabilised only Class II CYC/TB1 TCP. I cannot strictly confirm that $SAP11_{SPLL}$ and $SAP11_{WBDL}$ destabilised TCP proteins at this stage because I did not detect the effectors using anti-GFP. However, the destabilisation patterns of Class II CYC/TB1-TCP12, and Class I TCP14 are consistent with the Yeast-Two Hybrid assays for these two SAP11 effector homologs.

5.2.5 The intra-region and the TCP binding domain of $SAP11_{AYWB}$ are required for the binding specificity to CIN-TCPs

The interaction and destabilisation assays of SAP11 effector homologs toward the TCP proteins suggest that there is a region within SAP11 that provides the binding specificity toward the TCP (sub)classes. To determine this SAP11 region, I generated SAP11 chimaera constructs based on $SAP11_{AYWB}$ (able to bind to Class II CIN-TCP and Class II CYC/TB1-TCPs) and $SAP11_{MBSP}$ (able to bind to Class II CYC/TB1 only) and tested their binding to class II CYC/TB1-TCP18 and CIN-TCP2 in yeast. My lab previously showed that the 'KEEGSSSKQPDDSKK' sequence located at the C-terminus of $SAP11_{AYWB}$ did not affect the binding of this protein to TCP when deleted, but the deletion of the additional 'MEILKQKAEEETKNL', predicted to form coiled-coil structure, did (Sugio et al., 2014; Pecher et al., 2019). Thus, the region 'MEILKQKAEEETKNL' is required for TCP-binding and is called here "TCP binding domain".

The Y2H results of the chimaera constructs show that at least both the intra-region and the TCP binding domain of $SAP11_{AYWB}$ provided the binding specificity to the Class II CIN-TCPs (**Figure 5.2.5**). First, the control shows that $SAP11_{AYWB}$ interacted with both Class II CYC/TB1-TCP18 and CIN-TCP2 while $SAP11_{MBSP}$ interacted only with CYC/TB1-TCP18 in yeast, on the selective media SD-LWH (20mM 3AT). The addition of the intra-region + TCP binding domain of $SAP11_{AYWB}$ in the backbone of $SAP11_{MBSP}$ restored the interaction with Class II CIN-TCP2. However, the addition of the $SAP11_{AYWB}$ intra-region alone in the $SAP11_{MBSP}$ backbone did not show interaction with Class II CYC/TB1-TCP18 or Class II CIN-TCP2. We cannot exclude the possibility that the intra-region only can provide the interaction with the Class II CIN-TCPs as

there was no interaction with the positive control Class II CYC/TB1-TCP18. Moreover, the addition of the SAP11$_{AYWB}$ TCP binding domain alone in the SAP11$_{MBSP}$ backbone did not restore interaction with Class II CIN-TCP2 but did show interaction with Class II CYC/TB1-TCP18. The N-terminal and C-terminal region of SAP11$_{AYWB}$ are not required for the binding specificity towards Class II CIN-TCP2 (Supplementary data, Pecher et al., 2019).

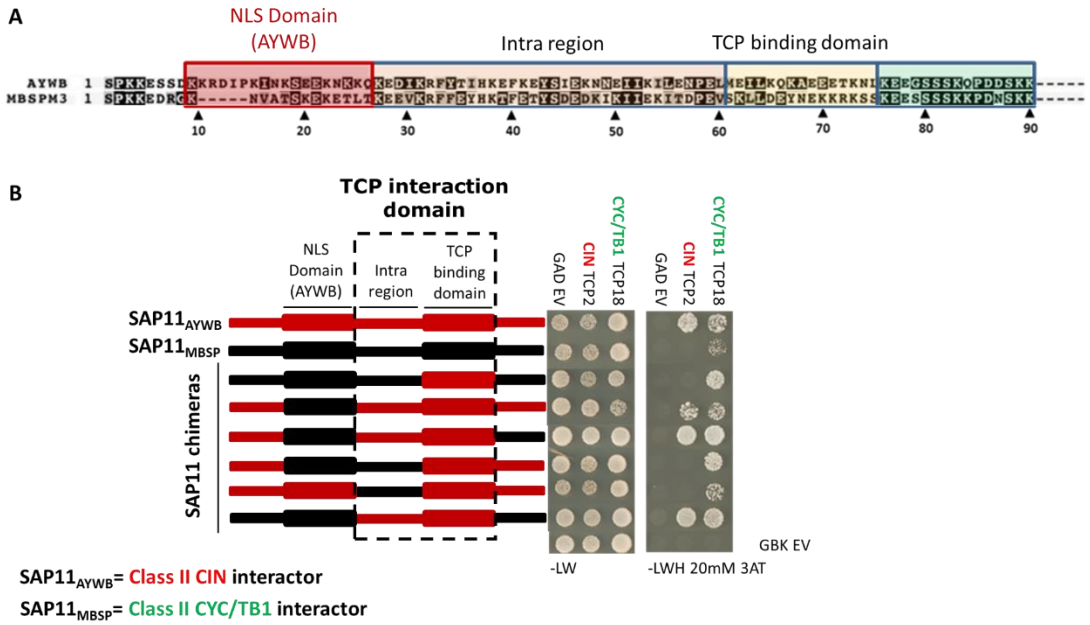


Figure 5.2.5: The intra-region and the TCP binding domain of SAP11$_{AYWB}$ are required for the binding specificity to CIN-TCPs. Yeast two-hybrid analysis of phytoplasma SAP11 chimeric versions based on SAP11$_{AYWB}$ and SAP11$_{MBSP}$ and *Arabidopsis thaliana* TCPs. TCP candidates selected from the Class II TCP (sub)classes, Class II CIN-TCP and Class II CYC/TB1-TCPs. SAP11 chimeric proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmid (pDEST-GADT7) combinations were grown in synthetic dropout (SD) media lacking either leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; or leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs or SAP11 chimaeras) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

Finally, the NLS domain of $SAP11_{AYWB}$ is not responsible for the binding specificity to Class II CIN-TCP2, as the construct bearing this domain in the backbone of $SAP11_{MBSP}$ did not show interaction with AtTCP2 or AtTCP18 in yeast (Pecher et al., 2019) (see **Appendix IV – Figure 1**)

To conclude, these results show that the intra-region and the TCP binding domain are both required for the binding specificity towards Class II CIN-TCPs. I called this region the TCP interaction domain. The NLS domain of $_{SAP11AYWB}$ is required for the targeting of the nucleus (**Figure 5.2.3**) but is not responsible for the binding specificity.

5.2.6 The intra-region and the TCP binding domain of $SAP11_{SPLL}$ are required for the binding specificity to Class I TCPs.

As I established the SAP11 region that provides the binding specificity toward Class II CIN-TCPs, I wished to investigate if the same region within Class I interactors was also responsible for the binding specificity toward Class I TCPs. The Y2H and destabilisation assays suggest that there is a tendency to bind to either Class II CIN or Class I TCPs. These patterns of interaction suggest that the same region within SAP11 could be responsible for binding to either Class II CIN or Class I TCPs. To test this, I generated SAP11 chimaeras constructs based on Class I interactor $SAP11_{SPLL}$ (able to bind to Class I and Class II CYC/TB1-TCPs) and $SAP11_{MBSP}$ (able to bind to Class II CYC/TB1 only) and tested their binding to Class II CYC/TB1-TCP18 and Class I TCP9 and TCP14 in yeast.

The Y2H results of the chimaeras constructs show that at least both the intra-region and the TCP binding domain of $SAP11_{SPLL}$ provides the binding specificity to the Class I TCP (**Figure 5.2.6**). First, the control shows that $SAP11_{AYWB}$ interacted with both Class II CYC/TB1-TCP18 and CIN-TCP2, $SAP11_{MBSP}$ interacted only with CYC/TB1-TCP18 while $SAP11_{SPLL}$ interacted with Class II CYC/TB1 TCP18 and Class I TCP9 and TCP14 in yeast, on the selective media SD-LWAH. When adding the intra-region + TCP binding domain of $SAP11_{SPLL}$ in the backbone of $SAP11_{MBSP}$ (SAP11 chimaeras B, **see Appendix IV – Figure 2**), the interaction with Class I TCP candidates AtTCP9 and AtTCP14 was restored. The addition of the $SAP11_{SPLL}$ intra-region alone in the

$SAP11_{MBSP}$ backbone does not show interaction with Class I TCP9 or TCP14 but still interacts with Class II CYC/TB1-TCP18.

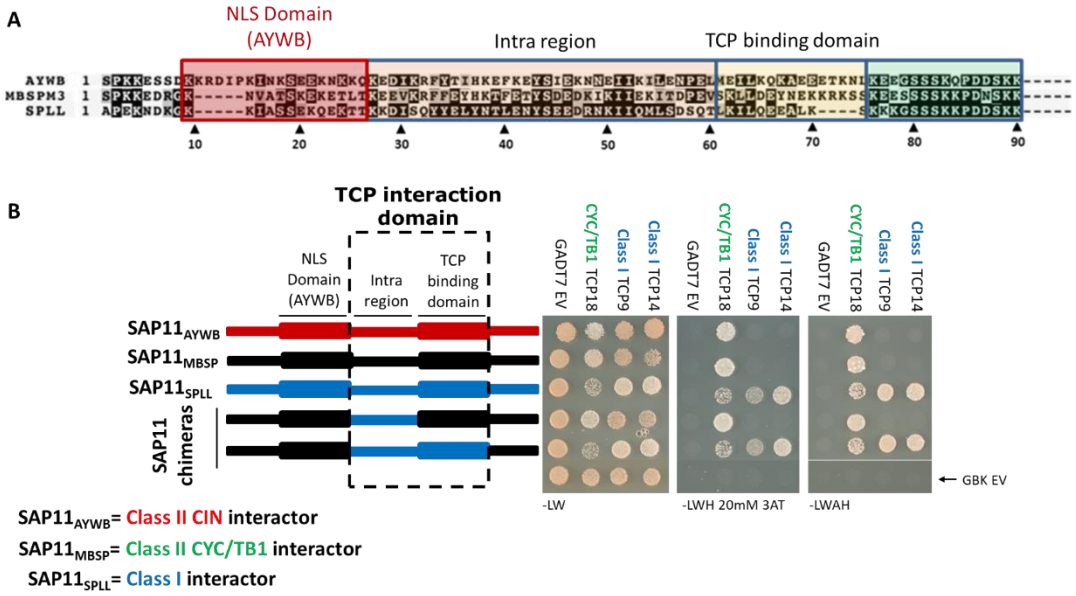


Figure 5.2.6: The intra-region and the TCP binding domain of $SAP11_{SPLL}$ are required for the binding specificity to Class I TCPs. Yeast two-hybrid analysis of phytoplasma SAP11 chimeric versions based on $SAP11_{SPLL}$ and $SAP11_{MBSP}$ and *Arabidopsis thaliana* TCPs. TCP candidates selected from the Class II CYC/TB1-TCP (sub)class and the Class I TCP. SAP11 chimeric proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmid (pDEST-GADT7) combinations were grown in synthetic dropout (SD) media lacking either leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (middle) or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs or SAP11 chimaeras) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

These results show that the same region within SAP11 provides the binding specificity to either Class II CIN TCPs or Class I TCPs. This region is also called TCP interaction domain. As SAP11 effector homologs all target the Class II CYC/TB1-TCPs, it is not possible to determine the binding specificity towards this (sub)class using

SAP11 chimaeras. The combined results of the SAP11 chimaeras study (**Figure 5.2.5** and **5.2.6**) led me to further investigate the residues responsible for the binding specificity of the TCP (sub)classes within the TCP interaction domain.

5.3 Discussion

This chapter complements the phylogenetic analysis of Chapter 4 and offered essential insights into the function of the different SAP11 effector homologs of the phytoplasma isolates. First, I showed that the SAP11 effector homologs from each *SAP11* clade (see **Figure 4.3.1**) have distinct patterns of interaction with the TCP (sub)classes in yeast. The patterns of interaction were mainly confirmed for four SAP11 effector homologs in protoplast degradation assays (see 5.3.1). Coupled with the radial phylogenetic tree of the SAP11 phylogeny, I confirmed the hypothesis that the ***SAP11* gene has evolved to target the TCP (sub)classes differentially.** Some SAP11 effector homologs bind to class I TCPs, which was not previously known. The distinct binding ranges of SAP11 effector homologs led to functional characterisation of the SAP11 region that binds specifically to either class I or class II CIN-TCPs. Therefore, I also confirmed the hypothesis that **SAP11 has a specific region that provides the binding specificity toward the class II CIN and class I TCP(sub)classes.**

5.3.1: Analysis of the Y2H and degradation assays

The main Y2H presented in this chapter shows clearly that the SAP11 effector homologs target the TCP (sub)classes differentially (see 5.3.4 for an overview of the results). First, the four SAP11 effector homologs used in Chapter 3 are good controls. The addition of class I TCPs in the analysis enhances the distinction between the patterns of interaction. Indeed, $SAP11_{AYWB}$ interacts with class II CIN-TCPs but does not show interaction with any members of class I TCP, while it is the opposite for $SAP11_{SPLL}$ and $SAP11_{WBDL}$.

These four SAP11 effector homologs are therefore good candidates for further analysis via the *in planta* degradation assays. The destabilisation assay in *A. thaliana* protoplasts globally confirms the interaction patterns of the Y2H assay. First, $SAP11_{AYWB}$ destabilises both class II CYC/TB1-TCP12 and class II CIN-TCP2 and TCP13

while SAP11$_{MBSP}$ destabilises only class II CYC/TB1-TCP12, consistent with the Y2H results and previous degradation assays (Pecher et al., 2019). However, the patterns of destabilisation of both SAP11$_{WBDL}$ and SAP11$_{SPLL}$ are more challenging to discern. Indeed, I did not detect either SAP11$_{WBDL}$ or SAP11$_{SPLL}$ using the anti-GFP antibody (**Figure 5.2.4.A**). This can be explained by the weaker proportion of transformed protoplasts (**Figure 5.2.3**). Thus, I cannot strictly confirm that the effector homologs destabilise the TCP (sub)classes. However, a previous study showed that SAP11$_{PnWB,}$ which is almost identical to SAP11$_{SPLL}$ binds and destabilises class II CYC/TB1-TCP (Chang et al., 2018). Therefore, this is most likely that both effectors or at least SAP11$_{SPLL}$ retain the function to bind and destabilise TCPs, similarly to other SAP11 effector homologs.

Nonetheless, using the anti-HA antibody, I showed that TCP candidates appear to be destabilised in the presence of SAP11$_{SPLL}$ or SAP11$_{WBDL}$ compared to the control. The patterns of destabilisation are consistent with Y2H assays concerning class II CYC/TB1 and class I TCPs. For example, class II CYC/TB1-TCP12 appears to be destabilised by SAP11$_{SPLL}$ and SAP11$_{WBDL}$. This is also true for class I TCP14, which appears to be destabilised by SAP11$_{SPLL}$ but not SAP11$_{WBDL}$, consistent with the Y2H assay.

Surprisingly, class II CIN-TCP2 and class II CIN-TCP13 appear to be destabilised in the presence of SAP11$_{SPLL}$ (**Figure 5.2.4.A**), which was not shown in yeast. However, using a different construct, I did not find that SAP11$_{SPLL}$ destabilises class II CIN-TCP13 (**Figure 5.2.4.B**). The next step should aim at repeating the transient protoplast degradation assays for SAP11$_{SPLL}$ and SAP11$_{WBDL}$ or perform co-expression assays in *Nicotiana benthamiana* leaves. Nevertheless, the apparent destabilisation of Class II CIN-TCP members in the presence of SAP11$_{SPLL}$ could be due to direct or indirect interaction. Class I and class II CIN-TCPs have antagonistic roles in plant development and plant defence through the regulation of the JA pathway (Kosugi and Ohashi, 2002; Li et al., 2005; Martín-Trillo and Cubas, 2010; Danisman et al., 2012; Lopez et al., 2015) (see Discussion point 5.3.3). Although, there is no evidence that class I form heterodimers with class II CIN-TCPs (Kosugi and Ohashi, 2002; Danisman et al., 2012),

the apparent destabilisation of class I TCPs could be indirect and due to the destabilisation of class II CIN-TCPs or other targets.

Finally, I opted for the destabilisation assay in *Arabidopsis thaliana* protoplasts over transient expression in *Nicotiana benthamiana* leaves as it allows me to correlate the findings with insect assays (no insect assays are shown in this thesis but some were performed during my PhD), that are also performed in *Arabidopsis thaliana* (Sugio et al., 2011a; Pecher et al., 2019). Furthermore, my lab generated transgenic lines that stably express $SAP11_{AYWB}$ and $SAP11_{MBSP}$, as well as mutant lines of *brc1 x brc2* or *35S::miR319a x 35S::miR3TCP* transgenic *A. thaliana* (Col-0) line that leads to a downregulation of CIN-TCPs (Pecher et al., 2019). These lines could be used for transient protoplast assays. For example, I could use the *35S::miR319a x 35S::miR3TCP* stable transgenic *A. thaliana* (Col-0) line and assess if $SAP11_{SPLL}$ retains its ability to destabilise class I TCPs.

5.3.2: The Ancestral Sequence Reconstruction as another approach to determine the key residues involved in the binding specificity toward the TCP (sub)classes

I used the Ancestral Sequence Reconstruction (ASR) method as another approach in my PhD but did not include them. The ASR analyses modern sequences that constitute a phylogenetic tree in order to resurrect the ancestral gene at particular nodes of that given tree (Thornton, 2004). The resurrection of an ancestral gene relies on various steps (**Figure 5.1.1;** Thornton, 2004). First, a well supported phylogenetic tree is inferred from a Multiple Sequence Alignment (MSA) of either protein or DNA sequences. Most of the time, the inferred tree is based on protein sequences (Thornton, 2004). Secondly, an ancient protein of a node in the tree is reconstructed *in-silico* via an algorithm that establishes the most probable residue at each position of the ancient protein sequence, based on the MSA and the phylogenetic tree. Thirdly, the ancient gene is "resurrected" *de novo* via gene synthesis, and then cloned into the appropriate vector for further molecular studies (Thornton, 2004).

I used the ASR to generate *in-silico* the most probable SAP11 sequences at each node of the SAP11 tree and test their interaction against TCP candidates in

yeast. I initially used the ASR approach based on a SAP11-tree generated via a smaller pool of eight SAP11 sequences, before I gained additional SAP11 sequences from collaborators. Back then, the initial objective was to determine the interaction ability of the ancestral SAP11 toward its TCP targets. Based on the observation that the four SAP11 effector homologs of Chapter 3 ($SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$) interact with Class II CYC/TB1 TCPs, the hypothesis was that the ancestral SAP11 effector protein could interact with Class II CYC/TB1 TCPs. Furthermore, the approach would allow investigating if $SAP11_{AYWB}$ acquired the ability to target *A. thaliana* Class II CIN-TCPs, an ability that the three other SAP11 effector homologs do not have. I would then synthesise the ancestral sequence located at the node between $SAP11_{AYWB}$ and $SAP11_{MBSP}$ and test the interaction with TCP (sub)classes using the Y2H system.

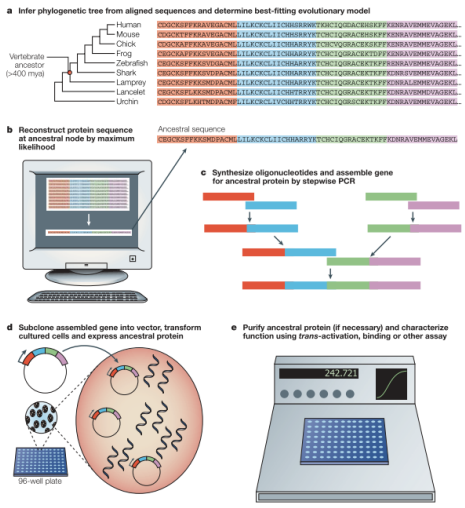


Figure 5.3.1: the steps for the Ancestral Sequence Reconstruction. See the review of Thornton et al., 2004. *Nature Reviews Genetics*. for more details. Figure taken from this review.

Finally, this would allow me to gain insights into the key residues or the region within SAP11 that provides the binding specificity towards TCP (sub)classes, as the SAP11 ancestral sequences are generated by computing the most probable residue at each position of the sequence, based on the inferred SAP11 Multiple Sequence Alignment and SAP11 phylogenetic tree. For example, Dong et al. (2014) used the ASR to resurrect the ancestral EPIC1 gene and test its interaction patterns with the target along the evolution of the effector. The ASR was also used to reconstruct the

predicted ancestral sequence of Tin2 effector and determine the functional evolution of the effector (Zess et al., 2019; Tanaka et al., 2019).

The ASR is a reliable approach; however, it presents some drawbacks. The major issue is that most of the time, the algorithms used for the resurrection might not take into account the biological relevance (Randall et al., 2016). The ASR approach did not yield much success in my case, as we do not know the SAP11 gene that behaves as an outgroup. It is therefore hard to determine confidently ancestral sequences.

5.3.3: The potential role of degradation of class I TCPs by SAP11 effector homologs

One of the novel findings of this chapter is that SAP11 effector from some phytoplasma isolates can bind to class I TCPs. This result is exciting considering the extensive role of class I TCPs in both plant development and plant immunity (see Chapter 1 General Introduction). The antagonistic regulation between class II and class I TCPs over the bioproduction of JA and the positive regulation of SA by class I TCPs provide a prime opportunity for a pathogen to target TCPs as it leads to a disruption of the hormonal cross-talk (Lopez et al., 2015; Garner et al., 2016; Nicolas and Cubas, 2016).

Class I TCPs were found to positively regulate the salicylic acid as at least TCP8 and TCP9 regulate the *ISOCHORIS- MATE SYNTHASE 1* (*ICS1*) gene involved in the biosynthesis of SA (Pieterse et al., 2012; Wang et al., 2015b).

JA production confers resistance to necrotrophic pathogens while the SA production enhances the resistance against biotrophic pathogens (Kazan and Lyons, 2014). The activation of the JA pathway attenuates the SA pathway and vice-versa. On the one hand, a biotrophic pathogen will manipulate the plant to activate JA in order to negatively regulate cell death and extend the life cycle and/or negatively regulate SA for the same purpose. On the other hand, a necrotrophic pathogen will aim at increasing the production of SA to promote the cell death and aim at limiting JA (Kazan and Manners, 2012; Xin and He, 2013; Kazan and Lyons, 2014).

In the case of phytoplasma, as a biotrophic pathogen, the stable expression of *SAP11$_{AYWB}$* in *A. thaliana* increases the fertility of *M. quadrilineatus* nymphs when the leafhoppers are feeding on SAP11 transgenic lines (Sugio et al., 2011a). The increase of the leafhopper-fertility is due to the destabilisation of Class II CIN-TCPs, leading to the downregulation of the *LOX2* gene and therefore the jasmonic acid synthesis (Bell et al., 1995; Schommer et al., 2008). The downregulation of Class I TCPs by SAP11$_{SPLL}$ might lead to an increase of the JA, leading to disruption in the hormone cross-talk of the plant (Kazan and Lyons, 2014; Yang et al., 2017), which could increase the phytoplasma or insect vectors virulence.

This strategy would be similar to the strategy employed by *P. syringae*. Indeed, HopBB1 effector interacts with TCP14 and JAZ3, two repressors of JA signalling, glues them together and are then degraded, leading to a fine-tuned activation of specific subsets of JA response, increasing the bacterial virulence (Yang et al., 2017). To test if SAP11$_{SPLL}$ or SAP11$_{WBDL}$ effectors activate subsets of JA response via destabilisation of Class I TCPs, I could compare the expression of marker genes in the JA pathway between *35S::SAP11$_{SPLL}$ A. thaliana* lines and Col-0 ecotype. I can also assess if SA-specific marker genes are down-regulated in response to the activation of JA.

We also know that class I TCP members are involved in the plant defence (Mukhtar et al., 2011; Kim et al., 2014; Wang et al., 2015b). TCP8, TCP14 and TCP15 are positive regulators of ETI through transcription regulation of defence genes. Indeed, triple mutants *tcp8*, *tcp14*, *tcp15* have a decreased ETI mediated by resistance genes RPS2, RPS4, RPS6 and RPM1 (Kim et al., 2014). Class I TCPs are also positive regulators of PTI components, through the regulation of the EFR pattern recognition receptor (Spears et al., 2019).

Targeting TCP13, TCP14, TCP15, TCP19, and TCP21 through effectors is a common strategy for very diverse pathogens. An interactome study showed that effector repertoires from *P. syringae* (Psy; eubacteria), *H. arabidopsidis* (Hpa; oomycete), and *Golovinomyces orontii* (Go; ascomycete)—converge specifically onto

TCP14 (Dreze et al., 2011; Mukhtar et al., 2011; Weßling et al., 2014) (see General Discussion).

Future work could aim at assessing if phytoplasma or insect vectors would see an increase of fitness in the presence of Class I interactors and test if ETI or PTI components are affected upon stable expression of *SAP11*$_{SPLL}$ or *SAP11*$_{WBDL}$. First, I could assess if phytoplasma is more virulent when infecting *A. thaliana* lines that stably produce SAP11$_{SPLL}$ or SAP11$_{WBDL}$. For this, I could measure the phytoplasma titer of infected *35S::SAP11*$_{SPLL}$ *A. thaliana* lines, infected *35S::SAP11*$_{WBDL}$ *A. thaliana* lines, infected *Class I tcp* mutant lines and compare it with infected Col-0 Wild Type control. I could measure the relative expression of *EFR* gene to assess if the PTI is hindered. I could also measure the RNA levels of defence genes related to ETI that are known to be regulated by Class I TCPs (Kim et al., 2014).

Secondly, I could assess if the insect vector sees an increase of its fertility when feeding and laying its eggs on *35S::SAP11*$_{SPLL}$ or *35S::SAP11*$_{WBDL}$ *A. thaliana* lines. I could perform fertility insect assays using *A. thaliana* lines that stably produce Class I interactors and test the fertility of *M. quadrilineatus*. As *M. quadrilineatus* produce more progeny when the level of JA is down upon production of SAP11$_{AYWB}$ and as SAP11$_{SPLL}$ destabilises Class I TCPs, I am expecting *M. quadrilineatus* to have no increase of fertility, on the contrary to SAP11$_{AYWB}$ (see the general discussion for details about different strategies of phytoplasma).

Finally, more experiments are needed to determine if the TCP destabilisation by SAP11 is dependent on the proteasome. A study showed that SAP11 effector homolog from wheat blue dwarf phytoplasma, SWP1, interacts with class II CYC/TB1 through a proteasome system but not through proteases (Wang et al., 2018b). To test if the destabilisation is proteasome-dependent, I could add a proteosome inhibitor in protoplast degradation assays. Additionally, does SAP11 require plant helper for the destabilisation of TCPs? Future global Y2H assays could indicate potential plant partners of SAP11. As an example, it was found that SAP54, an additional SAP effector of AY-WB phytoplasma, interacts with RAD23 (Farmer et al., 2010; MacLean et al., 2014).

SAP11 specifically binds the hlh motif of the TCP domain, which is involved in the homo/heterodimerisation of the TCPs (Chapter 3) (Aggarwal et al., 2010; Viola et al., 2012; Danisman et al., 2013). TCPs are stable in dimers and not stable in monomers (Aggarwal et al., 2010). The instability of the TCP monomers could lead to their degradation. Therefore, it is possible that SAP11 does not require to interact with proteasome components.

5.3.5: The same region within SAP11 provides binding specificity to either class II CIN-TCPs and Class I TCPs

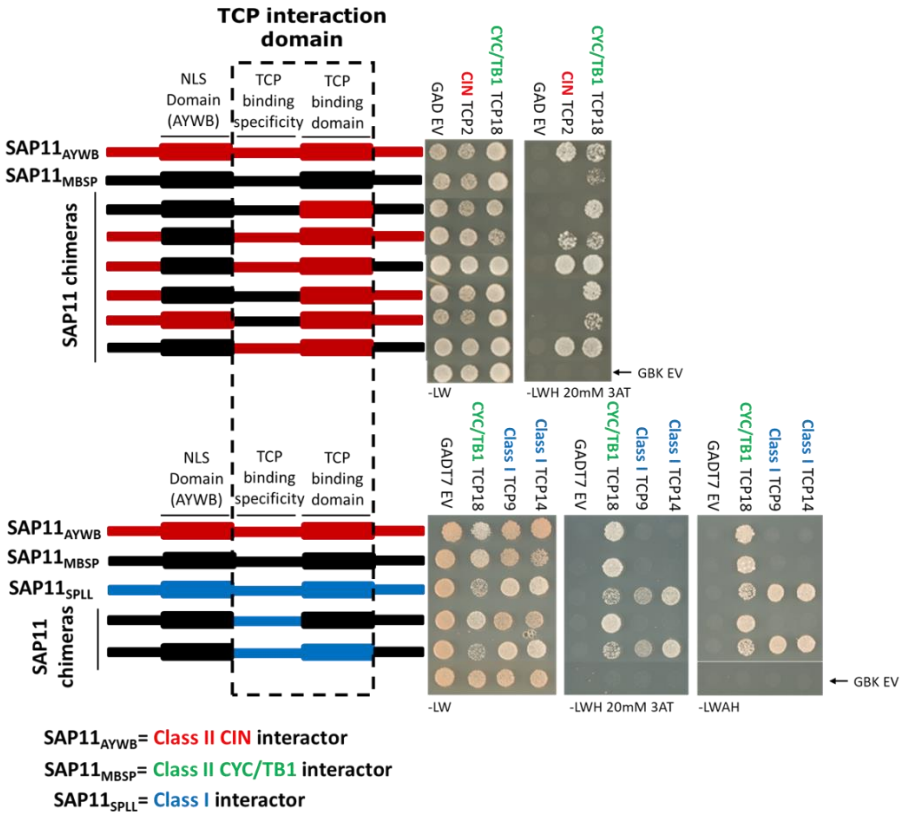


Figure 5.3.2: The TCP interaction domain of SAP11. Yeast two-hybrid analysis of phytoplasma SAP11 chimeric versions based on $SAP11_{AYWB}$, $SAP11_{SPLL}$, $SAP11_{MBSP}$ different SAP11 effector homologs and *Arabidopsis thaliana* TCPs. Yeast-Two-Hybrid showing the interaction between SAP11 chimeric versions based on $SAP11_{AYWB}$, $SAP11_{SPLL}$, $SAP11_{MBSP}$. TCP candidates selected from the different (sub)classes of TCPs, Class I, Class II CIN and Class II CYC/TB1. SAP11 chimeric proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4

(prey). Yeast colonies co-expressing bait and prey or empty plasmid (pDEST-GADT7) combinations were grown in synthetic dropout (SD) media lacking either leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; or leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (top right for SAP11 chimaeras analysis using $SAP11_{AYWB}$ and $SAP11_{MBSP}$ or bottom middle right for SAP11 chimaeras analysis using $SAP11_{SPLL}$ and $SAP11_{MBSP}$) or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (bottom right) with growths of colonies indicating interactions of bait (SAP11 homologs or SAP11 chimaeras) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

The **Figure 5.3.2** summarises the SAP11 chimaera results. The same region within SAP11 provides binding specificity to Class II CIN TCPs and Class I TCPs. This is consistent with the analysis of the SAP11-based radial tree (**Figure 5.3.3**) that suggest that, based on the SAP11 clades and their interaction patterns, the same region is responsible for the binding specificity of the two (sub)classes. The "TCP interaction" domain does not include the NLS domain, present at the N terminal part of the Class II CIN-interactors such as $SAP11_{AYWB}$ or at the C terminal part of the Class I interactors such as $SAP11_{SPLL}$. The position of the NLS domain within SAP11 does not impact the nuclear localisation of the SAP11 effector. Indeed, my results suggest that $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ localise in the nucleus of Arabidopsis protoplasts (**Figure 5.2.3**). Although I did not include a specific nucleus marker, my results are consistent with previous studies that showed the nuclear localisation of different SAP11 effector homologs from three distinct SAP11 clades; $SAP11_{AYWB}$ (Bai et al., 2009), $SAP11_{OY\text{-}M}$ and $SAP11_{ATP}$ (Chang et al., 2018).

I initially hypothesised that the intra-region of $SAP11_{AYWB}$ could provide the binding specificity toward class II CIN-TCPs, as the region is more diverse than the TCP binding domain (see Multiple Sequence Alignment of SAP11 effector homologs in Chapter 4) (Sugio et al., 2014). However, I could not detect yeast growth in the selective media when yeast co-expresses class II *TCP2* and a *SAP11* chimeric version based on the backbone of $SAP11_{MBSP}$ and harbouring $SAP11_{AYWB}$ intra-region (see **Appendix IV – Figure 1**). The absence of yeast growth suggests that there are

structural requirements between the intra-region and the TCP binding domain and both need to be present for the binding specificity. Alternatively, misfolding of the chimeric SAP11 in yeast could be possible.

Given the results of the SAP11 chimaera Y2H assays of **Figure 5.2.5**, I decided to determine further the residues involved in the binding specificity toward the TCP (sub)classes within the TCP interaction domain. To do this, I designed new SAP11 chimaera constructs based on the sequence variation between Class I interactors and SAP11 sequences that are not found to bind to Class I TCPs (see **Appendix IV – Figure 2)**. The block A is part of the TCP interaction domain and is conserved among SAP11 sequences of Class I interactors (in blue) and SAP11 sequences of Class II CIN interactors (in red), suggesting that this region might be the region that binds specifically to either class.

The predicted NLS domain of $SAP11_{SPLL}$ is "KSKKKGSSSKKPDDSKK" (65^{th} residue to 81^{st} residue), located at the C-terminal part of the SAP11 sequence, similarly to $SAP11_{MBSP}$. Thus, this allowed me to test if the C-terminal NLS domain of SAP11 effector homologs could be responsible for the binding specificity toward a TCP (sub)class. However, I have not found any positive interaction between SAP11 chimaera constructs and the TCP candidates, except for the SAP11 construct exhibiting the $SAP11_{SPLL}$ intra-region in the backbone of $SAP11_{MBSP}$ and the construct carrying the intra-region and the TCP binding domain of $SAP11_{SPLL}$ in the $SAP11_{MBSP}$ backbone (figure presented in **Appendix IV – Figure 2**).

For future work, I could add the TCP interaction domain of $SAP11_{AYWB}$ in the $SAP11_{MBSP}$ backbone and perform co-immunoprecipitation to see if there is interaction in planta. Alternatively, I could perform protoplast degradation assay to assess if the TCP interaction domain is sufficient for the destabilisation of cognate TCP (sub)classes.

5.3.5: SAP11 homolog genes cluster based on their patterns of interaction with the different TCP (sub)classes

I combined the results of the SAP11 phylogeny (chapter 4), the Y2H and degradation assays (chapter 5) and the phenotypes (predicted or established) of the

stable expression of the *SAP11* effector homolog genes as indicated in the phylogenetic tree. Globally, the Y2H results, the protoplast assays (Chang et al., 2018; Pecher et al., 2019) the phenotypes of the cognate *SAP11* over-expression lines (Sugio et al., 2011a; Chang et al., 2018) and the phylogeny are consistent and confirm that the ***SAP11* effector homolog genes cluster based on their patterns of interaction with the different TCP (sub)classes.**

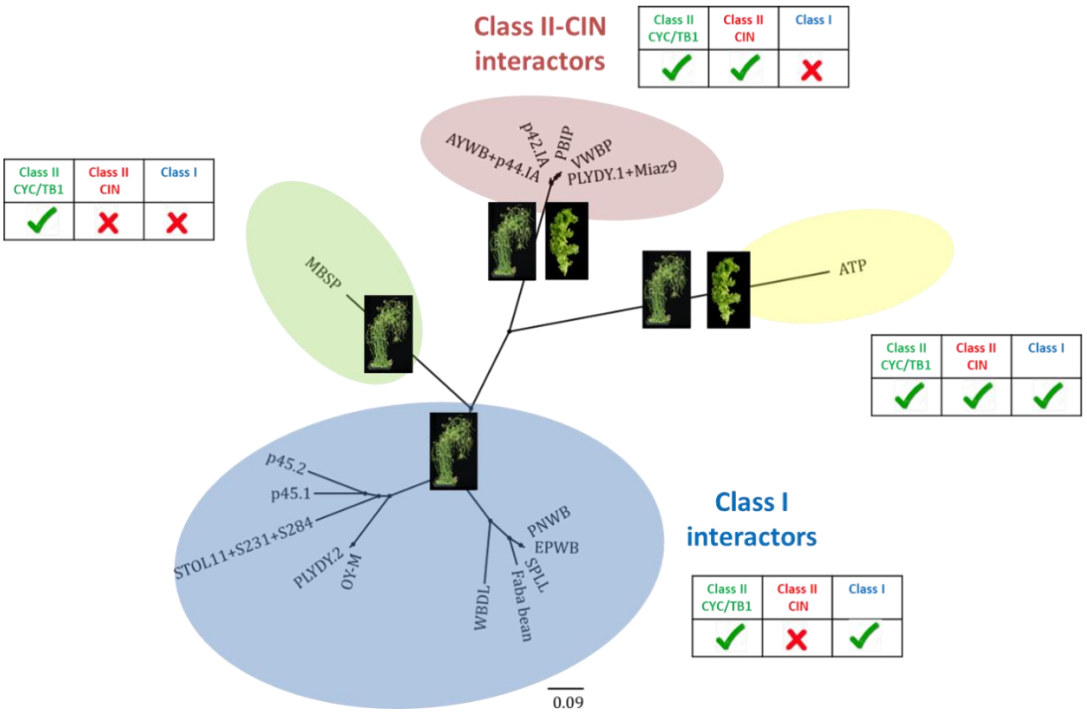


Figure 5.3.3: SAP11 homolog genes cluster based on their patterns of interaction with the different TCP (sub)classes. Radial phylogenetic tree based on the SAP11 homolog proteins from different phytoplasma isolates. The phylogenetic tree has been generated via the MEGA software using the Maximum Likelihood algorithm. The boxes indicate Y2H results. Based on the Y2H results, the protoplast assays, the phenotypes of the cognate SAP11 over-expression lines (Chang et al. 2018) and the phylogeny, different clades were established. The blue clade includes the SAP11 homologs that can interact with Class I TCPs, the red clade includes the SAP11 homologs that can interact with Class II CIN TCPs. Additionally, $SAP11_{MBSP}$ interacts only with CYC/TB1 (green circle) while $SAP11_{ATP}$ interacts with both Class I and Class II CIN (yellow). The pictures of *Arabidopsis thaliana* stable transgenic lines of the cognate SAP11 effector homologs as indicated in the phylogenetic tree are displayed for each clade. For example, $SAP11_{AYWB}$ stable expression in Arabidopsis lines lead to crinkled leaves and witches' broom symptoms

The clades established in Chapter 4 are now indicating the TCP range for each SAP11 subgroup. First of all, the SAP11 clades share the ability to interact with class II CYC/TB1-TCPs. The four SAP11 effector homologs $SAP11_{AYWB}$, $SAP11_{MBSP}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ from three distinct *SAP11* clades interact and destabilise members of class II CYC/TB1-TCPs. The stable expression of $SAP11_{AYWB}$ and $SAP11_{MBSP}$ in *A. thaliana* lines is consistent with this, as the increased stem is similar to the *brc1 brc2* double (Col-0) mutant line (Aguilar-Martinez et al., 2007; Pecher et al., 2019).

The red clade comprises closely related SAP11 effector homologs and $SAP11_{AYWB}$. The members of this clade share the ability to interact with class II CIN TCPs and class II CYC/TB1; however, they do not target class I TCPs. This was proved by Y2H and later complemented by protoplast degradation assays using $SAP11_{AYWB}$ as a candidate (Chapter 5). These patterns of interaction and destabilisation are reflected by the phenotype of the transgenic line that stably produces $SAP11_{AYWB}$ which shows crinkled leaves, phenocopying the phenotype of the *35S::miR319a x 35S::miR3TCP* line where CIN-TCPs are knocked down (Efroni et al., 2008; Sugio et al., 2011a; Pecher et al., 2019) and increased stem production (Aguilar-Martinez et al., 2007; Pecher et al., 2019).

The clade of the Class I interactors (blue) includes closely related SAP11 effector homologs that specifically interact with both Class II CYC/TB1 and Class I TCPs but not with Class II CIN TCPs. I predicted two distinct clades in Chapter 4 (**Figure 4.3.1**), but the two *SAP11* clades are combined into one through their ability to bind to class I TCPs in yeast. The protoplast degradation assays indicate but did not confirm that Class I interactors destabilise class I TCPs. I also showed that $SAP11_{SPLL}$ interacts specifically with the TCP domain of class I TCP members (**Figure 5.2.2**), which confirms that the SAP11 effector specifically interacts with the conserved TCP domain for all three (sub)classes. Although there are no transgenic lines that stably express $SAP11_{SPLL}$ or $SAP11_{WBDL}$, the phenotypes of the Arabidopsis lines that produce $SAP11_{OYM}$ or $SAP11_{PnWB}$ display increased stem production (Chang et al., 2018). There is little evidence that the *A. thaliana* transgenic lines that produce class I interactors such as $SAP11_{PnWB}$ and $SAP11_{OY\text{-}M}$ show symptoms that phenocopy the mutants of class I TCPs in *A. thaliana*, as class I TCPs are genetically redundant (Aguilar-Martínez

and Sinha, 2013). Clear phenotypes of class I TCP mutants require multiple knock-downs (Aguilar-Martínez and Sinha, 2013).

The clade of the Class II interactors (red) includes closely related *SAP11* gene homologs that can interact specifically with both Class II CYC/TB1 and Class II CIN TCPs but not with Class I TCPs. The interaction with either class II CIN or class I TCP is explained by the SAP11 chimaera study. The same SAP11 region provides the binding specificity toward either class I or class II CIN-TCP (see Discussion point 5.3.5).

The *SAP11$_{MBSP}$ and SAP11$_{ATP}$* genes form two distinct monophyletic groups and have unique abilities to interact with the TCP (sub)classes. Y2H analysis and protoplast degradation assays showed that SAP11$_{MBSP}$ binds and destabilises class II CYC/TB1 only but not interact class I or class II CIN-TCPs. This indicates that the sequence of *SAP11$_{MBSP}$* lacks the specific residues/region necessary for the binding to class I and class II CIN.

Surprisingly, SAP11$_{ATP}$ is the only SAP11 effector homolog that interacts with the three TCP(sub)classes in yeast. It would be important to test the destabilisation patterns of SAP11$_{ATP}$ toward the different TCP (sub)classes in *A. thaliana* mesophyll protoplasts. Another study showed that SAP11$_{ATP}$ did not destabilise class I TCP14 or TCP20 in Arabidopsis protoplasts, although the other class I TCP members have not been studied (Chang et al., 2018). Nonetheless, the paper found that SAP11$_{ATP}$ partially destabilised class II CIN-TCPs and also found that the stable expression of this effector homolog in *A. thaliana* triggers crinkled leaves, consistent with its ability to interact with class II CIN-TCPs (Chang et al., 2018). Additionally, a study found that SAP11$_{ATP}$ binds to class II CIN-TCP homologs MdTCP25 (AtTCP4 homolog) and MdTCP24 (AtTCP13 homolog) (Janik et al., 2017).

I initially hypothesised that SAP11$_{PLYDY.1}$ and SAP11$_{PLYDY.2}$ could be either duplicated genes or two versions that resulted from genetic transfer from another strain. I show here that SAP11$_{PLYDY.2}$ belongs to the same clade as SAP11$_{SPLL}$ and interacts with both class II CYC/TB1 and class I TCPs. SAP11$_{PLYDY.1}$ is almost identical to SAP11$_{AYWB}$ and therefore could interact with class II CIN and class II CYC/TB1, although I did not test it. It is highly probable that the two SAP11 effector homologs

have clear distinct patterns of interaction. On the other hand, sample p45 provides an example of gene duplication that results in two copies of *SAP11* that are very similar (see General Discussion about the evolution of SAP effectors).

To summarise, the hypothesis is confirmed, SAP11 effector homologs have evolved to bind and destabilise the TCP (sub)classes differentially. The analysis of the phylogenetic trees and the interaction assays establish the hypothesis that there is a SAP11 region that is responsible for the binding specificity to either Class I TCPs or Class II CIN-TCPs. This TCP interaction domain has been defined via chimaera studies and is comprised of both the intra-region and the TCP binding domain. The results presented in the last two chapters establish that some SAP11 effector homologs interact with Class I TCPs but $SAP11_{AYWB}$ cannot interact with Class I TCPs. This brings the hypothesis that AYWB phytoplasma may have additional effectors that interact with class I TCPs, extending the range of interaction with each TCP (sub)classes. The next chapter will address this hypothesis.

Chapter 6

SAP effectors from the SAP11$_{AYWB}$ genetic island interact with the TCP family, including Class I TCP (sub)class

6.1: Introduction

The two previous chapters show that *SAP11* effectors have evolved to interact with plant TCP (sub)classes differentially. The previous chapter shows that SAP11 effector homologs target all the three TCP (sub)classes: class I TCP, class II CIN-TCP and class II CYC/TB1-TCP. SAP11 effector homolog from AY-WB phytoplasma strain binds and destabilises class II CIN and class II CYC/TB1 leading to characteristic symptoms (Chapter 3 and Chapter 5) (Sugio et al., 2011a, 2014; Pecher et al., 2019). However, $SAP11_{AWYB}$ does not interact with class I TCPs, while other effector homologs such as $SAP11_{SPLL}$ do so. The two previous chapters suggest that the SAP11 ability to interact with the plant TCP family seems to be extended across the phytoplasma isolates. Considering the importance of the TCP in the plant development and plant defence and the phytoplasma ability to display its effector genes in PMU structures, allowing for potential horizontal gene transfer, it would be fair to presume that other SAP effectors could also target the TCP family.

The AY-WB phytoplasma genome was sequenced to completion in 2006 with 56 candidate effectors identified (Bai et al., 2006). The AY-WB genome contains four PMUs with $SAP54_{AYWB}$ and $SAP11_{AYWB}$ residing in PMU2 (Bai et al., 2006; Sugio and Hogenhout, 2012) (see also Chapter I General Introduction and Chapter 4 introduction). Aside from the characteristic genes that comprise PMU elements, multiple genes encoding for secreted proteins also reside in the four PMUs, including 41 out of the 56 predicted effectors (Bai et al., 2006, 2009). Interestingly, four genes encoding for secreted candidate effectors reside next to *SAP11* in PMU2 (Bai et al., 2006, 2009; Sugio and Hogenhout, 2012) (**Figure 6.1.1**). They were initially selected as candidate effectors because of the presence of signal peptide and the absence of additional transmembrane domain for each candidate (Bai et al., 2009). They are designated as SAP56, SAP66, SAP67 and SAP68.

A study from our lab showed that most SAP effectors are differentially expressed in either *A. thaliana* plant host or *M. quadrilineatus* insect host (MacLean et al., 2011). Along with SAP11, the four neighbouring effectors SAP56, SAP66, SAP67, SAP68 are up-regulated in *A. thaliana* compared to *M. quadrilineatus* at a

similar transcription level as they are in an operon-like region (MacLean et al., 2011). Based on this, the assumption would be that the four SAP candidate effectors are expressed inside the plant host and have a role in targeting plant components.

AY-WB phytoplasma infection modulates the insect-vector fitness when feeding to the plant host (Sugio et al., 2011b). Infected plants become more attractive to *M. quadrilineatus*, which is the main insect vector of the strain but also allow non-host insect vectors such as *D. maidis* to feed, reproduce and survive longer (Purcell, 1988; Kingdom and Hogenhout, 2007; Sugio et al., 2011b). Indeed, *D. maidis*, as a maize specialist, does not lay eggs on dicot plant species and usually die within a few days (Sugio et al., 2011b). The stable expression of SAP11$_{AYWB}$ in *A. thaliana* lines leads to dramatic changes in both the plant development via the characteristic symptoms and plant defence upon destabilisation of class II CIN-TCP members, resulting in the down-regulation of the *LOX2* gene and thus the JA synthesis (Schommer et al., 2008; Sugio et al., 2011a; Pecher et al., 2019a). The down-regulation of JA leads to the increase of fertility of *M. quadrilineatus* (Sugio et al., 2011a) but does not induce an increase in the survivability of *D. maidis*, suggesting that other SAP effectors could do so (Sugio et al., 2011b).

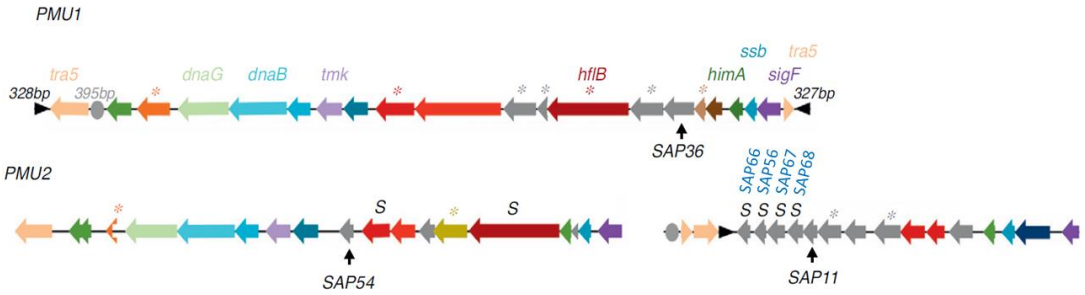


Figure 6.1.1: The PMU1 and PMU2 of the AY-WB phytoplasma strain. The black line indicates the AY-WB chromosome. ORFs are indicated with the coloured boxes and paralogous genes have the same colours. Grey boxes indicate unique genes. Within PMU2, SAP54 and SAP11 are indicated with black arrows and label. The letter S indicates candidate protein effectors, which latter were described as SAP56, SAP66, SAP67 and SAP68. The rest of the label describes the characteristic genes that feature PMUs (see General Introduction and introduction Chapter 4). Figure adapted from Sugio et al., 2012. *Current Opinion in Microbiology*.

A former PhD student of my lab previously studied SAP56, SAP66, SAP67 and SAP68 effectors to determine their potential role in leafhopper-Arabidopsis interactions (Kingdom, 2012). First, there is no visible change of phenotype when either of the four SAP effectors is stably expressed in *A. thaliana* lines compared to the wild-type. Nonetheless, she tested if the SAP effectors could induce a fertility boost in *M. quadrilineatus* or help *D.maidis* to survive longer in a non-host plant such as *A. thaliana*. As she noticed more variation in her insect assays compared to previous experiments, she included additional conditions, such as changes in temperature (Kingdom, 2012). She found that the stable expression of SAP67 or SAP68 in *Arabidopsis thaliana* lines induced the increase of *M. quadrilineatus* nymph production in cold conditions, but did not increase the survivability of *D.maidis* in *A. thaliana.*

On the other hand, the stable expression of SAP56 or SAP66 in *Arabidopsis thaliana* lines merely increased the fecundity of *M. quadrilineatus* but did increase strongly (34 and 37% respectively) the survivability of *D. maidis*. Although there is some variability in the increase of the fertility of *M. quadrilineatus*, her results confidently show that at least SAP56 and SAP66 increased the survivability of a non-host insect vector when overexpressed in *A. thaliana*. This result suggests that some of these SAP effectors target key plant components and modulate the plant host to the advantage of either the pathogen or insect vectors.

In parallel, our collaborator Dr Richard Imminck and his lab assessed if the phytoplasma effectors could target specific families of plant transcription factors in *A. thaliana* using a large-scale yeast-two hybrid screening (Correa Marrero and Capdevielle, unpublished). They screened 21 phytoplasma effectors, including SAP54 and SAP11 which were already found to interact with MADS-box transcription factors and TCP transcription factors respectively (MacLean et al., 2011; Sugio et al., 2011a, 2014; Pecher et al., 2019), and a library of *A. thaliana* transcription factors and transcription regulators (Pruneda-Paz et al., 2014). They found that the four effectors SAP56, SAP66, SAP67 and SAP68 interacted with members of plant TCPs.

Based on this, important questions arise. Do the SAP effectors from the same genetic island interact with plant TCPs using my yeast-two hybrid system? If yes, do they have distinct patterns of destabilisation with the TCP (sub)classes compared to SAP11$_{AYWB}$? Do the SAP effectors target the TCP domain in a similar fashion as SAP11 effector homologs?

The previous results establish the hypothesis that **SAP effectors from the same genetic island interact with plant TCPs and that they have a different interaction pattern than SAP11$_{AYWB}$, leading to new effects on the fitness of non-host insect vector *D.maidis*.**

I decided to test the interaction between the four SAP effectors and members of the TCP (sub)classes in yeast, using my established controls SAP11$_{AYWB}$, SAP11$_{MBSP}$, SAP11$_{WBDL}$ and SAP11$_{SPLL}$. I found that SAP56, SAP66 and SAP68 interacted with plant TCP (sub)classes, thus confirming the global yeast-two hybrid screening of our collaborators. Most importantly, I found that SAP56, SAP66 and SAP68 target class I TCPs in yeast. Then, I tested if the SAP effectors could target the TCP domain specifically, in a similar fashion than SAP11 effector homologs. Using TCP domain constructs of class I TCP members, I did not find a positive interaction between any of the SAP effectors and the TCP domain, suggesting that the SAP effectors target a different region of the TCP protein. I did not have time to test which TCP region is targeted by these SAP effectors during my PhD.

6.2: Results

6.2.1 SAP effectors from SAP11$_{AYWB}$ genetic island can interact with TCP, including Class I TCPs

Given the ability of SAP11 effector homologs to interact with TCP (sub)classes and based on the importance of the TCP family as a target of different effectors from different pathogens (Sugio et al., 2011a; Janik et al., 2017; Yang et al., 2017; Pecher et al., 2019), I wanted to extend my analysis and test if other phytoplasma effectors target the TCP family. The best candidates are four effectors from the model AY-WB phytoplasma strain. AY-WB SAP effector SAP56, SAP66, SAP67 and SAP68 belong to

the same genetic island as SAP11$_{AYWB}$, i.e., in the AY-WB Putative Mobile Unit PMU2 (**Figure 6.2.1 A**).

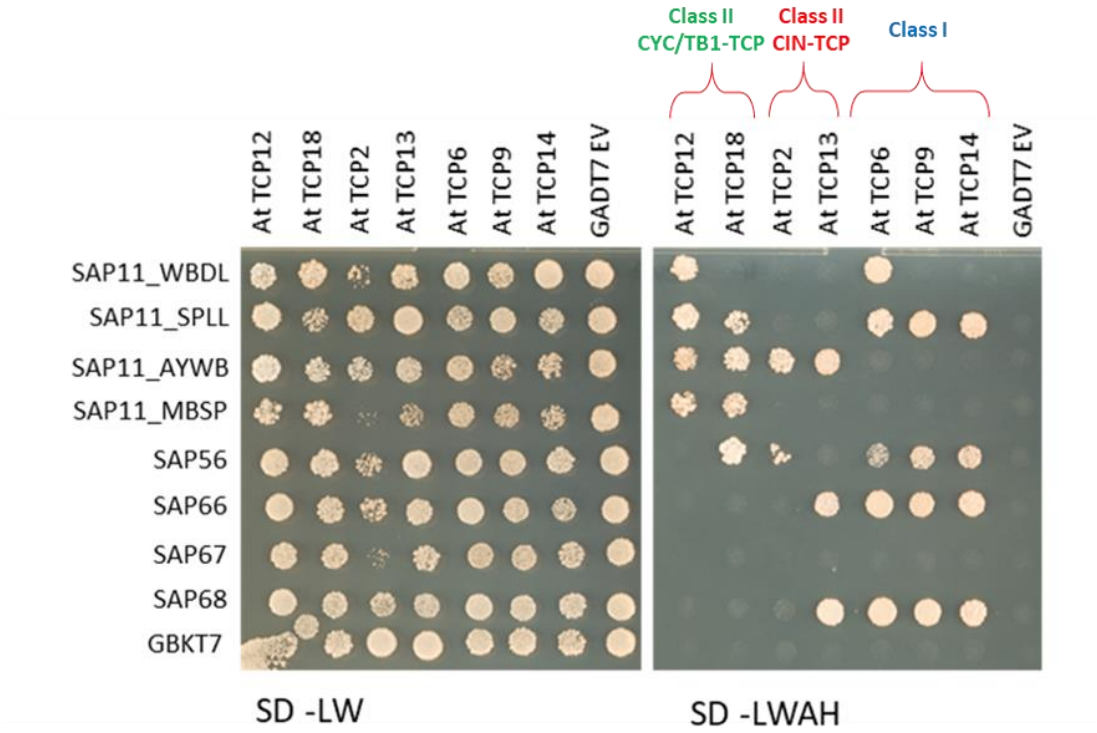


Figure 6.2.1: SAP effectors from SAP11$_{AYWB}$ genetic island can interact with TCP, including Class I TCPs. Yeast two-hybrid analysis of AY-WB phytoplasma SAP effectors SAP56, SAP66, SAP67 and SAP68, belonging to the genetic island of SAP11$_{AYWB}$ and *Arabidopsis thaliana* Class I, Class II CIN and Class II CYC/TB1-TCPs. SAP effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCP domains and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Phytoplasma SAP11 effector homologs SAP11$_{AYWB}$, SAP11$_{MBSP}$, SAP11$_{WBDL}$ and SAP11$_{SPLL}$ were used as positive control for interaction with the different TCP (sub)classes. Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking either leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SA homologs) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

The Y2H result shows that SAP56, SAP66 and SAP68 interacted with TCP proteins while SAP67 did not show interaction with any TCP candidate (**Figure 6.2.1 B**). First, the four SAP11 effector homologs bind TCP (sub)classes under the selective media SD-LWAH, according to their established interaction patterns (see previous Experimental Chapters). It is important to note that the yeast growth of the colonies overexpressing SAP56 effector and Class II CYC/TB1-TCP18 in the selective media SD-LWAH are due to contamination, as shown in another repetition in **Appendix V-figure 2**. The same is true for colonies expressing SAP56 and class II CIN-TCP2. SAP66 and SAP68 did show interaction with Class II CIN-TCP13.

Most interestingly, the three effectors SAP56, SAP66 and SAP68 interacted with Class I TCP6, TCP9 and TCP14. This result is interesting, as SAP11$_{AYWB}$ does not interact with Class I TCP members, but additional SAP effectors from the same genetic island can.

To conclude, additional SAP effectors also interacted with TCP members. Surprisingly, three SAP effector genes, *SAP56*, *SAP66* and *SAP68*, located in the genetic island of AY-WB phytoplasma *SAP11* gene, encode SAP effector proteins that bind to Class I TCPs, while SAP11$_{AYWB}$ cannot. This result suggests that the AY-WB phytoplasma strain can extend the binding range of SAP11$_{AYWB}$ toward the TCP (sub)classes via the secretion of additional SAP effectors that interact with Class I TCPs. The next step is to test if SAP56, SAP66 and SAP68 also target the TCP domain, similarly to SAP11$_{AYWB}$.

6.2.2 SAP effectors from SAP11$_{AYWB}$ genetic island cannot interact with the TCP domain of Class I TCPs, while SAP11$_{SPLL}$ can

Based on the previous result that shows that additional SAP effectors target the TCP family, I wished to investigate if they can also target the TCP domain, similarly to SAP11 homologs. This would indicate that the ability to target the TCP domain is a global strategy among phytoplasma isolates via the expression of multiple SAP effectors.

To do so, I used both Class I interactors $SAP11_{WBDL}$ and $SAP11_{SPLL}$ and the four SAP effectors SA56, SAP66, SAP67 and SAP68 and tested their interaction against the TCP domains of *A. thaliana* Class I TCP6, TCP8, TCP9, TCP14 and TCP15. The Yeast Two-Hybrid results show that $SAP11_{SPLL}$ targetted the TCP domain of Class I TCP6, consistent with the previous Yeast-Two Hybrid experiment (**Figure 6.2.2**). None of the four SAP effectors tested interacted with the TCP domain of Class I TCPs.

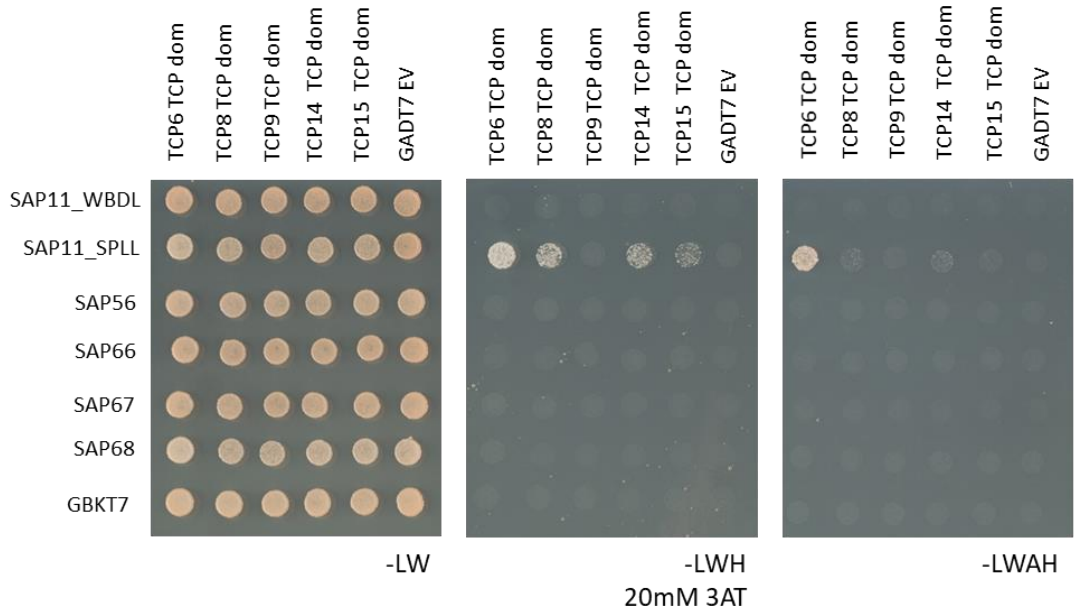


Figure 6.2.2: SAP effectors from $SAP11_{AYWB}$ genetic island does not interact with the TCP domain of Class I TCPs, while $SAP11_{SPLL}$ can. Yeast two-hybrid analysis of AY-WB phytoplasma SAP effectors SAP56, SAP66, SAP67 and SAP68, belonging to the genetic island of $SAP11_{AYWB}$ and the TCP domains of *Arabidopsis thaliana* Class I TCP6, TCP8, TCP9, TCP14 and TCP15. SAP effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCP domains and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmid (pDEST-GADT7) combinations were grown in synthetic dropout (SD) media lacking either leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; or leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (middle) or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs and other SAP effectors) and prey (TCP domains of TCP homologs). The experiment

was repeated three times with three different batches of transformed yeasts and these showed identical results as shown.

Taken together, these results suggest that SAP56, SAP66, SAP68 cannot bind to the TCP domain, unlike SAP11$_{AYWB}$. Additional experiments, such as co-immunoprecipitation between SAP effectors and the TCP domain of Class I TCPs would confirm the Yeast-Two Hybrid result. Future experiments need to be performed to investigate which TCP region is targeted by SAP56, SAP66 and SAP68.

6.3: Discussion

The results of this chapter show that three out of four effectors, SAP56, SAP66 and SAP68 from the same genetic island of SAP11$_{AYWB}$ do interact with TCP plant family in yeast. This result confirms the global yeast-two hybrid screening performed by collaborators. Furthermore, the three SAP effectors were found to interact specifically with class I TCP while SAP11$_{AYWB}$ does not. Furthermore, SAP66 and SAP68 interacted with class II CIN-TCP members. Based on this result, I hypothesised that the SAP effectors could interact with the TCP domain, in a similar manner than SAP11 effector homologs. When I tested this in Y2H, I found that none of the SAP effectors could interact with the TCP domain, suggesting that the effectors target other regions within TCP.

The Y2H results of this chapter refine the global yeast-two hybrid screening performed by our collaborators. Indeed, our collaborators also found that SAP56, SAP66, SAP67 and SAP68 interacted with the different TCP (subclasses). In my case, the Y2H system established in this thesis relies on strong positive and negative controls. Indeed, the four SAP11 effector homologs have distinct interactions and at least one effector homolog interacts with one different TCP (sub)classes. As such, the positive interactions between class I TCPs and SAP effectors are unlikely to be the result of unspecificity. However, these results are a good start but it is undeniable that additional experiments need to be performed to characterise the interaction more thoroughly. For example, transient protoplast degradation assays could be performed to test if the SAP effectors destabilise the TCPs in the same way as SAP11. I did not have time to test it.

Based on the initial Y2H presented in **Figure 6.2.1**, I wished to determine the specific region that is targeted by these effectors. As they reside in the same genetic island as SAP11 and ca interact with TCP targets, an intuitive guess would be that they target the TCP domain. This would indicate that the ability to interact with the TCP domain, the conserved region of TCP proteins, is a global strategy of AY-WB phytoplasma strain. Under this premise, the AY-WB strain would use an array of effectors to specifically target the TCP domain, leading to dramatic changes in the plant development and defence. However, I did not find an interaction between any of the four SAP effectors and the TCP domain of class I TCP (sub)classes. Further experiments need to be completed to determine the TCP region that is targeted by SAP56, SAP66 and SAP68. The TCP domain is usually located at the N-terminal part of the TCP proteins of the three (sub)classes (Martín-Trillo and Cubas, 2010). Considering this, I started to generate TCP mutants that either lack the N-terminal part with or without the TCP domain, and the C-terminal part with or without the TCP domain and test the interaction with the three SAP effectors in Y2H (not shown).

Nonetheless, the preliminary functional analysis of these SAP effectors, coupled with the insect assays generated previously (Kingdom, 2012), shed some light on the possible mechanisms behind the role of these effectors. We know that SAP11 destabilises CIN-TCPs, leading to a down-regulation of JA and therefore the increase of *M. quadrilineatus*. Some SAP effectors such as SAP68 also lead to an increase of fertility of *M.quadrilineatus* and also target class II CIN-TCPs. Following the same logic, could the increase of survivability of *D. maidis* be attributed to the targeting of class I TCPs by SAP effectors? Based on the review of the different roles of class I TCPs in the plant development and plant immunity (see Chapter 5 discussion and General Introduction), I hypothesised that the downregulation of class I TCPs could lead to a decrease in the ETI or PTI of the plant host. This downregulation could help non-host insect vectors such as *D.maidis* to perform better on *A. thaliana* when the SAP effectors are stably expressed. To answer this hypothesis, the initial insect assays using the *35S::SAP56*, *35S::SAP66*, *35S::SAP67*, *35S::SAP68 A. thaliana* lines need to be repeated. Indeed, it is hard to correlate the patterns of interaction of the SAP effectors with either class II CIN or class I and their ability to induce either *M.*

quadrilineatus fertility or *D. maidis* survivability. Nonetheless, I would need class I TCP mutant lines, such as *tcp 8,14,15,* triple mutants to assess if *D. maidis* survives longer when exposed to this TCP mutant line. This could confirm the role of the class I TCPs in the plant defence on the one hand, and the role in the insect non-host resistance on the other hand.

Chapter 7: General Discussion

7.1 Summary and implication of the key findings

The SAP11 effector homolog from AY-WB phytoplasma strain binds and destabilises the class II TCP(sub)classes, leading to dramatic changes in both the plant development and the plant defence (Sugio et al., 2011a; Sugio and Hogenhout, 2012; Sugio et al., 2014; Pecher et al., 2019). Concerning the plant development, the destabilisation of the class II CIN TCPs leads to the characteristic crinkled leaves (Efroni et al., 2008; Schommer et al., 2008), while the destabilisation of class II CYC/TB1 leads to the witches' broom phenotype (Aguilar-Martinez et al., 2007; Pecher et al., 2019). Concerning the plant defence, the class II CIN-TCPs destabilisation leads to a downregulation of the *LOX2* gene, involved in the early steps of the synthesis of JA, an important plant hormone that confers resistance against necrotrophic pathogens and herbivores (Reymond and Farmer, 1998; Li et al., 2001; Howe and Jander, 2008; Browse, 2009).

First, using Y2H, I confirmed that $SAP11_{AYWB}$ and $SAP11_{MBSP}$ have distinct but overlapping patterns of interaction with both *A. thaliana* and *Z. mays* class II TCP (sub)classes. I extended the analysis using two SAP11 effector homologs from distantly related phytoplasma strains (Chung et al., 2013). Testing four SAP11 effector homologs from distantly related phytoplasma strains allowed me to identify trends in the function of the effector across phytoplasma strains. As such, I also found that $SAP11_{WBDL}$ and $SAP11_{SPLL}$ have distinct patterns of interaction with the TCP (sub)classes. $SAP11_{WBDL}$ and $SAP11_{SPLL}$ interacted with *A. thaliana* and *Z. mays* class II CYC/TB1-TCPs but not class II CIN-TCPs. Additionally, $SAP11_{AYWB}$, $SAP11_{WBDL}$ and $SAP11_{SPLL}$ interacted with CII-TCPs, a maize-specific (sub)class.

I then determined the specific region within TCP that was targeted by SAP11 and found that the four SAP11 effector homologs target the TCP domain of their cognate TCP protein partners. This indicates that the ability to target the conserved TCP domain could be extended among the phytoplasma strains, allowing them to target the TCP proteins across dicots and monocot plant (Cubas et al., 1999). Using chimaera constructs, I found that $SAP11_{AYWB}$ and $SAP11_{MBSP}$ bind the helix-loop-helix region of the TCP domain. This region is required for the homo/hetero-dimerisation

of the TCP proteins and configuration of the TCP domain beta-sheets allowing the TCP dimers to bind to the promoter of the target gene (Aggarwal et al., 2010; Pecher et al., 2019).

Based on these findings, I performed an enlarged phylogenetic and functional analysis of the SAP11 effector homologs from various phytoplasma isolates. I found that the SAP11 gene has a different evolutionary history from the core genome, represented here by the 16S rDNA sequences that are used to define phytoplasma species (IRPCM, 2004). The phylogenetic analyses and further functional analyses showed that *SAP11* gene has evolved based on its ability to interact with the different TCP (sub)classes of the phytoplasma isolates. Furthermore, some SAP11 effector homologs interact with the class I TCPs, which was not established before. I correlated the TCP interaction patterns of four distinct SAP11 homologs with their TCP destabilisation patterns using protoplast degradation assays, although more repetition needs to be done for class I TCP degradation. I further characterised the region within SAP11 that provides the binding specificity toward class II CIN-TCP and class I TCP. I found that the NLS domain is not required for the binding specificity.

Finally, I found that other AY-WB effectors from the same genetic island as $SAP11_{AYWB}$ also interacted with TCP (sub)classes, highlighting the importance of targeting this plant family. This indicates that some phytoplasma isolates rely on a set of effectors to extend their binding range toward different classes of plant targets.

Altogether, I have generated a body of evidence that phytoplasma SAP11 effectors have diverged to target members of different TCP subclasses. This work highlights how experimental analyses of an effector can complement phylogenetic analyses. Moreover, this work can serve as a platform for future studies about SAP11 effector homologs. Indeed, the model exposed in **Figure 5.3.3** could serve as a predictive model for future studies. The binding range of newly sequenced SAP11 effector homologs could be determined via implementation in the phylogenetic tree.

We still have much to learn about the genetic features that phytoplasma genome display, especially the PMUs. We still do not know about the origins and

dynamics of the PMUs (Chung et al., 2013; Ku et al., 2013; Music et al., 2019). For this, we would need to sequence more phytoplasma genomes and perform comparative genome analyses. The findings of this thesis would be improved if more knowledge was available on PMU structures of each SAP11 gene across the phytoplasma isolates. Concerning the interaction between SAP11 and the TCP targets, the results generated via the chimaeras approach uncovered the regions that are involved in the binding specificity for both actors. However, crystal structures of the SAP11-TCP complex would help uncover the key residues involved in the interaction. Finally, as discussed, future studies to determine the mechanisms behind the destabilisation of TCPs by SAP11 are required.

In this section, I will first summarise and discuss the evolution of the SAP effectors across the phytoplasma isolates. Secondly, based on this, I will discuss the strategies that different phytoplasma isolates use to promote the bacterial virulence and/or the fitness of the leafhopper insect vectors.

7.2 Technical Discussion

7.2.1 Technical Discussion about Y2H

As I previously discussed, the Y2H experiments of this thesis relied on strong positive controls. For example, AtTCP2 and AtTCP13 were selected as positive controls for a screening of the interaction between SAP effectors and class II CIN-TCPs. Nonetheless, it is important to underline that the rest of the CIN-TCPs could still interact with the SAP effectors. As previously discussed, the absence of interaction could be due to either (1) absence of interaction due to protein misfolding in yeast, (2) inserts in plasmids, preventing the expression of proteins in yeast or (3) no interaction (see Discussion Chapter 3).

Some results indicate the limitations of the Y2H system. For example, the Y2H result of **figure 6.2.1** suggests that SAP67 does not show interaction with any of the TCP proteins tested. One reason could be that the construct is not functional in yeast. In that case, additional experiments such as protoplast assays are required to confirm the interaction patterns observed via Y2H. Future experiments should be aimed at

repeating protoplast degradation assays between SAP67 and the TCP candidates. The results of **figure 3.2.1** show that $SAP11_{WBDL}$ interacts with AtTCP12 but not AtTCP18. This was later proved by the protoplast degradation assays (Chapter 5). This indicates sub-specificities within the TCP domain, suggesting that multiple residues within the TCP domain might be involved in the interaction with SAP11.

The differences in the Y2H results between the TCP proteins from *A. thaliana* and *Z. mays* (figure 3.2.6) suggest that the interaction patterns between SAP effectors and TCP proteins might differ between the hosts. This raises the question of how generalized the Y2H results are across the plant species? The TCP family is conserved across the plant species. The results presented here show that different SAP11 effector homologs interact with TCP proteins from monocotyledon and dicotyledon species. Furthermore, the SAP11 effector homologs specifically interact with the TCP domain, which defines the TCP family across the plant species. This indicates that the results can be generalized to a certain extent across the plant species (More information on this on General Discussion 7.4).

7.3 The evolution of phytoplasma effectors

The evolution of *SAP11* within the phytoplasma genomes is fascinating and could be explained in part by the PMUs that feature phytoplasma genomes genes (Bai et al., 2006; Dickinson, 2010; Toruño et al., 2010). As mentioned, I cannot confirm that the *SAP11* genes of the 21 phytoplasma isolates are all expressed from PMU structures. However, at least '*Ca*. P. asteris' AY-WB strain, MBSP strain and Miaz9 isolate are found to express *SAP11* in PMUs (Bai et al., 2006; Sugio and Hogenhout, 2012; Orlovskis et al., 2017). Some *SAP11* genes of phytoplasma isolates could also be present in PMU-like clusters, which are degenerate versions of PMUs (Bai et al., 2006; Hogenhout and Loria, 2008).

In bacteria, the functionalisation of effectors from the same family relies on either gene duplication or horizontal gene transfer (Lerat et al., 2005; Treangen and Rocha, 2011; Remigi et al., 2011). The horizontal gene transfer allows the pathogen to acquire novel virulence genes, which for phytoplasma could be due to PMUs (Bai

et al., 2006; Dickinson, 2010; Chung et al., 2013; Ku et al., 2013). There are three outcomes in a gene duplication event that shapes the gene function: neo-functionalization, sub-functionalization or gene conservation (redundancy or dosage) (Ohno, 1970; Hahn, 2009). For example, the type III *GALA* effector family from *Ralstonia solanacearum* underwent functional diversification via the neofunctionalization of a member of the family (Remigi et al., 2011). Phytoplasma effectors have evolved and could follow these outcomes.

As discussed previously, the rate of evolution of PMU-related genes such as *SAP11* is likely to be faster than the rate of genes that reside outside PMU structures. This compartmentalisation resembles the "two-speed genome" model of filamentous pathogens (Croll and McDonald, 2012; Dong et al., 2015). Both *Phytophthora infestans* and phytoplasmas have a significant portion of repeat-rich regions within their genome and display their effector genes in mobile elements (Raffaele et al., 2010; Dickinson, 2010; Chung et al., 2013; Ku et al., 2013; Dong et al., 2015). This strategy allows several evolutionary advantages to phytoplasma (Croll and McDonald, 2012). For example, compartmentalised PMUs allow co-existing phytoplasma isolates to acquire new effector genes, thus expanding the function of the effector. Secondly, the higher rate of evolution within PMU structures could allow for a higher rate of substitution. In that case, the newly acquired effector gene could be beneficial for phytoplasma, allowing to expand its host range. On the other hand, a gene that does not benefit to phytoplasma in certain plant hosts might be negatively selected and its frequency drop in the population (Croll and McDonald, 2012).

As described in the review of Dong et al., 2015, several filamentous plant pathogens have evolved by jumping from one host to the next (Roy, 2001; Dong et al., 2015). The host jump is expected to impact the functional evolution of the effectors as they function inside the plant cell (Dong et al., 2015). For example, in *P. infestans*, some effector genes have adapted to the new plant host and either improved or expanded their activity (Raffaele et al., 2010; Dong et al., 2014, 2015).

Consistent with the two-speed genome, genome compartments underwent accelerated gene evolution after a host-jump (Raffaele et al., 2010). In the case of phytoplasma, the horizontal transfer of SAP effectors via PMUs between two phytoplasmas could facilitate host jump. One generalist strain could acquire a specialised effector from a second generalist strain that could expand the host range of the first strain.

7.3.1 The evolution of SAP11 effector homologs

A SAP11 effector homolog that can target another (sub)class in addition to class II CYC/TB1-TCPs is likely to be beneficial for the pathogen and therefore be positively selected over a *SAP11* gene that expresses an effector that destabilises class II CYC/TB1-TCPs only.

$SAP11_{p45.1}$ and $SAP11_{p45.2}$ are an example of gene duplication of an effector gene that results in redundancy. Indeed, both effectors are almost identical and are expressed from the same phytoplasma isolate. $SAP11_{PLYDY.1}$ and $SAP11_{PLYDY.2}$ could be duplicated gene that resulted in the neofunctionalization of the duplicated gene, allowing it to target a new (sub)class of TCP. However, this is unlikely considering how similar $SAP11_{PLYDY.2}$ is to $SAP11_{OY\text{-}M}$, which belong to the same SAP11 clade. This example underlines the possible horizontal transfer of *SAP11* from two generalist strains (Oshima et al., 2001; Chen et al., 2011). OY-M phytoplasma strain and PLYDY phytoplasma are likely to share a plant host.

Some phytoplasma isolates rely on other strategies, depending on their host range, the plant host and their insect vectors. For example, MBSP is a maize specialist, and therefore could be less likely to have been in contact with other phytoplasma strains and acquire a SAP11 effector homolog able to target other (sub)classes. $SAP11_{MBSP}$ has a narrower binding range toward the plant TCP (Pecher et al., 2019). However, sequencing of the genomes of MBSP isolates in various locations of Brazil showed that $SAP11_{MBSP}$ remains conserved (Orlovskis et al., 2017). We recently showed that the stable production of $SAP11_{MBSP}$ does not downregulate the JA pathway and does not significantly increase the fertility of *D. maidis* or *M. quadrilineatus* in maize (Pecher et al., 2019). This work suggests that the gene

underwent purifying selection via other means than the increase of the fertility of *D. maidis* (Pecher et al., 2019). So how does $SAP11_{MBSP}$ promote the fitness of MBSP phytoplasma strain or *D. maidis*? A few hypotheses emerge from this. First, *D.* maidis co-evolved with MBSP since the domestication of maize (Nault, 1980). Therefore, the leafhopper might not benefit from downregulation of JA in maize. Secondly, targeting class II maize CYC/TB1-TCPs leads to additional effects compared to *A. thaliana* CYC/TB1-TCPs, such as inhibition of female flower production (Pecher et al., 2019). Alternatively, the maize TCPs might have additional roles than *A.thaliana* TCPs, suggesting that downregulating class II CIN-TCP and CII-TCPs is not viable for the fitness of MBSP (Pecher et al., 2019).

SAP11 effector homolog from *'Ca.* P.mali' strain AT provides yet another insight into the functional diversification of the *SAP11* gene. $SAP11_{ATP}$ is also expressed from a specialist phytoplasma strain but differs from $SAP11_{MBSP}$ by its ability to interact with all the TCP (sub)classes. The strategy could be to have a multi-purpose SAP11 effector that has an extensive spectrum of interaction with the TCP(sub)classes (Janik et al., 2017). As discussed, *'Ca.* P.mali' has a few repeated regions, only one PMU (Ku et al., 2013) and no plasmid in its genome (Kube et al., 2008).

Finally, the evolution of the NLS domain of SAP11 effector homologs (see Chapter 4) suggests that the NLS domain got re-shuffled in between the N-terminal and the C-terminal part. The SAP11 effector homologs tend to cluster based on the localisation of their NLS domain. Since the NLS domain is essential for the function of the effector, it would be interesting to know if the ancestral SAP11 effector has its NLS domain at the N-terminal, C-terminal or both. Can the position of the NLS explain the difference in the function of the effector? We could test if the position of the NLS domain correlates different functionality. To test this, we could produce *in planta* SAP11 chimaeras where the N-terminal NLS has been replaced by the C-terminal NLS of a homolog and test for difference of phenotype. Most of the SAP11 effector homologs that display their NLS domain at the C terminal part can interact with Class I TCPs. Therefore, we could produce Class I interactors such as $SAP11_{SPLL}$ *in planta* and test for difference of phenotype. However, a recent study shows that the stable

expression of $SAP11_{PnWB}$, which is predicted to display its NLS domain at the C-termini, in *Arabidopsis* does not lead to characteristic phenotypes of the destabilisation of Class I TCPs (Chang et al., 2018). As discussed, this is probably due to the redundancy of the Class I TCPs. Therefore, such studies would be challenging. Nonetheless, the two clades of proteins that differ mainly in the location of their essential NLS may provide an interesting lead to study functional diversification of effector proteins.

7.3.2 The evolution of SAP56/66/68 in AY-WB phytoplasma

Chapter 6 showed that in AY-WB phytoplasma, multiple SAP effectors from the genetic island of SAP11 interact with TCP (sub)classes. This could be an example of another feature of phytoplasma genomes. As the TCP plant family is at the cross-road of the plant interactome, some phytoplasma effectors might have been selected to target these proteins (Dreze et al., 2011; Mukhtar et al., 2011; Weßling et al., 2014). Moreover, the sequences of the five SAP effectors are dissimilar (see **Appendix V – figure 1**), and preliminary data suggest that they do not interact with the same region within TCPs, suggesting that they have independently evolved to interact with different regions of TCP (sub)classes.

7.3.3 The evolution of SAP05 effector homologs

The SAP05 effector is another example of phytoplasma effector that has evolved based on its interaction pattern with the plant targets. Indeed, Dr Weijie Huang from our lab has found that SAP05 interacts with two zinc-finger plant transcription factors families: GATA and SQUAMOSA promoter-binding protein-like (SPL) (Huang, personal communication). Briefly, he found that $SAP05_{AYWB}$ degrades SPL and GATA transcription factors, leading to a disturbance of plant phase transitions, reproduction and branching. He extended the analysis using SAP05 effector homologs from multiple phytoplasma isolates, including phytoplasmas that I studied in my PhD. He found that the SAP05 effector homologs differentially interacted with the plant SPL and GATA transcription factor families. Some SAP05

effector homologs target only SPLs or only GATAs, in a similar fashion than SAP11 effector homologs with TCP (sub)classes.

The multiple sequence alignment indicates that SAP05 are more homologous than SAP11. Furthermore, some phytoplasma isolates display two copies of SAP05 that have distinct patterns of interaction with the two plant transcription factor families. This suggests that SAP05 effector homologs underwent functional diversification, with potential gene duplication, leading to neo-functionalization or sub-functionalization of the new copy. As an example, the WBDL phytoplasma strain expresses two copies of *SAP05*. One of them interacts with the GATA family while the other one targets the SPL family. This suggests a neofunctionalization of the new copy of *SAP05* in WBDL phytoplasma strain. In the case of VWBP phytoplasma, the two *SAP05* copies interact with both GATA and SPL, suggesting redundancy in the function of the gene.

Alternatively, horizontal transfer of *SAP05* genes could also occur in specific phytoplasma isolates. For example, AY-WB phytoplasma *SAP05* gene resides in a PMU-like structure (Bai et al., 2006).

7.4 The role of the phytoplasma effectors in the promotion of the bacterial virulence and insect vector fitness

In order to promote their fitness, plant pathogens must either avoid the plant host defence mechanisms or suppress these defence responses. In addition, plant pathogens that rely on insect vectors may increase the fitness of their insect vector(s), thus favouring pathogen spread. Diverse strategies are used, from modulating the plant immunity (Boller and He, 2012) to altering the plant development (Mescher, 2012). Phytoplasma relies on its arsenal of SAP effectors to put in place their strategies. Based on the wide variety of symptoms caused by phytoplasmas, these strategies are likely to be extremely varied. Symptoms of most phytoplasmas indicate a preference to a more biotrophic lifestyle (witches broom symptoms). However, some phytoplasmas cause symptoms more associated to a necrotrophic lifestyle (Palm wilt, lethal yellowing). To elucidate how these different

strategies are implemented by the pathogen we need clear case studies for several of phytoplasma isolates. In my thesis, I focussed on those phytoplasma isolates with a more biotrophic lifestyle. Therefore, caution should taken when extrapolating the results in this thesis to those type of infections that appear to favour a radically different lifestyle, most notably a necrotrophic symptomology. I will summarise here the role of SAP effectors to increase the virulence of phytoplasma with different symptomology and the insect vector and see how the different strategies fulfil these goals. As this thesis underlines, the phytoplasma isolates use distinct means to modulate the plant development and defence. I will, therefore, elaborate case studies of different phytoplasma isolates for each strategy.

7.4.1 Phytoplasma targets transcription factor families to target additional layers of regulation

This thesis shows that different SAP effectors from multiple phytoplasmas target the three (sub)classes of the plant TCP family. This underlines how important this family of transcription factors is for the correct development and defence of the plants (Martín-Trillo and Cubas, 2010; Lopez et al., 2015; Nicolas and Cubas, 2016). Indeed, TCPs are highly connected in the plant interactome and are the target of a wide range of pathogens (Dreze et al., 2011; Mukhtar et al., 2011; Weßling et al., 2014). As discussed in this thesis, TCPs are involved in multiple key processes, from hormonal biosynthesis and signalling, transport, immunity, regulation of circadian clock, floral development, shoot branching (Lopez et al., 2015; Nicolas and Cubas, 2015, 2016; Bemer et al., 2017; Dhaka et al., 2017).

TCPs interact with multiple transcription factor families; thereby, phytoplasma effectors might indirectly affect additional pathways upon targeting TCPs (Nicolas and Cubas, 2016; Dhaka et al., 2017). Multiple studies have shown the cross-family interaction between TCP and other essential transcription factors families. For example, class II CIN TCP4, TCP10 and TCP24 interact with ASYMETRIC LEAVES2 (AS2) in order to repress *KNOX* genes and regulate the shoot apical meristem (Li et al., 2012). Class I TCP14 interacts with GA-regulated DELLA proteins and is involved in GA-mediated plant height regulation (Davière et al., 2014). TCP8

also interacts with WRKY28, which is implicated in the plant pathogen responses (Wang et al., 2015b; Bemer et al., 2017).

Furthermore, SAP effectors target other plant transcription factor families. SAP54 targets the MADS-box transcription factor family, which are involved in flower development (MacLean et al., 2014). TCP proteins are known to regulate MADS-box transcription factors, with OsMADS57 positively regulated by OsTB1, leading to rice tillering (Guo et al., 2013). Maize TB1-TCP indirectly regulates MADS-box transcription factors via TGA1 (Studer et al., 2017). As discussed, our lab found that SAP05 targets two plant transcription factor families SPL and GATA (Huang, personal communication).

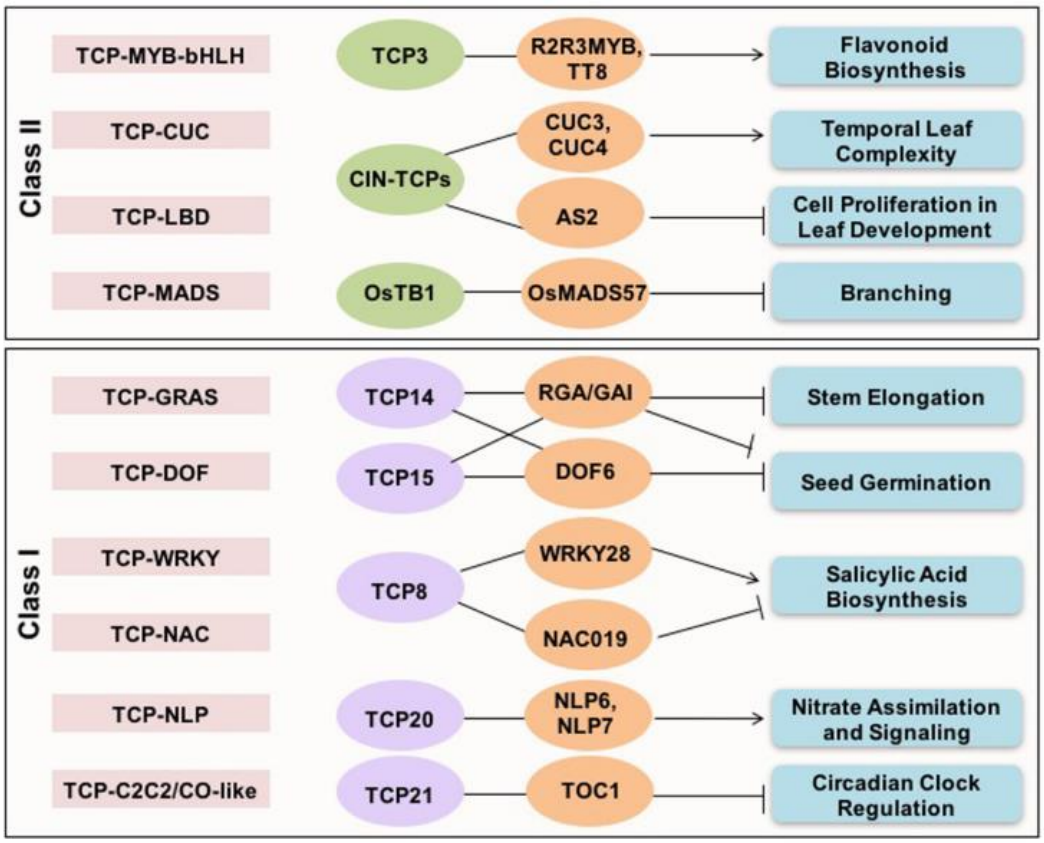


Figure 7.1: TCP (sub)classes form interaction with multiple transcription factor families. The TCP interactions with key transcription factor families lead to distinct regulations of the plant development and defence. Figure pasted from Dhaka et al., 2017. *Frontiers in Plant Science*.

7.4.2 Phytoplasma alters the plant development via a set of SAP effectors

7.4.2.1 Phytoplasma promotes the plant stem proliferation to increase its niche

Phytoplasma isolates from diverse phytoplasma groups express SAP11. Based on this thesis, all SAP11 effector homologs interact with class II CYC/TB1-TCPs in yeast. Furthermore, different SAP11 effector homologs destabilise these TCPs, suggesting that the ability to destabilise class II CYC/TB1-TCPs is conserved across the phytoplasma groups (see Chapter 5) (Sugio et al., 2011a, 2014; Chang et al., 2018; Pecher et al., 2019). *A. thaliana* BRC1 and BRC2 and *Z. mays* TB1 are suppressors of axillary bud growth (Finlayson, 2007; Dong et al., 2017; Gonzalez-Grandio et al., 2017). As discussed, the destabilisation of class II CYC/TB1-TCP is consistent with the phenotype showing an increase in the stem branching (see Chapter 3) (Sugio et al., 2011a; Chang et al., 2018; Pecher et al., 2019).

These findings strongly suggest that the increase in stem branching is a common strategy across phytoplasmas. Lateral branching is a critical process in determining the shoot architecture of plants (Uberti Manassero et al., 2013). One goal behind this strategy would be to increase the niche of phytoplasma within the infected plant. This would increase the habitat of phytoplasma, with more stems, therefore more sieve elements.

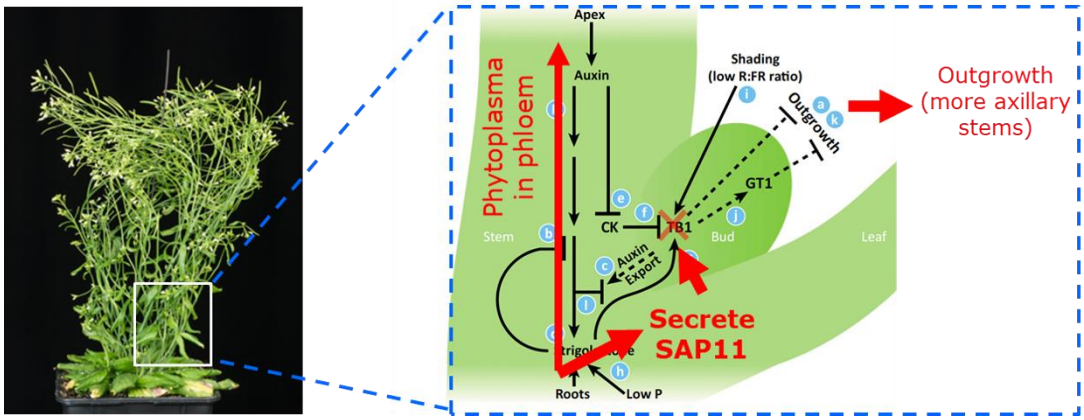


Figure 7.2: Phytoplasma promotes plant stem proliferation to increase its niche. Phytoplasma express SAP11 effector to bind and destabilise class II CYC/TB1-TCPs, leading to an increase in the stem number. On the left, the stable expression of $SAP11_{MBSP}$ leads to an increase in stem branching (see Chapter 3). On the right, a model of the regulation of bud outgrowth. When a plant is not infected by phytoplasma, the bud outgrowth is inhibited by

the activity of TB1 (class II CYC/TB1-TCP) and GT1. When the plant is infected by phytoplasma, phytoplasma secrete SAP11, which will translocate to the axillary buds and destabilise TB1-TCPs. The destabilisation of TB1-TCPs results in an outgrowth of the buds, leading to an increase in the number of axillary stems. Model modified from Kebrom et al., 2013. *Trends Plant Sci.*

7.4.2.2 The zombie plant strategy to extend the lifespan of phytoplasmas in plant host

Many phytoplasma isolates have a biotrophic lifestyle that requires a living host to survive. Therefore, those phytoplasma needs to delay cell death by keeping the plant alive longer (Sugio et al., 2011b). Increasing the lifespan of the plant host could also help for the phytoplasma acquisition by insect vectors. Many of the shytoplasma isolates studied to date rely on SAP11, SAP54, TENGU effectors to achieve this goal.

The AY-WB phytoplasma strain case study

In the case of AY-WB phytoplasma strain, SAP54 induces the phyllody symptoms, where flowers turn into leaf-like tissue (MacLean et al., 2011, 2014). The phyllody symptoms are due to the degradation of MADS-box transcription factors, which are involved in floral development (MacLean et al., 2011, 2014) (see Chapter 1 General Introduction). The alteration of the structure of the flower often lead to sterility. Therefore, the plant becomes a host that only propagates for the benefit of phytoplasma, hence the term “zombie” plant (MacLean et al., 2014).

AY-WB phytoplasma strain also relies on $SAP11_{AYWB}$, which binds and destabilises class II CIN-TCPs. As discussed in Chapter 3 and Chapter 5, class II CIN-TCPs are involved in the promotion of leaf maturation (Efroni et al., 2008; Danisman et al., 2012). Therefore, one of the effects of the destabilisation of class II CIN-TCPs by $SAP11_{AYWB}$ would be to extend the vegetative phase by delaying senescence.

The MBSP phytoplasma strain case study

Our lab sequenced the genome of diverse MBSP isolates in Brazil (Orlovskis et al., 2017). They found that the MBSP genome has 36 candidate effector genes, but *SAP54* effector gene is not present (Orlovskis et al., 2017). In order to extend the lifespan of the plant host, MBSP could rely on the SAP11 effector. Indeed, the

destabilisation of maize class II CYC/TB1-TCPs leads to additional effects compared to *A. thaliana* CYC/TB1-TCPs, such as inhibition of female flower production (Pecher et al., 2019). This is because TB1 TCP regulates the flowering development in maize via regulation of MADS-box transcription factors (Studer et al., 2017). Alternatively, the maize TCPs might have other roles than *A.thaliana* TCPs, suggesting that downregulating class II CIN-TCP and CII-TCPs is not viable for the fitness of MBSP (Pecher et al., 2019).

7.4.3 Phytoplasma increases the fitness of leafhopper insect vectors via a set of SAP effectors

7.4.3.1 Phytoplasma attracts and increase the fertility of leafhoppers

Phytoplasma requires insect vectors for spreading before the plant host dies (Weintraub and Beanland, 2006). For this, the obligate parasite uses its set of effectors for at least two known goals: (1) increase the insect fertility and (2) increase the leafhopper attraction to infected plants. As previously discussed, the first goal is effective via $SAP11_{AYWB}$ that induces an increase in the fertility of *M. quadrilineatus* in *A. thaliana* (Sugio et al., 2011a) (see General Introduction).

The second goal is achieved with SAP54 that induces the attraction of leafhoppers (Orlovskis and Hogenhout, 2016). Indeed, our lab found that the stable production of SAP54 in *A. thaliana* attracts *M. quadrilineatus* compared to wild-type (Orlovskis and Hogenhout, 2016). Furthermore, we have evidence that the witches' broom phenotype of infected lime trees, coupled with yellowing of leaves in symptomatic parts of the plant, increase the attraction of leafhopper insect vectors (Hogenhout, personal communication). This could indicate that the destabilisation of CYC/TB1-TCPs could be involved in part in the attraction of leafhoppers.

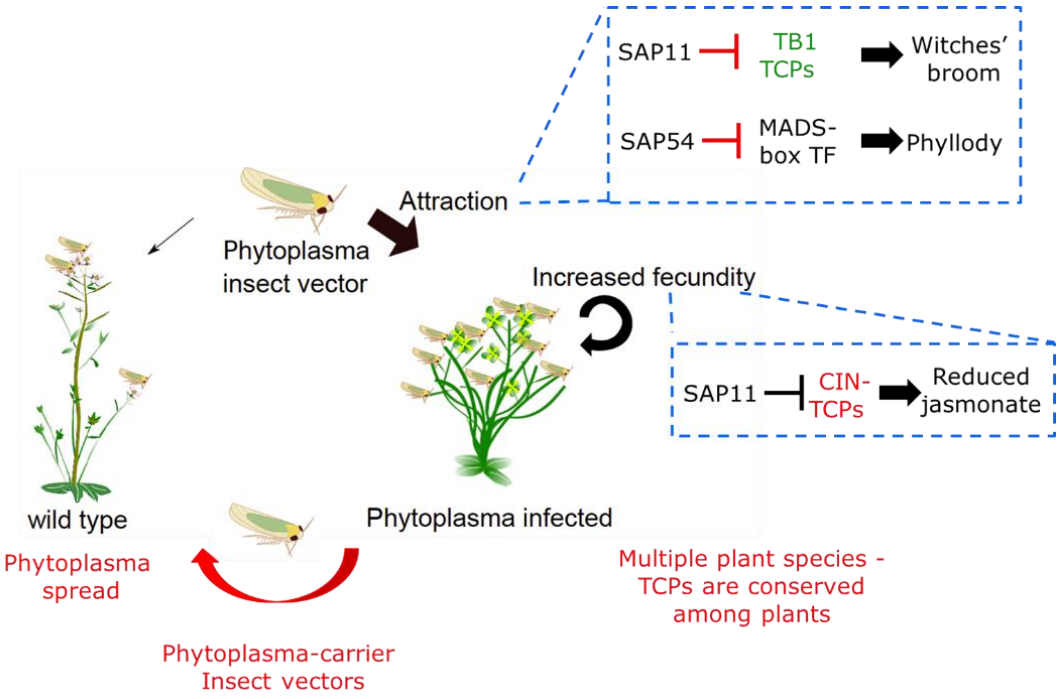


Figure 7.3: Phytoplasma increases the fitness of leafhopper insect vectors via a set of SAP effectors. AY-WB phytoplasma strain expresses SAP11 and SAP54 effectors to increase the fitness of *M. quadrilineatus*. $SAP11_{AYWB}$ promotes the fecundity of *M. quadrilineatus*, while SAP54 increases the attraction of *M. quadrilineatus*. Both effects lead to an increase in the spreading of phytoplasma. Adapted from Tomkins et al., 2018. *Curr. Opin. Plant Biol.*

7.4.3.2 Phytoplasma could increase the non-host insect fitness

Phytoplasma could also increase the fitness of non-host insect vectors. AYWB-infected *A. thaliana* plants increase the survivability and reproduction of *D. maidis* (Purcell, 1988; Sugio et al., 2011b). *D. maidis* lays eggs on AY-WB infected Arabidopsis plants and nymphs emerge two weeks later, suggesting that the plant is a feeding host (Sugio et al., 2011b). The stable production of $SAP11_{AYWB}$ does not promote the survivability and reproduction of *D. maidis* in *A. thaliana.* However, as discussed in Chapter 6, some AY-WB effectors from the SAP11 genetic island increase the survivability of non-host *D. maidis* in *A. thaliana*. This result could be due to the targeting of class I TCPs (see Chapter 6).

Class I TCPs are involved in the plant immunity, in both PTI and ETI (Kim et al., 2014; Spears et al., 2019). Targeting this (sub)class could lead to a reduction in the plant defence against insects. Therefore, non-host insect vectors could perform

better on class I TCP defective plants. This would increase the chances of non-host leafhoppers to acquire phytoplasma when feeding on the infected plant. Although *D. maidis* does not vector AY-WB phytoplasma (Sugio et al., 2011b), other non-host leafhoppers could acquire a phytoplasma isolate and transmit it to new plant hosts.

Targeting class I TCPs could allow phytoplasma to indirectly expand its range of insect vectors. Consequently, phytoplasma could expand its host range through insect vectors. Most of the generalist phytoplasmas studied in this thesis seem to destabilise class I TCPs, either via SAP11 in the case of SPLL, OY-M, PLYDY, the solani isolates or via other SAP effectors in the case of AY-WB phytoplasma (chapter 6). One of the reasons behind their ability to interact with multiple hosts could be through destabilisation of class I TCPs. The destabilisation of this TCP class could lead to an increased number of potential new non-host insect vectors. The targeting of class I could increase the chances of phytoplasma to be shared between plant hosts, and also increase the chances of horizontal transfer of effector genes between the strains.

Appendices

Appendix I – Supplemental Tables for Chapter 2

Appendix I table 1: Sequence IDs of TCPs from *Zea mays* (Zm), *Arabidopsis thaliana* (At).

This table can be seen in our latest paper (Pecher et al., 2019).

Zea mays (Zm) TCPs			
Class/group		Name	Sequence ID
class II	CII	ZmTCP01	GRMZM2G166687
		ZmTCP05	GRMZM2G110242
		ZmTCP13	GRMZM2G060319
	CYC/ TB1	ZmTCP02	AC233950.1_FG002
		ZmTCP18	AC190734.2_FG003

Arabidopsis thaliana (At) TCPs			
Class/group		Name	Sequence ID
ClassI		AtTCP6	GRMZM2G078077
		AtTCP7	GRMZM2G003944
		AtTCP8	GRMZM2G445944
		AtTCP9	GRMZM2G142751
		AtTCP11	GRMZM2G113888
		AtTCP14	GRMZM2G096610
		AtTCP15	GRMZM2G465091
		AtTCP16	GRMZM2G092214
		AtTCP19	GRMZM2G093895
		AtTCP20	GRMZM2G178603
		AtTCP21	AC199782.5_FG003
		AtTCP22	GRMZM2G034638
		AtTCP23	GRMZM2G107031
ClassII	CIN	AtTCP2	At4g18390
		AtTCP3	At1g53230
		AtTCP4	At3g15030
		AtTCP5	At5g60970
		AtTCP13	At3g02150
		AtTCP10	At2g31070
		AtTCP17	At5g08070
		AtTCP24	At1g30210
	CYC/ TB1	AtTCP1	At1g67260
		AtTCP12	At1g68800
		AtTCP18	At3g18550

Appendix I table 2: Oligonucleotide sequences (5´ > 3´) for cloning. the green sequences indicate the start codon while the red sequence indicates the stop codon. The underlined sequences indicate the gene-specific region of the primer. The sequences before the start codon (forward primers) or the stop codon (reverse primer) are the attB1 or attB2 sequences.

Oligonucleotide	Sequence (5 -> 3)
attB1	GGGACAAGTTTGTACAAAAAAGCAGGCTTC
attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTC
TCP domain of AtTCP6 forward	GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAAGATCGTCACCTTAAAGTTG
TCP domain of AtTCP6 reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAAGTAGCAGAGAGTATCGATGGC
TCP domain of AtTCP8 forward	GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAAGACCGTCACACGAAAGTCG
TCP domain of AtTCP8 reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATGTAGCAGCAACAATAGCTGG
TCP domain of AtTCP9 forward	GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAAGACCGTCACACGAAGGTTG
TCP domain of AtTCP9 reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTACGTGGCGGCTATAATCGC
TCP domain of AtTCP14 forward	GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAAGACCGACACACGAAAGTAG
TCP domain of AtTCP14 reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGGTGGCGGCGATTACAGAT
TCP domain of AtTCP15 forward	GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAAGACCGTCACACGAAAGTCG
TCP domain of AtTCP15 reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATGTAGCGGCTATAACCGCTG
TCP hlh motif of AtTCP2 forward	GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCTCTTCAATTCTATGATCTTC
TCP hlh motif of AtTCP2 reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATTCAGCAGCTTTGATTAACC
TCP hlh motif of AtTCP18 forward	GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCCAAAGAGTTGTTTGGCTTAC
TCP hlh motif of AtTCP18 reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATTTTGCTTGTGTAAGCAACC

Appendix I table 3: Synthesised CDS (underlined) flanked by gateway compatible attL1 and attL2 sites. Nucleotide sequences for gene syntheses of SAP11 chimaeras, TCP chimaeras, TCP domains and SAP11 effector homologs for expression in yeast. The figure number of the thesis is also indicated.

Name of the construct:	Thesis Figure #
SAP11-Chimaera 1: chimeric *SAP11$_{MBSP}$* with *SAP11$_{AYWB}$* TCP binding domain used in Y2H analysis. Same sequence as (2) of appendix.	**5.2.5**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCTCACCTAAAAAAGAAGATCGCGGAAAAAATGTTGCAACTTCAAAAGAAAAAGAAACACTAACTAAAGAAGAAGTGAAACGTTTTTTTGAATACCATAAAACATTTGAAACATATTCTGATGAAGACAAAATTAAAATTATTGAAAAAATTACCGACCCAGAAGTAATGGAAATATTAAAACAAAAAGCCGAAGAGGAAACGAAAAATTTAAAGGAAGAAAGTTCTTCTAGCAAAAAACCTGATAATTCAAAAAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11-Chimaera 2: chimeric *SAP11$_{AYWB}$* with *SAP11$_{MBSP}$* NLS domain used in Y2H analysis. Same sequence as (3) of appendix.	**5.2.5**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCTCACCTAAAAAAGAATCTAGTGATAAAAATGTTGCAACTTCAAAAGAAAAAGAAACACTAACTAAAGAAGATATAAAAAGATTTTATACAATACATAAAGAATTTAAAGAATATTCAATTGAAAAAAATAATGAAATTATAAAAATTTTAGAAAACCCTGAATTAATGGAAATATTAAAACAAAAAGCCGAAGAGGAAACGAAAAATTTAAAAGAAGAAGGTTCTTCTTCAAAACAACCTGATGATTCTAAAAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11-Chimaera 3: chimeric *SAP11$_{AYWB}$* with *SAP11$_{MBSP}$* NLS domain and *SAP11$_{MBSP}$* intra region 3 used in Y2H analysis. Same sequence as (7) of appendix.	**5.2.5**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCTCACCTAAAAAAGAATCTAGTGATAAAAATGTTGCAACTTCAAAAGAAAAAGAAACACTAACTAAAGAAGATATAAAAAGATTTTATACAATACATAAAGAATTTAAAGAATATTCAATTGAAAAAAATAATGAAATTATAAAAATTTTAGAAAACCCTGAATTAATGGAAATATTAAAACAAAAAGCCGAAGAGGAAACGAAAAATTTAAAGGAAGAAAGTTCTTCTAGCAAAAAACCTGATAATTCAAAAAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	

SAP11-Chimaera 4: chimeric *SAP11$_{AYWB}$* with *SAP11$_{MBSP}$* NLS domain and *SAP11$_{MBSP}$* intra region 2 used in Y2H analysis. Same sequence as (8) of appendix.	**5.2.5**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCTCAC CTAAAAAAGAATCTAGTGATAAAAATGTTGCAACTTCAAAAGAAAAAGAAACACTAAC TAAAGAAGAAGTGAAACGTTTTTTTGAATACCATAAAACATTTGAAACATATTCTGATG AAGACAAAATTAAAATTATTGAAAAAATTACCGACCCAGAAGTAATGGAAATATTAAA ACAAAAAGCCGAAGAGGAAACGAAAAATTTAAAAGAAGAAGGTTCTTCTTCAAAACAA CCTGATGATTCTAAAAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCAT TATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATA AAATCATTATTTGCCATC	
SAP11-Chimaera 5: chimeric *SAP11$_{AYWB}$* with *SAP11$_{MBSP}$* intra region 2 used in Y2H analysis. Same sequence as (9) of appendix.	**5.2.5**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCTCAC CTAAAAAAGAATCTAGTGATAAAAAAAGAGATATTCCGAAAATTAATAAATCAGAAGA AAAAAACAAAAAACAAAAAGAAGAAGTGAAACGTTTTTTTGAATACCATAAAACATTT GAAACATATTCTGATGAAGACAAAATTAAAATTATTGAAAAAATTACCGACCCAGAAGT AATGGAAATATTAAAACAAAAAGCCGAAGAGGAAACGAAAAATTTAAAAGAAGAAGG TTCTTCTTCAAAACAACCTGATGATTCTAAAAAATAAGGCAAGCTTGACCCAGCTTTCTT GTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCA CTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11-Chimaera 6: chimeric *SAP11$_{MBSP}$* with *SAP11$_{AYWB}$* TCP binding domain and *SAP11$_{AYWB}$* intra region 2 used in Y2H analysis. Same sequence as (11) of appendix.	**5.2.5**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCTCAC CTAAAAAAGAAGATCGCGGAAAAAATGTTGCAACTTCAAAAGAAAAAGAAACACTAAC TAAAGAAGATATAAAAAGATTTTATACAATACATAAAGAATTTAAAGAATATTCAATTG AAAAAAATAATGAAATTATAAAAATTTTAGAAAACCCTGAATTAATGGAAATATTAAAA CAAAAAGCCGAAGAGGAAACGAAAAATTTAAAGGAAGAAAGTTCTTCTAGCAAAAAA CCTGATAATTCAAAAAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCAT TATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATA AAATCATTATTTGCCATC	
SAP11-Chimaera A: chimeric *SAP11$_{MBSP}$* with *SAP11$_{SPLL}$* TCP intra-region 2 used in Y2H analysis.	**5.2.6**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGT CCCCGAAAAAAGAGGATAGGGGCAAAAACGTAGCCACAAGCAAGGAGAAAGAGACC TTGACGAAGAAGGACATCAGTCAGTATTATGAACTATATAATACTTTAGAAAACTATTC CGAGGAAGACCGAAATAAGATTATACAGATGCTGAGTGACTCTCAGACGAGTAAACTA TTGGATGAGTATAATGAAAAGAAAAGAAAGTCATCCAAGGAGGAGAGCAGCTCTTCA	

AAGAAACCAGACAACTCCAAGAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAG TTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGT CAAAATAAAATCATTATTTGCCATC	
SAP11-Chimaera B: chimeric *SAP11$_{MBSP}$* with *SAP11$_{SPLL}$* TCP intra-region 2 and *SAP11$_{SPLL}$* TCP binding domain used in Y2H analysis.	**5.2.6**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGA GCCCGAAGAAAGAGGATAGAGGGAAGAACGTCGCCACGTCTAAGGAAAAAGAAACG CTAACTAAGAAAGACATTAGCCAATACTATGAATTATACAATACTTTAGAAAATTATTCT GAAGAGGACCGTAATAAGATAATACAAATGCTAAGCGATAGTCAGACACTAAAGATCC TCCAGGAGGAAGCCCTGAAGAGCAAGGAGGAAAGTTCATCATCAAAAAAGCCGGATA ATTCCAAGAAGTAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAA AATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCAT TATTTGCCATC	
Sequence of TCP domain *AtTCP2* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCAAAG ATAGACATTCTAAAGTTTTGACTTCCAAGGGTCCAAGAGATAGAAGAGTTAGATTGTCT GTTTCTACCGCCTTGCAATTTTACGACTTGCAAGATAGATTGGGTTACGACCAACCATCT AAAGCTGTTGAATGGTTGATTAAGGCTGCCGAAGATTCCATTTCTGAATTGCCATAAGG CAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAAT TTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
Sequence of TCP domain of *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCACTG ATAGACATTCTAAAATCAAAACTGCCAAGGGTACTAGAGATAGAAGAATGAGATTGTC CTTGGATGTCGCCAAAGAATTATTCGGTTTACAAGACATGTTGGGTTTCGATAAGGCTT CTAAAACTGTCGAATGGTTGTTGACTCAAGCCAAGCCAGAAATTATCAAGATTGCCTGA GGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATC AATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
TCP-Chimaera 1: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCACTG ATAGACATTCTAAAATCAAAACTGCCAAGGGTACTAGAGATAGAAGAATGAGATTGTC CGTTTCTACCGCCAAAGAATTATTCGGTTTACAAGACATGTTGGGTTTCGATAAGGCTT CTAAAACTGTCGAATGGTTGTTGACTCAAGCCAAGGATTCCATTTCTGAATTGCCATAA GGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATC AATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	

TCP-Chimaera 2: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCAAAG ATAGACATTCTAAAGTTTTGACTTCCAAGGGTCCAAGAGATAGAAGAGTTAGATTGTCT GTTTCTACCGCCAAAGAATTATTCGGTTTACAAGACATGTTGGGTTTCGATAAGGCTTC TAAAACTGTCGAATGGTTGTTGACTCAAGCCAAGGATTCCATTTCTGAATTGCCATAAG GCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCA ATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCAT	
TCP-Chimaera 3: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCACTG ATAGACATTCTAAAATCAAAACTGCCAAGGGTACTAGAGATAGAAGAATGAGATTGTC TGTTTCTACCGCCTTGCAATTTTACGACTTGCAAGATAGATTGGGTTACGACCAACCATC TAAAGCTGTTGAATGGTTGATTAAGGCTGCCGAAGATTCCATTTCTGAATTGCCATAAG GCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCA ATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
TCP-Chimaera 4: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCAAAG ATAGACATTCTAAAGTTTTGACTTCCAAGGGTCCAAGAGATAGAAGAGTTAGATTGTCT GTTTCCACCGCTAAAGAATTATTCGGTTTACAAGACATGTTGGGTTACGACCAACCATC TAAAGCTGTTGAATGGTTGATTAAGGCTGCCGAAGATTCCATTTCTGAATTGCCATAAG GCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCA ATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
TCP-Chimaera 5: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCAAAG ATAGACATTCTAAAGTTTTGACTTCCAAGGGTCCAAGAGATAGAAGAGTTAGATTGTCT GTTTCTACCGCCTTGCAATTTTACGACTTGCAAGATAGATTGGGTTTCGATAAGGCTTCT AAAGCTGTTGAATGGTTGATTAAGGCTGCCGAAGATTCCATTTCTGAATTGCCATAAGG CAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAAT TTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
TCP-Chimaera 6: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCAAAG	

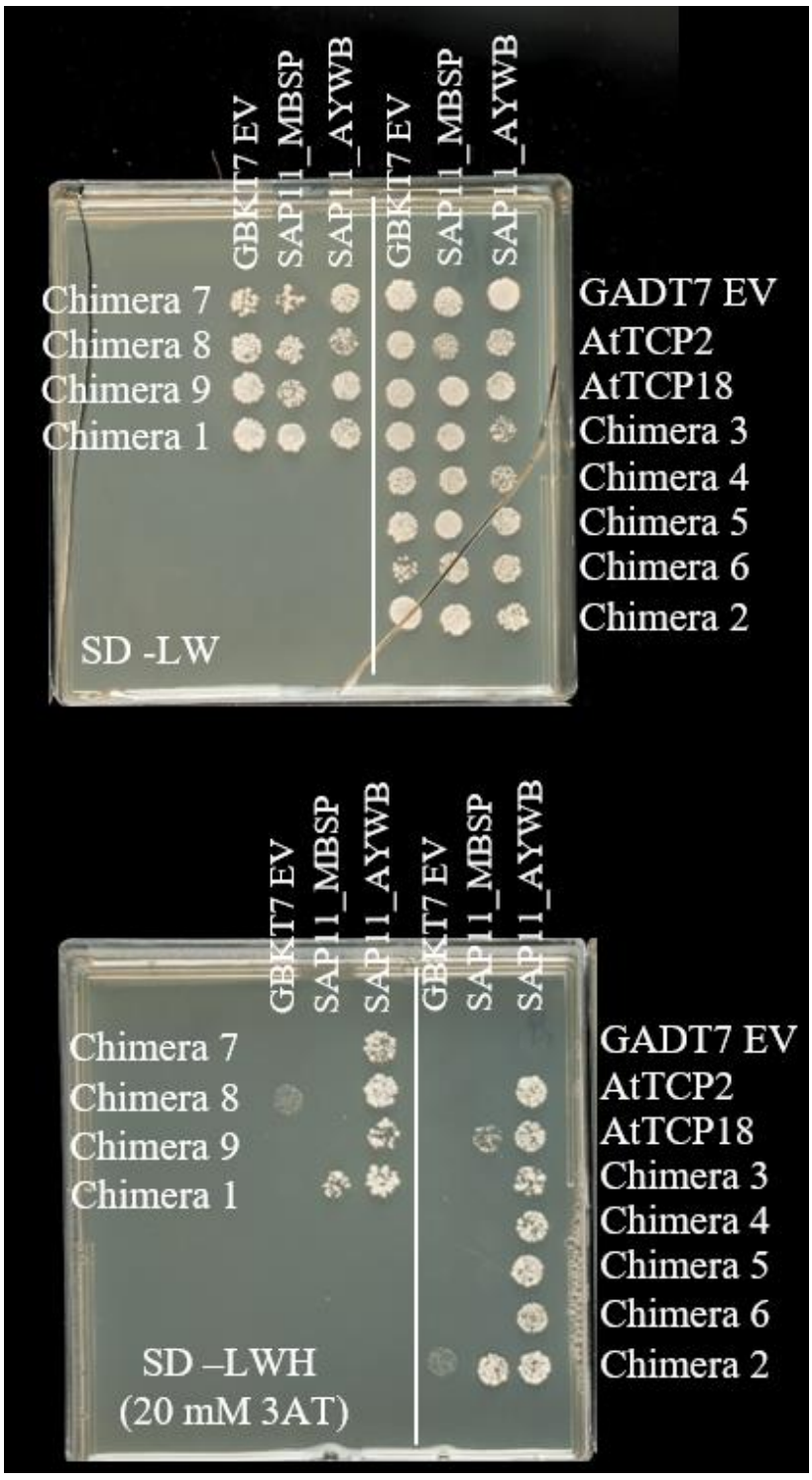
ATAGACATTCTAAAGTTTTGACTTCCAAGGGTCCAAGAGATAGAAGAGTTAGATTGTCT GTTTCTACCGCCTTGCAATTTTACGACTTGCAAGATAGATTGGGTTACGACCAACCATCT AAAGTTGAATGGTTGTTGACTCAAGCCAAGGACTCTATTTCTGAATTGCCATGAGGCAA GCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTG TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
TCP-Chimaera 7: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCAAAG ATAGACATTCTAAAGTTTTGACTTCCAAGGGTCCAAGAGATAGAAGAGTTAGATTGTCT GTTTCTACCGCCTTGCAATTTTACGACTTGCAAGATAGATTGGGTTTCGATAAGGCTTCT AAAACTGTCGAATGGTTGTTGACTCAAGCCAAGGATTCCATTTCTGAATTGCCATAAGG CAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAAT TTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
TCP-Chimaera 8: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCAAAG ATAGACATTCTAAAGTTTTGACTTCCAAGGGTCCAAGAGATAGAAGAGTTAGATTGTCT GTTTCTACCGCCAAAGAATTATTCGGTTTACAAGACATGTTGGGTTTCGATAAGGCTTC TAAAGCTGTTGAATGGTTGATTAAGGCTGCCGAAGATTCCATTTCTGAATTGCCATAAG GCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCA ATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
TCP-Chimaera 9: chimeric TCP domain of *AtTCP2* and *AtTCP18* used in Y2H analysis	**3.2.7**
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCACTG ATAGACATTCTAAAATCAAAACTGCCAAGGGTACTAGAGATAGAAGAATGAGATTGTC CGTTTCTACCGCCAAAGAATTATTCGGTTTACAAGACATGTTGGGTTACGACCAACCAT CTAAAACTGTCGAATGGTTGTTGACTCAAGCCAAGGATTCCATTTCTGAATTGCCATAA GGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATC AATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11$_{ATP}$: codon optimised version of SAP11 effector homolog from *Ca.* P. mali strain AT used in Y2H analysis.	5.2.1
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGT CACCACCTAAGAAAGATAGTAATAAAGGTAAGTCCATAGACAAGAGTGTCTCCTCCAA ACGTGAAACTGTGTCAATACGAGAATACCGAGAATTAGAGATGGCCTTGAACCAGCTC CCAGAGGAGGAGAGGAACACCATTATGGAAACATTAAACAACCCGGAGAAGATGGAG GTACTGTTAAAGAAAGCTCAAGACGAGGCAAATAAAAAACGAGGTGGTTCTAGCAGT AGCCAGCACGACGACAATAATAAGGACAAAGGTAAGAAGTAAGGCAAGCTTGACCCA	

GCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAA CAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11$_{PBIP}$: codon optimised version of SAP11 effector homolog from Poinsettia branch-inducing phytoplasma (PBIP) strain PoiBI used in Y2H analysis.	5.2.1
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGT CACCTAAGAAGGAATCAAGTGACAAAAAGAGAGACATTTCTAAGATCAATAAGAGCG AAGAAAAGAACAAAAAACAGAAGGAGGATATTAAGAGGTTCTACACAATACATAAAG AGTTTAAAGAGTATAGCATAGAAAAGAATAATGAGATAATTAAGATACTAGAGAATCC CGAACTGATGAAAATCCTAAAGCAGAAAGCTGAGGAAGAAACAAAAAACTTAAAGGA GGAGGGGTCCAGTAGCAAGCAGAGTGACGATTCTAAGAAGTAAGGCAAGCTTGACCC AGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGA ACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11$_{solani}$: codon optimised version of SAP11 effector homolog from *Ca.* P. solani strain STOLBUR (AF) used in Y2H analysis.	5.2.1
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGA CCCCCAAGAAGAATCACGGCAAGGACATAATTAGTAGCAAGGAGGAAGCAAAGAAAG ATGTCAAGAATTTCTACGAGCTGTATAACACGTTAGAAAACTATAGTGAGGAGGACCG TTCTAAGATCATCCAGATGCTCTCCAACCCAGAGATAATTAAAAAGCTGAAGGAGAAA ATTGAGGAGGAGAAAAAACAAGAGAAAGGGTCTTCTAGTCGTCAACCGGACAATTGT CATAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAA TTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTG CCATC	
SAP11$_{PLYDY.2}$: codon optimised version of SAP11 effector homolog from *Ca.* P. asteris Periwinkle-leaf yellowing phytoplasma (PLYDY) used in Y2H analysis.	5.2.1
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGG CTCCGAAAAAGCATGGTAAGGACATAATTTCTTCTAAGGAGGAGGTCAAGGACAACGT CAAAAAGTACTACGAACTGTACAATACACTGGAGAATTATTCTGAGAAAGAAAGAAAC AAGATTATCCAGATGCTGAGTGACCCTGCTATCATCAAGACCTTAGAAGAGAAGATAA AAGAAACCAAAACACACGAGAAGGGCTCTTCTTATAAGAAACCCGATAATCTTAAGAA GTAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCT CATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11$_{p45}$: codon optimised version of SAP11 effector homolog from *Ca.* P. asteris sample p45.IC used in Y2H analysis.	5.2.1
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGA GCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGG CGTCAAAAAAGGACCACGAGAAAAACATTATATCCTTGAAGGAGGAAGATAAGAAGG ACGTAAAGAATTACTATGAACTGTACAACACGCTGGAGAACTACAGTGAAGAAGACAG AAACAAAATTATCCAAATGCTTAGCAATCCCGAGATAATAAAGATCCTAGGGGAGAAG	

GTTAAAGAGGCTCAAAATCAAAAGAAAGGGTCCAGTAGCAAAAAGCCGGATGACCTATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11$_{OYM}$: codon optimised version of SAP11 effector homolog from *Ca.* P. asteris Onion Yellows mild-symptom (OY-M) phytoplasma used in Y2H analysis.	5.2.1
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGGCTCCAAAAAAACACGGGAAAGACATCATTTCTTCAAAAGAGGAGGTCAAAGATAACGTGAAAAAATACTATGAACTCTACAATACACTAGAGAACTACTCAGAGGAAGAGAGGAACAAGATTATACAGATGCTGAGCGATCCCGCAATTATAAAGACACTTGAAGAGAAGATCAAAGAAACGAAAACACACGAGAAGGGTAGTAGCTATAAAAAACCAGACAATCTAAAAAAGTAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	
SAP11$_{OYM}$: codon optimised version of SAP11 effector homolog from Faba bean phyllody (FBP) (Faba bean) used in Y2H analysis.	5.2.1
GCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGGTACCGCCATGGCGCCGGAAAAGAACGATAAGGGCAAAAAGATCGCCTCTAGCCAAAAAAAAGAAAAGACCACGAAGAAGGACATCAGCCAGTACTACGAACTGTATAATACTCTGGAGAACTACTCAGAGGAGAACAGGAACAAGATCATTAAGATGTTGAGCGACCCGAAGACTTTAAAGACTTTACAGGAAAAGGCGCTCAAGTCTAAGAAAAAGGGCTCAAGCTCCAAGAAACCTGATGACAGCAAGAAATAAGGCAAGCTTGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATC	

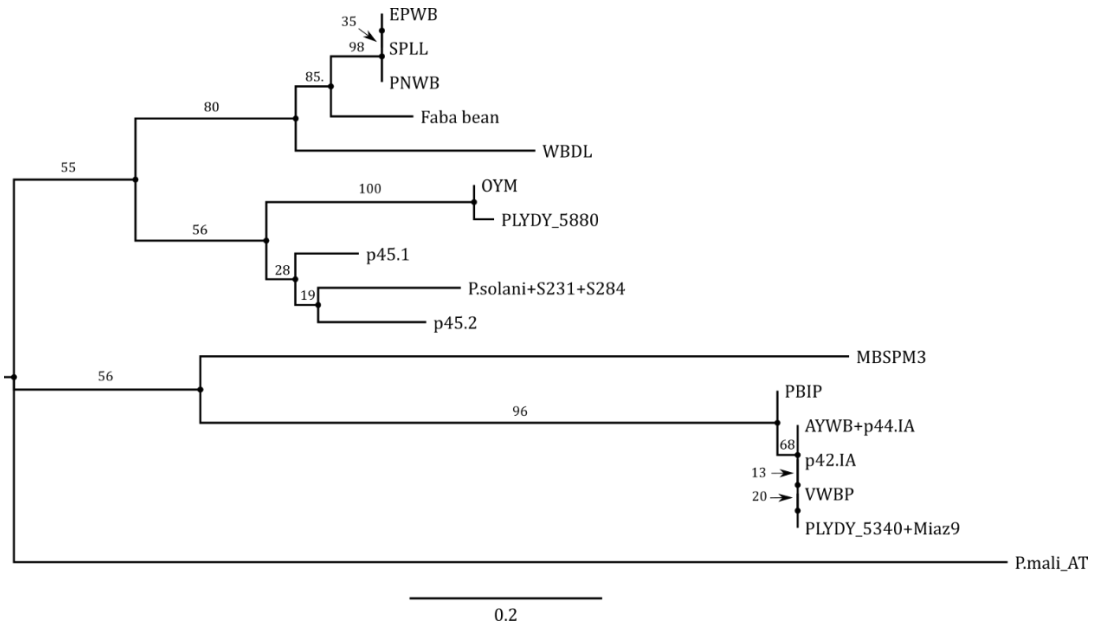
Appendix II – Supplemental Figure for Chapter 3

Appendix II figure 1: Yeast two-hybrid analysis of phytoplasma SAP11$_{AYWB}$, SAP11$_{MBSP}$ and *Arabidopsis thaliana* TCP chimaeras of the TCP domain of AtTCP2 and AtTCP18. (Pecher et al., 2019)



Appendix III – Supplemental Figure for Chapter 4

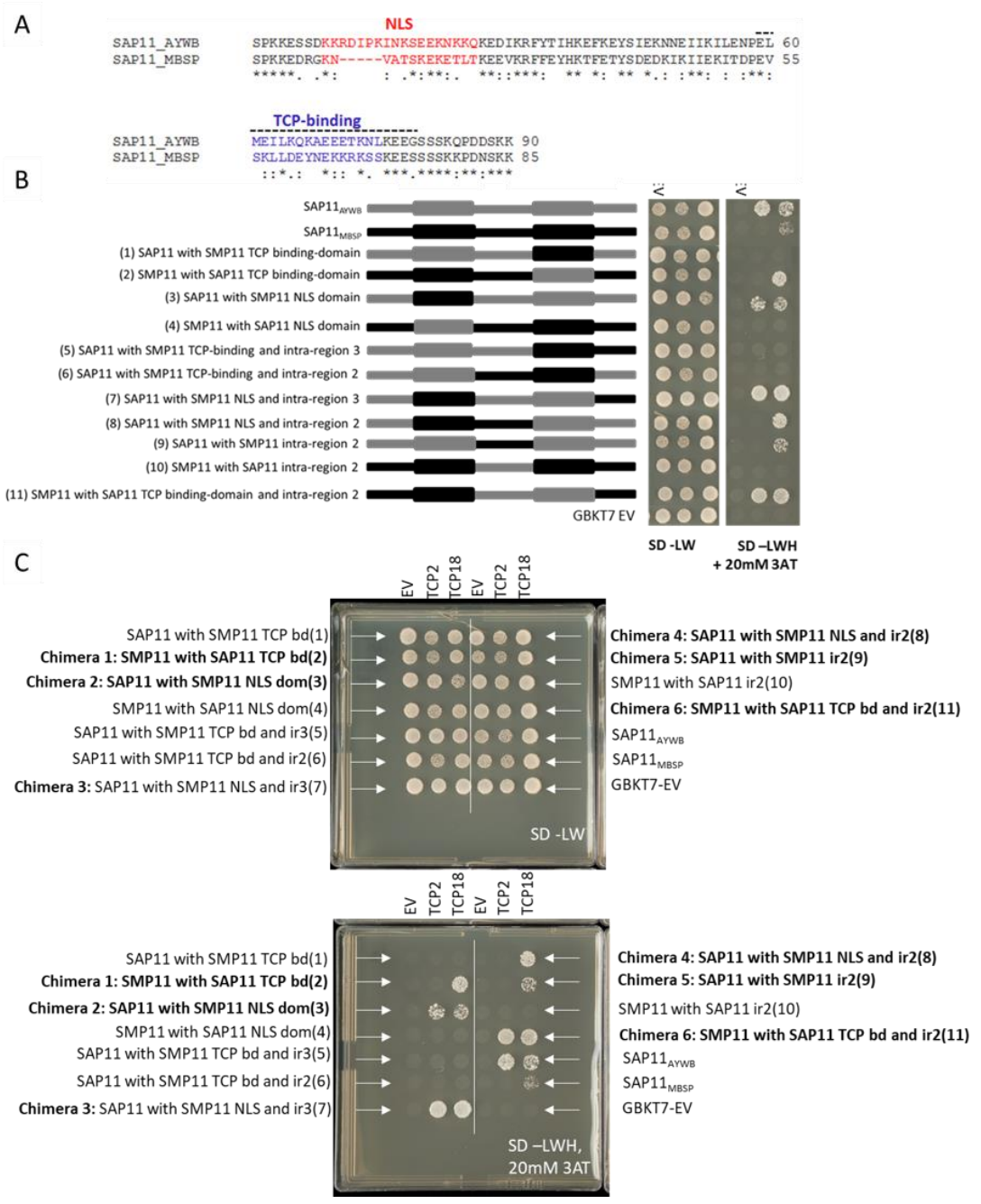
Appendix III figure 1: The phylogenetic tree based on the TCP interaction domains of SAP11 effector homologs is consistent with the SAP11-based phylogenetic tree. Phylogenetic tree generated via the MEGA software using the Maximum Likelihood algorithm. (A) Phylogenetic tree based on the intra-region and the TCP binding domain of SAP11 proteins from different phytoplasma groups. (B) Multiple Alignment of the intra-region and the TCP binding domain of SAP11 proteins from different phytoplasma groups.




```
AYWB+p44.IA          1 KEDIKRFYTIHKEFKEYSIEKNNEIIKILENPELMEILKQKAEEETKNL
p42.IA               1 KEDIKRFYTIHKEFKEYSIEKNNEIIKILENPELMEILKQKAEEETKNL
PBIP                 1 KEDIKRFYTIHKEFKEYSIEKNNEIIKILENPELMKILKQKAEEETKNL
VWBP                 1 KEDIKRFYTIHKEFKEYSIEKNNEIIKILENPELMEILKQKAEEETKNL
PLYDY_5340+Miaz      1 KEDIKRFYTIHKEFKEYSIEKNNEIIKILENPELMEILKQKAEEETKNL
P.mali_AT            1 -VSIREYRELEMALNQLPEEERNTIMETLNNPEKMEVLLKKAQDEANK-
MBSPM3               1 KEEVKRFFEYHKTFETYSDEDKIKIIEKITDPEVSKILDEYNEKKRKSS
PNWB                 1 KKDISQYYELYNTLENYSEEDRNKIIQMLSDSQTLKILQEEALKSKKK-
EPWB                 1 KKDISQYYELYNTLENYSEEDRNKIIQMLSDSQTLKILQEEALKSKKK-
SPLL                 1 KKDISQYYELYNTLENYSEEDRNKIIQMLSDSQTLKILQEEALKSKKK-
Faba_bean            1 KKDISQYYELYNTLENYSEENRNKIIKMLSDPKTLKTLQEKALKSKKK-
WBDL                 1 KKDISQYYELYNTLKDYSEEDQNKIIQILSDPEISKILQEQKLKSQKT-
OYM                  1 KDNVKKYYELYNTLENYSEEERNKIIQMLSDPAIIKTLEEKIKETKTH-
PLYDY_5880           1 KDNVKKYYELYNTLENYSEKERNKIIQMLSDPAIIKTLEEKIKETKTH-
P.solani+S231+S      1 KKDVKNFYELYNTLENYSEEDRSKIIQMLSNPEIIKKLKEKIEEEKKQ-
p45.1                1 KKDVKNYYELYNTLENYSEEDRNKIIQMLSNPEIIKILGEKVKEAQNQ-
p45.2                1 KKDIKNFYELYNTLENYSEEERNKIIQMLSNPEITKILEEKTKEIKNQ-
```

Appendix IV – Supplemental Figures for Chapter 5

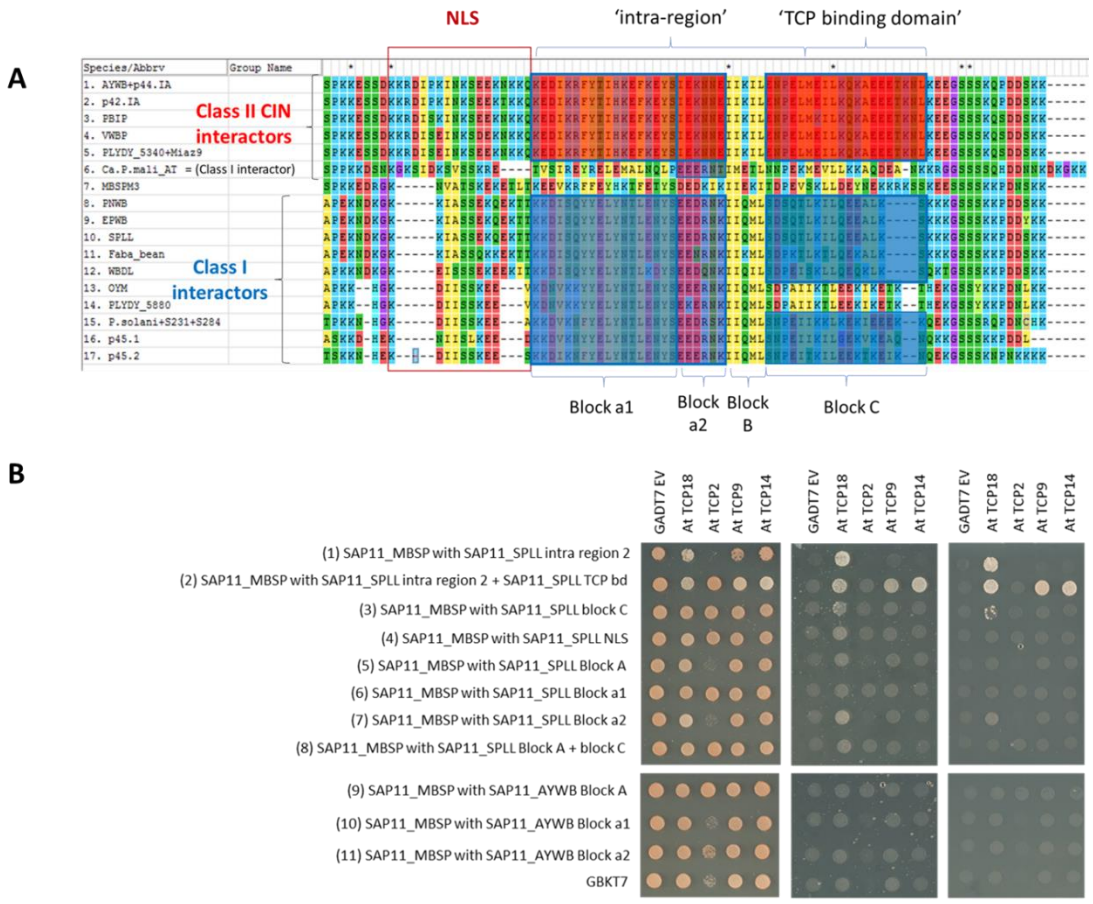
Appendix IV figure 1: Yeast two-hybrid analysis of phytoplasma SAP11 chimeric versions based on SAP11$_{AYWB}$ and SAP11$_{MBSP}$ and *Arabidopsis thaliana* TCPs.



Legend of Appendix IV figure 1: The details of the assay are identical to **Figure 5.2.5**. The SAP11 chimaeras are based from AYWB phytoplasma SAP11 effector (SAP11$_{AYWB}$) and MBSP phytoplasma SAP11 effector (SAP11$_{MBSP}$). The signal peptide is removed from the sequences. (**A**) Multiple Sequence Alignment using Clustal 2.1 program of AYWB and MBSP phytoplasma

SAP11 amino acid sequences, without the signal peptide. The nuclear localisation signal of $SAP11_{AYWB}$ (NLS) is highlighted in red, the coiled-coil domain with a dashed line, and part of the coiled-coil structure required for binding TCP in blue. *, Fully conserved residues, :, conservation of residues with strongly similar properties and ., conservation of residues with weakly similar properties. (**B**) Schematic representation of the eleven chimeric constructs used for the Y2H assay. As an example, "SAP11 with SMP11 TCP bd" indicates the chimeric construct #1, which is the backbone of $SAP11_{AYWB}$ and the addition of the TCP binding domain of $SAP11_{MBSP}$. SAP11= $SAP11_{AYWB}$; SMP11= $SAP11_{MBSP}$; TCP bd= TCP binding domain (Sugio et al., 2014); NLS dom= NLS domain; ir2=intra-region 2; ir3= intra-region 3 (**C**) Original pictures of the Y2H assay of the SAP11 chimaeras based on $SAP11_{AYWB}$ and $SAP11_{MBSP}$ (**Figure 5.2.5**). In bold are the constructs displayed in **Figure 5.2.5**. The rest of the constructs are not in the result chapters as no yeast growth was visible in the selective media SD -LWH (20mM 3AT). The figure **C** can be found in our latest paper (Pecher et al., 2019).

Appendix IV figure 2: Further characterisation of the binding specificity region of Class I interactor SAP11$_{SPLL}$.



Legend of Appendix IV figure 2: (**A**) Based on the previous results of the SAP11 chimaera Y2H assays (**Figure 5.2.5**), I designed the chimeric constructs depending on block A, block B and block C, with block A further subdivided into block a1 and block a2. (**B**) Yeast-Two-Hybrid analysis showing the interaction between SAP11 chimeric versions based on SAP11$_{AYWB}$, SAP11$_{SPLL}$ and SAP11$_{MBSP}$ and *Arabidopsis thaliana* TCPs. TCP candidates selected from the Class II CYC/TB1-TCP (sub)class and the Class I TCP. SAP11 chimeric versions are based on the Multiple Sequence Alignment of SAP11 effector homologs. SAP11 chimeric proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCPs and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Yeast colonies co-expressing bait and prey or empty plasmid (pDEST-GADT7) combinations were grown in synthetic dropout (SD) media lacking either leucine and tryptophan (-L,-W) (left); leucine, tryptophan and histidine with the addition of 20 mM 3-Amino-1,2,4-triazole (3AT) (used to suppress auto-activation) (-L, -W, -H) (middle) or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H)

(right); the former indicating the presence of both plasmids and the two later indicating interactions of bait and prey. The experiment has been repeated three times with three different batches of transformed yeasts.

Appendix V – Supplemental Figures for Chapter 6

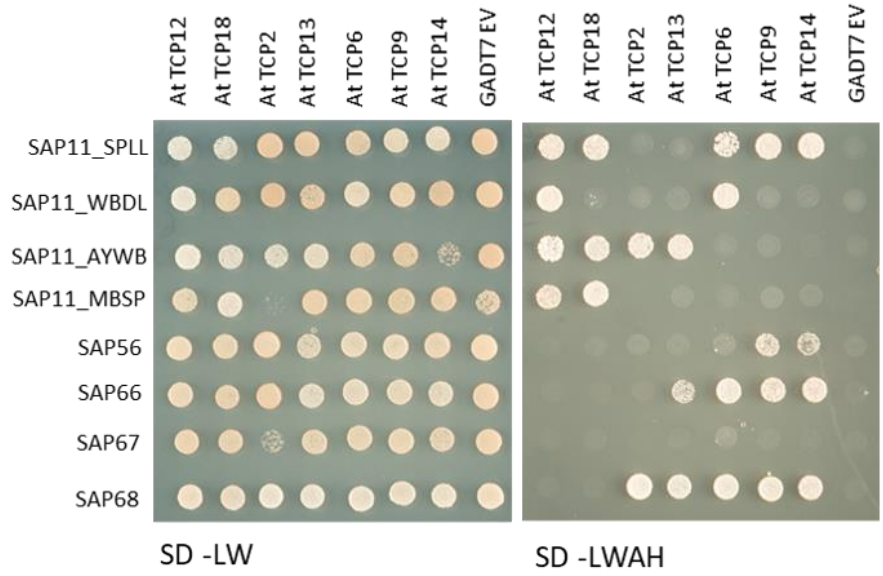
Appendix V figure 1: Multiple Sequence Alignment of AY-WB effectors SAP11, SAP56, SAP66, SAP67 and SAP68. The AY-WB sequences of SAP11, SAP56, SAP66, SAP67 and SAP68 were obtained from NCBI. The signal peptides of the protein sequences were removed. The Genbank accession numbers are ABC65484.1 (SAP56), ABC65483.1 (SAP66), ABC65485.1 (SAP67) and ABC65486.1 (SAP68). Then, I aligned the protein sequences using T coffee, default parameters (http://tcoffee.crg.cat/apps/tcoffee/do:regular). I then used Boxshade (http://www.ch.embnet.org/software/BOX_form.html) to produce the alignment displayed here.

```
SAP11_AYWB     1 SPKKESS-DKKRDIPK--IN-------------KSEEKNKKQKEDIKRFYTIHKEFKEYS
SAP56_AYWB     1 KIRLEERKNE-L---------------------------------KSQEIVLSQEP--RN
SAP66_AYWB     1 MNNDNN--NNEIRNE------------------IREINNRIL-ELALEK----KQL--SN
SAP67_AYWB     1 MGNKNSN-NNEISNDEEYALYVQAEFSIQNELTNFKFNNEKKSELLYKQTCITKAIQKYN
SAP68_AYWB     1 MHNGNATPNNGHHNT------------------NFETQARILQEMNREQAIIVQQI--FN

SAP11_AYWB    45 IEKNNEIIK-ILEN-PELMEILKQKAEEETKNLKEE------------------GSSSKQ
SAP56_AYWB    25 NVNIERLCH--VRT--EIV-KINMEIYMIMQQIQMRD-II---Q----------------
SAP66_AYWB    34 QMSLNPIRS--FAN-LELN-IRNRAIYLELNNLYDRLRII---G--------QQPINPNQ
SAP67_AYWB    60 KRKNNHQLSTPYDNRPESFQNNLHSLTENNNNLSGDLQYINLTKNNSNIYSVDKPES-YK
SAP68_AYWB    41 ARNNNASKE--IIN--NLV-RQNIQLSQRI---SNQQ-II---LH----NAMPHENNRNQ

SAP11_AYWB    85 PDDSKK-------
SAP56_AYWB       -------------
SAP66_AYWB    79 PNNPLN--R----
SAP67_AYWB   119 KNDSKNKGVIIND
SAP68_AYWB    85 LNNSNNRRR----
```

Appendix V figure 2: SAP effectors from SAP11$_{AYWB}$ genetic island can interact with TCP, including Class I TCPs. Yeast two-hybrid analysis of AY-WB phytoplasma SAP effectors SAP56, SAP66, SAP67 and SAP68, belonging to the genetic island of SAP11$_{AYWB}$ and *Arabidopsis thaliana* Class I, Class II CIN and Class II CYC/TB1-TCPs. SAP effector proteins and pDEST-GBKT7 empty vector control were fused to the DNA binding domain of the GAL4 transcriptional activator (bait) while TCP domains and pDEST-GADT7 empty vector control were fused to the transcription activation domain of GAL4 (prey). Phytoplasma SAP11 effector homologs SAP11$_{AYWB}$, SAP11$_{MBSP}$, SAP11$_{WBDL}$ and SAP11$_{SPLL}$ were used as positive control for interaction with the different TCP (sub)classes. Yeast colonies co-expressing bait and prey or empty plasmids combinations were grown in synthetic dropout (SD) media lacking either leucine and tryptophan (-L,-W) (left) with growths of colonies indicating the presence of both plasmids; or leucine, tryptophan, adenine and histidine (-L, -W, -A, -H) (right) with growths of colonies indicating interactions of bait (SAP11 homologs) and prey (TCP homologs). The experiment was repeated three times with three different batches of transformed yeasts and these showed identical results as shown. This repetition does not include GBKT7 combinations so is kept as an appendix.



Appendix VI – Published work

Pecher, P., Moro, G., Canale, M. C., Capdevielle, S., Singh, A., MacLean, A., … Hogenhout, S. A. (2019). Phytoplasma SAP11 effector destabilization of TCP transcription factors differentially impact development and defence of Arabidopsis versus maize. *PLoS Pathogens*, in press . See **Appendix VI**.

Abstract:

Phytoplasmas are insect-transmitted bacterial pathogens that colonize a wide range of plant species, including vegetable and cereal crops, and herbaceous and woody ornamentals. Phytoplasma-infected plants often show dramatic symptoms, including proliferation of shoots (witch's brooms), changes in leaf shapes and production of green sterile flowers (phyllody). Aster Yellows phytoplasma Witches' Broom (AY-WB) infects dicots and its effector, secreted AYWB protein 11 (SAP11), was shown to be responsible for the induction of shoot proliferation and leaf shape changes of plants. SAP11 acts by destabilizing TEOSINTE BRANCHED 1-CYCLOIDEA-PROLIFERATING CELL FACTOR (TCP) transcription factors, particularly the class II TCPs of the CYCLOIDEA/TEOSINTE BRANCHED 1 (CYC/TB1) and CINCINNATA (CIN)-TCP clades. SAP11 homologs are also present in phytoplasmas that cause economic yield losses in monocot crops, such as maize, wheat and coconut. Here we show that a SAP11 homolog of Maize Bushy Stunt Phytoplasma (MBSP), which has a range primarily restricted to maize, destabilizes only TB1/CYC TCPs. $SAP11_{MBSP}$ and $SAP11_{AYWB}$ both induce axillary branching and $SAP11_{AYWB}$ also alters leaf development of *Arabidopsis thaliana* and maize. However, only in maize, $SAP11_{MBSP}$ prevents female inflorescence development, phenocopying maize tb1 lines, whereas $SAP11_{AYWB}$ prevents male inflorescence development and induces feminization of tassels. $SAP11_{AYWB}$ promotes fecundity of the AY-WB leafhopper vector on A. thaliana and modulates the expression of A. thaliana leaf defence response genes that are induced by this leafhopper, in contrast to $SAP11_{MBSP}$. Neither of the SAP11 effectors promote fecundity of AY-WB and MBSP leafhopper vectors on maize. These data provide evidence that class II TCPs have overlapping but also distinct roles in regulating development and defence in a dicot and a monocot plant species that is likely to shape SAP11 effector evolution depending on the phytoplasma host range.

References

Aggarwal, P., Das Gupta, M., Joseph, A.P., Chatterjee, N., Srinivasan, N., and Nath, U. (2010). Identification of Specific DNA Binding Residues in the TCP Family of Transcription Factors in Arabidopsis. Plant Cell **22**: 1174–1189.

Aguilar-Martinez, J. a., Poza-Carrion, C., and Cubas, P. (2007). Arabidopsis *BRANCHED1* Acts as an Integrator of Branching Signals within Axillary Buds. Plant Cell **19**: 458–472.

Aguilar-Martínez, J.A. and Sinha, N. (2013). Analysis of the role of Arabidopsis class I *TCP* genes *AtTCP7*, *AtTCP8*, *AtTCP22*, and *AtTCP23* in leaf development. Front. Plant Sci. **4**: 1–13.

Alfano, J.R. (2009). Roadmap for future research on plant pathogen effectors. Mol Plant Pathol **10**: 805–813.

Alvarez, E., Mejía, J.F., Contaldo, N., Paltrinieri, S., Duduk, B., and Bertaccini, A. (2014). '*Candidatus* Phytoplasma asteris' Strains Associated with Oil Palm Lethal Wilt in Colombia. Plant Dis. **98**: 311–318.

Andersen, M.T., Liefting, L.W., Havukkala, I., and Beever, R.E. (2013). Comparison of the complete genome sequence of two closely related isolates of "*Candidatus* Phytoplasma australiense" reveals genome plasticity. BMC Genomics **14**: 1–15.

Asai, T., Stone, J.M., Heard, J.E., Kovtun, Y., Yorgey, P., Sheen, J., and Ausubel, F.M. (2000). Fumonisin B1-induced cell death in Arabidopsis protoplasts requires jasmonate-, ethylene-, and salicylate-dependent signaling pathways. Plant Cell **12**: 1823–1835.

Bai, X., Ammar, E.D., and Hogenhout, S.A. (2007). A secreted effector protein of AY-WB phytoplasma accumulates in nuclei and alters gene expression of host plant cells, and is detected in various tissues of the leafhopper *Macrosteles quadrilineatus*. Bull. Insectology **60**: 217–218.

Bai, X., Correa, V.R., Toruño, T.Y., Ammar, E.-D., Kamoun, S., and Hogenhout, S. a (2009). AY-WB phytoplasma secretes a protein that targets Plant Cell Nuclei. Mol. Plant. Microbe. Interact. **22**: 18–30.

Bai, X., Zhang, J., Ewing, A., Miller, S.A., Radek, A.J., Shevchenko, D. V., Tsukerman, K., Walunas, T., Lapidus, A., Campbell, J.W., and Hogenhout, S.A. (2006). Living with genome instability: The adaptation of phytoplasmas to diverse environments of their insect and plant hosts. J. Bacteriol. **188**: 3682–3696.

Beanland, L., Hoy, C.W., Miller, S. a., and Nault, L.R. (2000). Influence of Aster Yellows Phytoplasma on the Fitness of Aster Leafhopper (Homoptera: Cicadellidae). Ann. Entomol. Soc. Am. **93**: 271–276.

Bell, E., Creelman, R.A., and Mullet, J.E. (1995). A chloroplast lipoxygenase is required for wound-induced jasmonic acid accumulation in *Arabidopsis*. Proc. Natl. Acad. Sci. **92**: 8675–8679.

Bemer, M., van Dijk, A.D.J., Immink, R.G.H., and Angenent, G.C. (2017). Cross-Family Transcription Factor Interactions: An Additional Layer of Gene Regulation. Trends Plant Sci. **22**: 66–80.

Bertaccini, A. (2007). Phytoplasmas: diversity, taxonomy, and epidemiology. Front. Biosci. **12**: 673–689.

Blanc, G., Barakat, A., Guyot, R., Cooke, R., and Delseny, M. (2000). Extensive Duplication and Reshuffling in the Arabidopsis Genome. Plant Cell Online **12**: 1093–1102.

Blanc, G., Hokamp, K., and Wolfe, K.H. (2003). A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. Genome Res. **13**: 137–144.

Boller, T. and He, S.Y. (2012). Innate Immunity in Plants: An Arms Race Between Pattern Recognition Receptors in Plants and Effectors in Microbial Pathogens. Science (80-.). **742**: 742–745.

Boonrod, K., Munteanu, B., Jarausch, B., Jarausch, W., and Krczal, G. (2012). An Immunodominant Membrane Protein (Imp) of ' *Candidatus* Phytoplasma mali' Binds to Plant Actin. Mol. Plant-Microbe Interact. **25**: 889–895.

Bowers, J.E., Chapman, B.A., Rong, J., and Paterson, A.H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature **422**: 433–438.

Bresso, E.G., Chorostecki, U., Rodriguez, R.E., Palatnik, J.F., and Schommer, C. (2017). Spatial control of gene expression by miR319-regulated TCP transcription factors in leaf development. Plant Physiol. **176**: pp.00823.2017.

Browse, J. (2009). Jasmonate Passes Muster: A Receptor and Targets for the Defense Hormone. Annu. Rev. Plant Biol. **60**: 183–205.

Buxa, S. V., Degola, F., Polizzotto, R., De Marco, F., Loschi, A., Kogel, K.-H., Toppi, L.S. di, Bel, A.J.E. va., and Musetti, R. (2015). Phytoplasma infection in tomato is associated with re-organization of plasma membrane, ER stacks, and actin filaments in sieve elements. Front. Plant Sci. **6**: 1–11.

Capinera, J.L. (2008). Encyclopedia of Entomology J.L. Capinera, ed.

Chai, W., Jiang, P., Huang, G., Jiang, H., and Li, X. (2017). Identification and expression profiling analysis of TCP family genes involved in growth and development in maize. Physiol. Mol. Biol. Plants **23**: 779–791.

Chang, S.-H., Cho, S.-T., Chen, C.-L., Yang, J.-Y., and Kuo, C.-H. (2015). Draft Genome Sequence of a 16SrII-A Subgroup Phytoplasma Associated with Purple Coneflower (*Echinacea purpurea*) Witches' Broom Disease in Taiwan. Genome Announc. **3**: e01398-15.

Chang, S.H., Tan, C.M., Wu, C.-T., Lin, T.-H., Jiang, S.-Y., Liu, R.-C., Tsai, M.-C., Su, L.-W., and Yang, J.-Y. (2018). Alterations of plant architecture and phase transition by the phytoplasma virulence factor SAP11. J. Exp. Bot.: 1–13.

Chen, W., Li, Y., Wang, Q., Wang, N., and Wu, Y. (2014). Comparative genome analysis of wheat blue dwarf phytoplasma, an obligate pathogen that causes wheat

blue dwarf disease in China. PLoS One **9**: 1–11.

Chen, W.Y., Huang, Y.C., Tsai, M.L., and Lin, C.P. (2011). Detection and identification of a new phytoplasma associated with periwinkle leaf yellowing disease in Taiwan. Australas. Plant Pathol. **40**: 476–483.

Chiykowski, L.N. (1990). Vector-pathogen-host plant relationships of clover phyllody mycoplasmalike organism and the vector leafhopper *Paraphlepsius irroratus*. Can. J. Plant Pathol. **13**: 11–18.

Christensen, N.M., Axelsen, K.B., Nicolaisen, M., and Schulz, A. (2005). Phytoplasmas and their interactions with hosts. Trends Plant Sci. **10**: 526–535.

Chung, W.C., Chen, L.L., Lo, W.S., Lin, C.P., and Kuo, C.H. (2013). Comparative Analysis of the Peanut Witches'-Broom Phytoplasma Genome Reveals Horizontal Transfer of Potential Mobile Units and Effectors. PLoS One **8**: 1–10.

Croll, D. and McDonald, B.A. (2012). The accessory genome as a cradle for adaptive evolution in pathogens. PLoS Pathog. **8**: 8–10.

Cubas, P., Lauter, N., Doebley, J., and Coen, E. (1999). The TCP domain: A motif found in proteins regulating plant growth and development. Plant J. **18**: 215–222.

Danisman, S. (2016). TCP Transcription Factors at the Interface between Environmental Challenges and the Plant's Growth Responses. Front. Plant Sci. **7**: 1–13.

Danisman, S., Van Dijk, A.D.J., Bimbo, A., Van Der Wal, F., Hennig, L., De Folter, S., Angenent, G.C., and Immink, R.G.H. (2013). Analysis of functional redundancies within the Arabidopsis TCP transcription factor family. J. Exp. Bot. **64**: 5673–5685.

Danisman, S., van der Wal, F., Dhondt, S., Waites, R., de Folter, S., Bimbo, A., van Dijk, A.D.J., Muino, J.M., Cutri, L., Dornelas, M.C., Angenent, G.C., and Immink, R.G.H. (2012). Arabidopsis class I and class II TCP transcription factors regulate jasmonic acid metabolism and leaf development antagonistically. Plant Physiol. **159**: 1511–23.

Davière, J.M., Wild, M., Regnault, T., Baumberger, N., Eisler, H., Genschik, P., and Achard, P. (2014). Class I TCP-DELLA interactions in inflorescence shoot apex determine plant height. Curr. Biol. **24**: 1923–1928.

Davis, R.E. and Dally, E.L. (2001). Revised Subgroup Classification of Group 16SrV Phytoplasmas and Placement of Flavescence Dorée -Associated Phytoplasmas in Two Distinct Subgroups. Plant Dis. **85**.

Davis, R.E., Jomantiene, R., Kalvelyte, A., and Dally, E.L. (2003). Differential amplification of sequence heterogeneous ribosomal RNA genes and classification of the "*Fragaria multicipita*" phytoplasma. Microbiol. Res. **158**: 229–236.

Davis, R.E., Zhao, Y., Dally, E.L., Lee, I.M., Jomantiene, R., and Douglas, S.M. (2013). "*Candidatus* Phytoplasma pruni", a novel taxon associated with X-disease of stone fruits, Prunus spp.: Multilocus characterization based on 16S rRNA, secY, and ribosomal protein genes. Int. J. Syst. Evol. Microbiol. **63**: 766–776.

Dhaka, N., Bhardwaj, V., Sharma, M.K., and Sharma, R. (2017). Evolving Tale of TCPs:

New Paradigms and Old Lacunae. Front. Plant Sci. **8**: 1–8.

Dickinson, M. (2010). Mobile units of DNA in phytoplasma genomes. Mol. Microbiol. **77**: 1351–1353.

Doebley, J., Stec, A., and Gustus, C. (1995). *teosinte branched1* and the Origin of Maize: Evidence for Epistasis and the Evolution of Dominance. Maize Genet. Coop. Newsl. **33**: 74.

Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize.: 485–488.

Doi, Y., Teranaka, M., Yora, K., and Asuyama, H. (1967). Mycoplasma or PLT-group-like microorganisms found in the phloem el- ements of plants infected with mulberry dwarf, potato witches' broom, aster yellows, or paulownia witches' broom. Ann. Phytopathol. Soc. Jpn. **33**: 259–266.

Dong, S. et al. (2014). Effector Specialization in a Lineage of the Irish Potato Famine Pathogen. Science (80-.). **343**: 552–555.

Dong, S., Raffaele, S., and Kamoun, S. (2015). The two-speed genomes of filamentous pathogens: waltz with plants. Curr. Opin. Genet. Dev. **35**: 021774.

Dong, Z., Li, W., Unger-Wallace, E., Yang, J., Vollbrecht, E., and Chuck, G. (2017). Ideal crop plant architecture is mediated by tassels replace upper ears1, a BTB/POZ ankyrin repeat gene directly targeted by *TEOSINTE BRANCHED1*. Proc. Natl. Acad. Sci. U. S. A. **114**: E8656–E8664.

Dumonceaux, T.J., Green, M., Hammond, C., Perez, E., and Olivier, C. (2014). Molecular diagnostic tools for detection and differentiation of phytoplasmas based on *chaperonin-60* reveal differences in host plant infection patterns. PLoS One **9**: 1–21.

Ebbert, M.A., Jeffers, D.P., Harrison, N.A., and Nault, L.R. (2001). Lack of specificity in the interaction between two maize stunting pathogens and field collected *Dalbulus* leafhoppers. Entomol. Exp. Appl. **101**: 49–57.

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

Efroni, I., Blum, E., Goldshmidt, A., and Eshed, Y. (2008). A Protracted and Dynamic Maturation Schedule Underlies Arabidopsis Leaf Development. Plant Cell Online **20**: 2293–2306.

Errampalli, D., Fletcher, J., and Claypool, P. (1991). Incidence of Yellows in Carrot and Lettuce and Characterization of Mycoplasmalike Organism Isolates in Oklahoma. Plant Dis. **75**: 579–584.

Evans, A.S. (1976). Causation and disease: the Henle-Koch postulates revisited. Yale J. Biol. Med. **49**: 175–195.

Faghihi, S.M.S.M.I.K.S.M.B.A. (2007). Transmission of "*Candidatus* Phytoplasma aurantifolia" to Bakraee (*Citrus reticulata* Hybrid) by Feral *Hishimonus phycitis* Leafhoppers in Iran. Plant Dis. **v. 91**.

Farmer, L.M., Book, A.J., Lee, K.-H., Lin, Y.-L., Fu, H., and Vierstra, R.D. (2010). The RAD23 family provides an essential connection between the 26S proteasome and ubiquitylated proteins in Arabidopsis. Plant Cell **22**: 124–142.

Finlayson, S.A. (2007). Arabidopsis TEOSINTE BRANCHED1-LIKE 1 regulates axillary bud outgrowth and is homologous to monocot TEOSINTE BRANCHED1. Plant Cell Physiol. **48**: 667–677.

Fischer, A., Santana-cruz, I., Wambua, L., Olds, C., Midega, C., Dickinson, M., Kawicha, P., Khan, Z., Masiga, D., Jores, J., and Schneider, B. (2016). Draft Genome Sequence of " *Candidatus* Phytoplasma oryzae " Strain Mbita1 , the Causative Agent of Napier Grass Stunt Disease in Kenya. Genome Announc. **4**: 1–2.

Frost, K.E., Willis, D.K., and Groves, R.L. (2011). Detection and Variability of Aster Yellows Phytoplasma Titer in Its Insect Vector, *Macrosteles quadrilineatus* (Hemiptera: Cicadellidae). J Econ Entomol **104**: 1800–1815.

Galetto, L., Bosco, D., Balestrini, R., Genre, A., Fletcher, J., and Marzachì, C. (2011). The major antigenic membrane protein of "*Candidatus* Phytoplasma asteris" selectively interacts with ATP synthase and actin of leafhopper vectors. PLoS One **6**: e22571.

Garner, C.M., Kim, S.H., Spears, B.J., and Gassmann, W. (2016). Express yourself: Transcriptional regulation of plant innate immunity. Semin. Cell Dev. Biol. **56**: 150–162.

Gasparich, G.E. (2010). Spiroplasmas and phytoplasmas: Microbes associated with plant hosts. Biologicals **38**: 193–203.

Gedvilaite, a and Sasnauskas, K. (1994). Control of the expression of the *ADE2* gene of the yeast Saccharomyces cerevisiae. Curr Genet **25**: 475–479.

Gibb, K.S., Padovan, A.C., and Mogen, B.D. (1995). Studies on sweet potato little-leaf phytoplasma detected in sweet potato and other plant species growing in Northern Australia. Phytopathology **85**: 169–174.

Gonsalves, D. (1998). Control of papaya ringspot virus in papaya: a case study. Annu. Rev. Phytopathol. **36**: 415–437.

Gonzalez-Grandio, E., Pajoro, A., Franco-Zorrilla, J.M., Tarancon, C., Immink, R.G.H., and Cubas, P. (2017). Abscisic acid signaling is controlled by a *BRANCHED1/HD-ZIP I* cascade in Arabidopsis axillary buds. Proc. Natl. Acad. Sci. U. S. A. **114**: E245–E254.

Green, R. and Noller, H.F. (1997). Ribosomes and Translation. Annu. Rev. Biochem. **66**: 679–716.

Guindon, S. and Gascuel, O. (2003). A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Syst. Biol. **52**: 696–704.

Gundersen, D.E., Lee, I.M., Rehner, S.A., Davis, R.E., and Kingsbury, D.T. (1994). Phylogeny of mycoplasma-like organisms (phytoplasmas): A basis for their classification. J. Bacteriol. **176**: 5244–5254.

Guo, S., Xu, Y., Liu, H., Mao, Z., Zhang, C., Ma, Y., Zhang, Q., Meng, Z., and Chong, K. (2013). The interaction between OsMADS57 and OsTB1 modulates rice tillering via

DWARF14. Nat. Commun. **4**: 1566.

Guttman, D.S., Vinatzer, B.A., and Sarkar, S.F. (2002). A Functional Screen for the Type III (Hrp) Secretome of the Plant Pathogen *Pseudomonas syringae*. Science (80-.). **295**: 1722–1726.

Hahn, M.W. (2009). Distinguishing among evolutionary models for the maintenance of gene duplicates. J. Hered. **100**: 605–617.

He, P., Shan, L., Lin, N., Martin, G.B., Kemmerling, B., and Nu, T. (2006). Specific Bacterial Suppressors of MAMP Signaling Upstream of MAPKKK in Arabidopsis Innate Immunity. Cell **125**: 563–575.

He, P., Shan, L., and Sheen, J. (2007). The Use of Protoplasts to Study Innate Immune Responses. In Plant-Pathogen Interactions: Methods and Protocols, P.C. Ronald, ed (Humana Press: Totowa, NJ), pp. 1–9.

Hervé, C., Dabos, P., Bardet, C., Jauneau, A., Auriac, M.C., Ramboer, A., Lacout, F., and Tremousaygue, D. (2009). In vivo interference with AtTCP20 function induces severe plant growth alterations and deregulates the expression of many genes important for development. Plant Physiol. **149**: 1462–1477.

Hogenhout, S. a., Oshima, K., Ammar, E.-D., Kakizawa, S., Kingdom, H.N., and Namba, S. (2008). Phytoplasmas: bacteria that manipulate plants and insects. Mol. Plant Pathol. **9**: 403–423.

Hogenhout, S.A. and Loria, R. (2008). Virulence mechanisms of Gram-positive plant pathogenic bacteria. Curr. Opin. Plant Biol. **11**: 449–456.

Hogenhout, S.A. and Music, M.S. (2010). Phytoplasma Genomics, from Sequencing to Comparative and Functional Genomics - What Have We Learnt? In Phytoplasmas: Genomes, Plant Hosts and Vector, P.G. Weintraub and P. Jones, eds, pp. 19–36.

Hoshi, A., Oshima, K., Kakizawa, S., Ishii, Y., Ozeki, J., Hashimoto, M., Komatsu, K., Kagiwada, S., Yamaji, Y., and Namba, S. (2009). A unique virulence factor for proliferation and dwarfism in plants identified from a phytopathogenic bacterium. Proc. Natl. Acad. Sci. U. S. A. **106**: 6416–6421.

Howarth, D.G. and Donoghue, M.J. (2005). DUPLICATIONS IN *CYC*-like Genes from Dipsacales Correlate with Floral Form. Int. J. Plant Sci. **166**: 357–370.

Howe, G.A. and Jander, G. (2008). Plant Immunity to Insect Herbivores. Annu. Rev. Plant Biol. **59**: 41–66.

Hu, C.C., Hsu, Y.H., and Lin, N.S. (2009). Satellite RNAs and satellite viruses of plants. Viruses **1**: 1325–1350.

Hubbard, L., McSteen, P., Doebley, J., and Hake, S. (2002). Expression patterns and mutant phenotype of *teosinte branched1* correlate with growth suppression in maize and teosinte. Genetics **162**: 1927–1935.

IRPCM (2004). "*Candidatus* Phytoplasma", a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. Int. J. Syst. Evol. Microbiol. **54**: 1243–1255.

Ishii, Y., Kakizawa, S., and Oshima, K. (2013). New ex vivo reporter assay system reveals that σ factors of an unculturable pathogen control gene regulation involved in the host switching between insects and plants. Microbiologyopen **2**: 553–565.

Iwabuchi, N., Maejima, K., Kitazawa, Y., Miyatake, H., Nishikawa, M., Tokuda, R., Koinuma, H., Miyazaki, A., Nijo, T., Oshima, K., Yamaji, Y., and Namba, S. (2019). Crystal structure of phyllogen, a phyllody-inducing effector protein of phytoplasma. Biochem. Biophys. Res. Commun. **513**: 952–957.

Janik, K., Mithöfer, A., Raffeiner, M., Stellmach, H., Hause, B., and Schlink, K. (2017). An effector of apple proliferation phytoplasma targets TCP transcription factors—a generalized virulence strategy of phytoplasma? Mol. Plant Pathol. **18**: 435–442.

Johansson, K.-E., Heldtander, M., and Pettersson, B. (2002). Characterization of Mycoplasmas by PCR and Sequence Analysis with Universal 16S rDNA Primers. In Mycoplasma Protocols, pp. 145–166.

Jomantiene, R., Davis, R.E., Maas, J., and Dally, E.L. (1998). Classification of new phytoplasmas associated with diseases of strawberry in Florida, based on analysis of 16S rRNA and ribosomal protein gene operon sequences. Int. J. Syst. Bacteriol. **48**: 269–277.

Jomantiene, R., Davis, R.E., Valiunas, D., and Alminaite, A. (2002). New group 16SrIII phytoplasma lineages in Lithuania exhibit rRNA interoperon sequence heterogeneity. Eur. J. Plant Pathol. **108**: 507–517.

Kakizawa, S., Makino, A., Ishii, Y., Tamaki, H., and Kamagata, Y. (2014). Draft Genome Sequence of "*Candidatus* Phytoplasma asteris" Strain OY-V, an Unculturable Plant-Pathogenic Bacterium. Genome Announc. **2**: 9–10.

Kakizawa, S., Oshima, K., Nishigawa, H., Jung, H.Y., Wei, W., Suzuki, S., Tanaka, M., Miyata, S.I., Ugaki, M., and Namba, S. (2004). Secretion of immunodominant membrane protein from onion yellows phytoplasma through the Sec protein-translocation system in Escherichia coli. Microbiology **150**: 135–142.

Kazan, K. and Lyons, R. (2014). Intervention of Phytohormone Pathways by Pathogen Effectors. Plant Cell **26**: 2285–2309.

Kazan, K. and Manners, J.M. (2012). JAZ repressors and the orchestration of phytohormone crosstalk. Trends Plant Sci. **17**: 22–31.

Kim, S.H., Son, G.H., Bhattacharjee, S., Kim, H.J., Nam, J.C., Nguyen, P.D.T., Hong, J.C., and Gassmann, W. (2014). The Arabidopsis immune adaptor SRFR1 interacts with TCP transcription factors that redundantly contribute to effector-triggered immunity. Plant J. **78**: 978–89.

Kingdom, H.N. (2012). AY-WB phytoplasma manipulations of host and non--host leafhopper interactions.

Kingdom, H.N. and Hogenhout, S.A. (2007). Aster yellows phytoplasma witches' broom (AY-WB" '*Candidatus* Phytoplasma asteris') increases survival rates of *Macrosteles quadrilineatus* and *Dalbulus maidis* on various plant species. Bull. Insectology **60**: 225–226.

Koshiba, T., Ballas, N., Wong, L.M., and Theologis, A. (1995). Transcriptional regulation of *PS-IAA4/5* and *PS-IAA6* early gene expression by indoleacetic acid and protein synthesis inhibitors in pea (*Pisum sativum*). J. Mol. Biol. **253**: 396–413.

Kosugi, S. and Ohashi, Y. (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. Plant J. **30**: 337–348.

Kosugi, S. and Ohashi, Y. (1997). PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. Plant Cell **9**: 1607–19.

Ku, C., Lo, W.-S., and Kuo, C.-H. (2013). Horizontal transfer of potential mobile units in phytoplasmas. Mob. Genet. Elements **3**: e26145.

Kube, M., Schneider, B., Kuhl, H., Dandekar, T., Heitmann, K., Migdoll, A.M., Reinhardt, R., and Seemüller, E. (2008). The linear chromosome of the plant-pathogenic mycoplasma "*Candidatus* Phytoplasma mali". BMC Genomics **9**: 306.

Kumar, S., Nei, M., Dudley, J., and Tamura, K. (2008). MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform. **9**: 299–306.

Kunkel, L. 0. (1926). Studies on Aster Yellows. Am. J. Bot. **13**: 646–705.

Lee, I.-M., Davis, R.E., and Gundersen-Rindal, D.E. (2000). phytoplasma: Phytopathogenic Mollicutes. Annu. Rev. Microbiol. **54**: 221–255.

Lee, I.-M., Gundersen-Rindal, D.E., Davis, R.E., and Bartoszyk, Irena M. (1998). Revised Classification Scheme of Phytoplasmas based on RFLP Analyses of 16S rRNA and Ribosomal Protein Gene Sequences. Int. J. Syst. Bacteriol. **48**: 1153–1169.

Lee, I.-M., Hammond, R.W., Davis, R.E., and Gundersen, D.E. (1993). Universal Amplification and Analysis of Pathogen 16S rDNA for Classification and Identification of Mycoplasmalike Organisms. Mol. Plant Pathol. **83**: 834–842.

Lee, I.-M., Shao, J., Bottner-Parker, K.D., Gundersen-Rindal, D.E., Zhao, Y., and Davis, R.E. (2015). Draft Genome Sequence of "*Candidatus* Phytoplasma pruni" Strain CX, a Plant-Pathogenic Bacterium. Genome Announc. **3**: 9–10.

Lee, I.M., Bottner-Parker, K.D., Zhao, Y., Davis, R.E., and Harrison, N.A. (2010). Phylogenetic analysis and delineation of phytoplasmas based on *secY* gene sequences. Int. J. Syst. Evol. Microbiol. **60**: 2887–2897.

Lee, I.M., Gundersen-Rindal, D.E., Davis, R.E., Bottner, K.D., Marcone, C., and Seemüller, E. (2004). "*Candidatus* Phytoplasma asteris", a novel phytoplasma taxon associated with aster yellows and related diseases. Int. J. Syst. Evol. Microbiol. **54**: 1037–1048.

Lee, I.M., Zhao, Y., and Bottner, K.D. (2006). *SecY* gene sequence analysis for finer differentiation of diverse strains in the aster yellows phytoplasma group. Mol. Cell. Probes **20**: 87–91.

Lerat, E., Daubin, V., Ochman, H., and Moran, N.A. (2005). Evolutionary origins of genomic repertoires in bacteria. PLoS Biol. **3**: 0807–0814.

Li, C., Potuschak, T., Colón-Carmona, A., Gutiérrez, R.A., and Doerner, P. (2005).

Arabidopsis TCP20 links regulation of growth and cell division control pathways. Proc. Natl. Acad. Sci. U. S. A. **102**: 12978–83.

Li, L., Li, C., and Howe, G.A. (2001). Genetic analysis of wound signaling in tomato. Evidence for a dual role of jasmonic acid in defense and female fertility. Plant Physiol. **127**: 1414–1417.

Li, Z., Li, B., Shen, W.H., Huang, H., and Dong, A. (2012). TCP transcription factors interact with AS2 in the repression of class-I KNOX genes in *Arabidopsis thaliana*. Plant J. **71**: 99–107.

Liefting, L.I.A.W., Andersen, M.T., Beever, R.E., Gardner, R.C., and Forster, R.L.S. (1996). Sequence Heterogeneity in the Two 16S rRNA Genes of Phormium Yellow Leaf Phytoplasma. Microbiology **62**: 3133–3139.

Lim, P.O. and Sears, B.B. (1989). 16S rRNA sequence indicates that plant-pathogenic mycoplasmalike organisms are evolutionarily distinct from animal mycoplasmas. J. Bacteriol. **171**: 5901–5906.

Lim, P.O. and Sears, B.B. (1992). Evolutionary relationships of a plant-pathogenic mycoplasmalike organism and Acholeplasma laidlawii deduced from two ribosomal protein gene sequences. J. Bacteriol. **174**: 2606–2611.

Linnaeus, C. and Rudberg, D. (1744). Dissertatio Botanica de Peloria (Upsaliae).

Lius, S., Manshardt, R.M., Fitch, M.M.M., Slightom, J.L., Sanford, J.C., and Gonsalves, D. (1997). Pathogen-derived resistance provides papaya with effective protection against papaya ringspot virus. Mol. Breed. **3**: 161–168.

Lopez, J.A., Sun, Y., Blair, P.B., and Mukhtar, M.S. (2015). TCP three-way handshake: Linking developmental processes with plant immunity. Trends Plant Sci. **20**: 238–245.

Lu, Y.-T., Li, M.-Y., Cheng, K.-T., Tan, C.M., Su, L.-W., Lin, W.-Y., Shih, H.-T., Chiou, T.-J., and Yang, J.-Y. (2014). Transgenic plants that express the phytoplasma effector SAP11 show altered phosphate starvation and defense responses. PLANT Physiol. **164**: 1456–1469.

Luo, D., Carpenter, R., Copsey, L., Vincent, C., Clark, J., and Coen, E. (1999). Control of organ asymmetry in flowers of Antirrhinum. Cell **99**: 367–376.

Luo, D., Carpenter, R., Vincent, C., Copsey, L., and Coen, E. (1996). Origin of floral asymmetry in Antirrhinum. Nature **383**: 794–799.

MacLean, a. M., Sugio, A., Makarova, O. V., Findlay, K.C., Grieve, V.M., Toth, R., Nicolaisen, M., and Hogenhout, S. a. (2011). Phytoplasma Effector SAP54 Induces Indeterminate Leaf-Like Flower Development in Arabidopsis Plants. Plant Physiol. **157**: 831–841.

MacLean, A.M., Orlovskis, Z., Kowitwanich, K., Zdziarska, A.M., Angenent, G.C., Immink, R.G.H., and 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 Biol. **12**.

Makarova, O., MacLean, A.M., and Nicolaisen, M. (2015). Phytoplasma adapt to the diverse environments of their plant and insect hosts by altering gene expression.

Physiol. Mol. Plant Pathol. **91**: 81–87.

Marcone, C., Lee, I.M., Davis, R.E., Ragozzino, A., and Seemüller, E. (2000). Classification of aster yellows-group phytoplasmas based on combined analyses of rRNA and *tuf* gene sequences. Int. J. Syst. Evol. Microbiol. **50**: 1703–1713.

Martín-Trillo, M. and Cubas, P. (2010). TCP genes: a family snapshot ten years later. Trends Plant Sci. **15**: 31–39.

Martini, M., Lee, I.M., Bottner, K.D., Zhao, Y., Botti, S., Bertaccini, A., Harrison, N.A., Carraro, L., Marcone, C., Khan, A.J., and Osler, R. (2007). Ribosomal protein gene-based phylogeny for finer differentiation and classification of phytoplasmas. Int. J. Syst. Evol. Microbiol. **57**: 2037–2051.

Mescher, M.C. (2012). Manipulation of plant phenotypes by insects and insect-borne pathogens. Host Manip. by parasites: 73–92.

Mitrovic, J., Kakizawa, S., Duduk, B., Oshima, K., Namba, S., and Bertaccini, A. (2011). The *groEL* gene as an additional marker for finer differentiation of '*Candidatus* Phytoplasma asteris'-related strains. Ann. Appl. Biol. **159**: 41–48.

Mitrovic, J., Siewert, C., Duduk, B., Hecht, J., Mölling, K., Broecker, F., Beyerlein, P., Büttner, C., Bertaccini, A., and Kube, M. (2014). Generation and Analysis of Draft Sequences of 'Stolbur' Phytoplasma from Multiple Displacement Amplification Templates. J. Mol. Microbiol. Biotechnol. **24**: 1–11.

Mukhtar, M.S. et al. (2011). Independently Evolved Virulence Effectors Converge onto Hubs in a Plant Immune System Network. Science (80-.). **333**: 596–601.

Munkvold, K.R. and Martin, G.B. (2009). Advances in experimental methods for the elucidation of *Pseudomonas syringae* effector function with a focus on AvrPtoB. Mol. Plant Pathol. **10**: 777–93.

Musetti, R., Buxa, S. V, De Marco, F., Loschi, A., Polizzotto, R., Kogel, K.-H., and van Bel, A.J.E. (2013). Phytoplasma-triggered Ca(2+) influx is involved in sieve-tube blockage. Mol. Plant. Microbe. Interact. **26**: 379–86.

Musetti, R., Pagliari, L., Buxa, S. V., Degola, F., de Marco, F., Loschi, A., Kogel, K.H., and van Bel, A.J.E. (2016). OHMS**: Phytoplasmas dictate changes in sieve-element ultrastructure to accommodate their requirements for nutrition, multiplication and translocation. Plant Signal. Behav. **11**: 1–5.

Mushegian, A.R. and Koonin, E. V. (1996). A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc. Natl. Acad. Sci. **93**: 10268–10273.

Music, M.S., Samarzija, I., Hogenhout, S.A., Haryono, M., Cho, S.T., and Kuo, C.H. (2019). The genome of '*Candidatus* Phytoplasma solani' strain SA-1 is highly dynamic and prone to adopting foreign sequences. Syst. Appl. Microbiol. **42**: 117–127.

Nakagawa, T., Ishiguro, S., and Kimura, T. (2009). Gateway vectors for plant transformation. Plant Biotechnol. **26**: 275–284.

Namba, S., Oyaizu, H., Kato, S., Iwanami, S., and Tsuchizaki, T. (1993). Phylogenetic diversity of phytopathogenic mycoplasmalike organisms. Int. J. Syst. Bacteriol. **43**:

461–7.

Nath, U., Crawford, B.C.W., Carpenter, R., and Coen, E. (2003). Genetic Control of Surface Curvature. Science (80-.). **299**: 1404–1407.

Nault, L.R. (1980). Maize Bushy Stunt and Corn Stunt: A Comparison of Disease Symptoms, Pathogen Host Ranges, and Vectors.: 1–4.

Navaud, O., Dabos, P., Carnus, E., Tremousaygue, D., and Hervé, C. (2007). TCP Transcription Factors Predate the Emergence of Land Plants. J. Mol. Evol. **65**: 23–33.

Nguyen Ba, A.N., Pogoutse, A., Provart, N., and Moses, A.M. (2009). NLStradamus: A simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinformatics **10**: 1–11.

Nicolas, M. and Cubas, P. (2016). TCP factors: New kids on the signaling block. Curr. Opin. Plant Biol. **33**: 33–41.

Nicolas, M. and Cubas, P. (2015). The Role of TCP Transcription Factors in Shaping Flower Structure, Leaf Morphology, and Plant Architecture. In Plant Transcription Factors: Evolutionary, Structural and Functional Aspects, pp. 1–422.

Nishigawa, H., Oshima, K., Kakizawa, S., Jung, H.Y., Kuboyama, T., Miyata, S.I., Ugaki, M., and Namba, S. (2002). Evidence of intermolecular recombination between extrachromosomal DNAs in phytoplasma: A trigger for the biological diversity of phytoplasma? Microbiology **148**: 1389–1396.

Ohno, S. (1970). Evolution by gene duplication. Springer.

Olsen, G.J. and Woese, C.R. (1993). Ribosomal RNA : a key to phylogeny. FASEB J. **7**: 113–123.

Ori, N. et al. (2007). Regulation of *LANCEOLATE* by *miR319* is required for compound-leaf development in tomato. Nat. Genet. **39**: 787–791.

Orlovskis, Z., Canale, M.C., Haryono, M., Lopes, J.R.S., Kuo, C.H., and Hogenhout, S.A. (2017). A few sequence polymorphisms among isolates of Maize bushy stunt phytoplasma associate with organ proliferation symptoms of infected maize plants. Ann. Bot. **119**: 869–884.

Orlovskis, Z. and Hogenhout, S.A. (2016). A bacterial parasite effector mediates insect vector attraction in host plants independently of developmental changes. Front. Plant Sci. **7**: 036186.

Oshima, K., Ishii, Y., Kakizawa, S., Sugawara, K., Neriya, Y., Himeno, M., Minato, N., Miura, C., Shiraishi, T., Yamaji, Y., and Namba, S. (2011). Dramatic transcriptional changes in an intracellular parasite enable host switching between plant and insect. PLoS One **6**.

Oshima, K., Kakizawa, S., Nishigawa, H., Jung, H.-Y., Wei, W., Suzuki, S., Arashida, R., Nakata, D., Miyata, S., Ugaki, M., and Namba, S. (2004). Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. Nat. Genet. **36**: 27–29.

Oshima, K., Maejima, K., and Namba, S. (2013). Genomic and evolutionary aspects

of phytoplasmas. Front. Microbiol. **4**: 1–8.

Oshima, K. and Nishida, H. (2007). Phylogenetic relationships among mycoplasmas based on the whole genomic information. J. Mol. Evol. **65**: 249–258.

Oshima, K., Shiomi, T., Kuboyama, T., Sawayanagi, T., Nishigawa, H., Kakizawa, S., Miyata, S., Ugaki, M., and Namba, S. (2001). Isolation and Characterization of Derivative Lines of the Onion Yellows Phytoplasma that Do Not Cause Stunting or Phloem Hyperplasia. Phytopathology **91**: 1024–1029.

Pacifico, D., Galetto, L., Rashidi, M., Abbà, S., Palmano, S., Firrao, G., Bosco, D., and Marzachì, C. (2015). Decreasing Global Transcript Levels over Time Suggest that Phytoplasma Cells Enter Stationary Phase during Plant and Insect Colonization. Appl. Environ. Microbiol. **81**: 2591–2602.

Pagliari, L., Martini, M., Loschi, A., and Musetti, R. (2016). Looking inside phytoplasma-infected sieve elements: A combined microscopy approach using *Arabidopsis thaliana* as a model plant. Micron **89**: 87–97.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. Nature **425**: 257–263.

Paterson, A.H., Bowers, J.E., Burow, M.D., Draye, X., Elsik, C.G., Jiang, C.-X., Katsar, C.S., Lan, T.-H., Lin, Y.-R., Ming, R., and Wright, R.J. (2000). Comparative Genomics of Plant Chromosomes. Plant Cell **12**: 1523.

Pecher, P., Moro, G., Canale, M.C., Capdevielle, S., Singh, A., MacLean, A., Sugio, A., Kuo, C.-H., Lopes, J.R.S., and Hogenhout, S.A. (2019). Phytoplasma SAP11 effector destabilization of TCP transcription factors differentially impact development and defence of Arabidopsis versus maize. PLoS Pathog. **1**.(In press)

Pérez-López, E., Olivier, C.Y., Luna-Rodríguez, M., Rodríguez, Y., Iglesias, L.G., Castro-Luna, A., Adame-García, J., and Dumonceaux, T.J. (2016). Maize bushy stunt phytoplasma affects native corn at high elevations in Southeast Mexico. Eur. J. Plant Pathol.: 1–9.

Pieterse, C.M.J., Van der Does, D., Zamioudis, C., Leon-Reyes, A., and Van Wees, S.C.M. (2012). Hormonal Modulation of Plant Immunity. Annu. Rev. Cell Dev. Biol. **28**: 489–521.

Pruneda-Paz, J.L., Breton, G., Nagel, D.H., Kang, S.E., Bonaldi, K., Doherty, C.J., Ravelo, S., Galli, M., Ecker, J.R., and Kay, S.A. (2014). A Genome-Scale Resource for the Functional Characterization of Arabidopsis Transcription Factors. Cell Rep. **8**: 622–632.

Purcell, A.H. (1988). Increased survival of *Dalbulus maidis*, a specialist on maize, on non-host plants infected with mollicute plant pathogens. Entomol. Exp. Appl. **46**: 187–196.

Quaglino, F., Kube, M., Jawhari, M., Abou-jawdah, Y., Siewert, C., Choueiri, E., Sobh, H., Casati, P., Tedeschi, R., Lova, M.M., Alma, A., and Bianco, P.A. (2015). '*Candidatus* Phytoplasma phoenicium' associated with almond witches '-broom

disease : from draft genome to genetic diversity among strain populations. BMC Microbiol. **15:148**: 1–15.

Raffaele, S. et al. (2010). Genome Evolution Following Host Jumps in the Irish Potato Famine Pathogen Lineage. Science (80-.). **330**: 1540–1543.

Randall, R.N., Radford, C.E., Roof, K.A., Natarajan, D.K., and Gaucher, E.A. (2016). An experimental phylogeny to benchmark ancestral sequence reconstruction. Nat. Commun. **7**: 1–6.

Razin, S., Yogev, D., and Naot, Y. (1998). Molecular biology and pathogenicity of mycoplasmas. S. Razin and R. Herrmann, eds (Kluwer Academic Publishers).

Remigi, P., Anisimova, M., Guidot, A., Genin, S., and Peeters, N. (2011). Functional diversification of the GALA type III effector family contributes to *Ralstonia solanacearum* adaptation on different plant hosts. New Phytol. **192**: 976–987.

Resentini, F., Felipo-Benavent, A., Colombo, L., Blázquez, M.A., Alabadí, D., and Masiero, S. (2015). TCP14 and TCP15 mediate the promotion of seed germination by gibberellins in *Arabidopsis thaliana*. Mol. Plant **8**: 482–485.

Reymond, P. and Farmer, E.E. (1998). Jasmonate and salicylate as global signals for defense gene expression. Curr. Opin. Plant Biol. **1**: 404–411.

Robicsek, A., Jacoby, G.A., and Hooper, D.C. (2006). The worldwide emergence of plasmid-mediated quinolone resistance. Lancet Infect. Dis. **6**: 629–640.

Roy, B.A. (2001). Patterns of association between crucifers and their flower-mimic pathogens: Host jumps are more common than coevolution or cospeciation. Evolution (N. Y). **55**: 41–53.

Saccardo, F., Martini, M., Palmano, S., Ermacora, P., Scortichini, M., Loi, N., and Firrao, G. (2012). Genome drafts of four phytoplasma strains of the ribosomal group 16SrIII. Microbiol. (United Kingdom) **158**: 2805–2814.

Dos Santos Maraschin, F., Memelink, J., and Offringa, R. (2009). Auxin-induced, SCFTIR1-mediated poly-ubiquitination marks AUX/IAA proteins for degradation. Plant J. **59**: 100–109.

Sarvepalli, K. and Nath, U. (2011). Hyper-activation of the TCP4 transcription factor in *Arabidopsis thaliana* accelerates multiple aspects of plant maturation. Plant J. **67**: 595–607.

Schneider, B. and Seemüller, E. (1994). Presence of two sets of ribosomal genes in phytopathogenic mollicutes. Appl. Environ. Microbiol. **60**: 3409–3412.

Schommer, C., Debernardi, J.M., Bresso, E.G., Rodriguez, R.E., and Palatnik, J.F. (2014). Repression of cell proliferation by miR319-regulated TCP4. Mol. Plant **7**: 1533–1544.

Schommer, C., Palatnik, J.F., Aggarwal, P., Chételat, A., Cubas, P., Farmer, E.E., Nath, U., and Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. PLoS Biol. **6**: e230.

Seemüller, E., Garnier, M., and Schneider, B. (2002). Mycoplasmas of plants and

insects. In Molecular Biology and Pathology of Mycoplasmas, pp. 91–116.

Seemüller, E. and Schneider, B. (2004). "*Candidatus* Phytoplasma mali", "*Candidatus* Phytoplasma pyri" and *Candidatus* Phytoplasma prunorum', the casual agents of apple proliferation, pear decline and European stone fruit yellows, respectively. Int. J. Syst. Evol. Microbiol. **54**: 1217–1226.

Simillion, C., Vandepoele, K., Van Montagu, M.C.E., Zabeau, M., and Van de Peer, Y. (2002). The hidden duplication past of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. **99**: 13627–13632.

Sparks, M.E., Bottner-parker, K.D., Gundersen-rindal, D.E., and Lee, I. (2018). Draft genome sequence of the New Jersey aster yellows strain of ' *Candidatus* Phytoplasma asteris .' PLoS One: 1–16.

Spears, B.J., Howton, T.C., Gao, F., Garner, C.M., Mukhtar, M.S., and Gassmann, W. (2019). Direct Regulation of the *EFR* -Dependent Immune Response by Arabidopsis TCP Transcription Factors. Mol. Plant-Microbe Interact. **32**: 540–549.

Studer, A.J., Wang, H., and Doebley, J.F. (2017). Selection during maize domestication targeted a gene network controlling plant and inflorescence architecture. Genetics **207**: 755–765.

Sugawara, K., Honma, Y., Komatsu, K., Himeno, M., Oshima, K., and Namba, S. (2013). The alteration of plant morphology by small peptides released from the proteolytic processing of the bacterial peptide TENGU. Plant Physiol. **162**: 2005–2014.

Sugio, A. and Hogenhout, S.A. (2012). The genome biology of phytoplasma: modulators of plants and insects. Curr. Opin. Microbiol. **15**: 247–254.

Sugio, A., Kingdom, H.N., MacLean, A.M., Grieve, V.M., and Hogenhout, S. a (2011a). Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. Proc. Natl. Acad. Sci. U. S. A. **108**: E1254-63.

Sugio, A., Maclean, A.M., and Hogenhout, S. a. (2014). The small phytoplasma virulence effector SAP11 contains distinct domains required for nuclear targeting and CIN-TCP binding and destabilization. New Phytol. **202**: 838–848.

Sugio, A., MacLean, A.M., Kingdom, H.N., Grieve, V.M., Manimekalai, R., and Hogenhout, S. a. (2011b). Diverse Targets of Phytoplasma Effectors: From Plant Development to Defense Against Insects. Annu. Rev. Phytopathol. **49**: 175–195.

Suzuki, S., Oshima, K., Kakizawa, S., Arashida, R., Jung, H.-Y., Yamaji, Y., Nishigawa, H., Ugaki, M., and Namba, S. (2006). Interaction between the membrane protein of a pathogen and insect microfilament complex determines insect-vector specificity. Proc. Natl. Acad. Sci. U. S. A. **103**: 4252–4257.

Tanaka, S., Schweizer, G., Rössel, N., Fukada, F., Thines, M., and Kahmann, R. (2019). Neofunctionalization of the secreted Tin2 effector in the fungal pathogen *Ustilago maydis*. Nat. Microbiol. **4**: 251–257.

Tatematsu, K., Nakabayashi, K., Kamiya, Y., and Nambara, E. (2008). Transcription

factor AtTCP14 regulates embryonic growth potential during seed germination in *Arabidopsis thaliana*. Plant J. **53**: 42–52.

Thornton, J.W. (2004). Resurrecting ancient genes: experimental analysis of extinct molecules. Nat. Rev. Genet. **5**: 366–375.

Toruño, T.Y., Seruga Musić, M., Simi, S., Nicolaisen, M., and Hogenhout, S.A. (2010). Phytoplasma PMU1 exists as linear chromosomal and circular extrachromosomal elements and has enhanced expression in insect vectors compared with plant hosts. Mol. Microbiol. **77**: 1406–1415.

Tran-Nguyen, L.T.T., Kube, M., Schneider, B., Reinhardt, R., and Gibb, K.S. (2008). Comparative Genome Analysis of "*Candidatus* Phytoplasma australiense" (Subgroup tuf-Australia I; rp-A) and "*Ca*. Phytoplasma asteris" Strains OY-M and AY-WB. J. Bacteriol. **190**: 3979–3991.

Treangen, T.J. and Rocha, E.P.C. (2011). Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. PLoS Genet. **7**.

Turner, J.G., Ellis, C., and Devoto, A. (2002). The jasmonate signal pathway. Plant Cell **14 Suppl**: S153–S164.

Uberti Manassero, N.G., Viola, I.L., Welchen, E., and Gonzalez, D.H. (2013). TCP transcription factors: Architectures of plant form. Biomol. Concepts **4**: 111–127.

Ulmasov, T., Murfett, J., Hagen, G., and Guilfoyle, T.J. (1997). Aux/lAA Proteins Repress Expression of Reporter Genes Containing Natural and Highly Active Synthetic Auxin Response Elements. Society **9**: 1963–1971.

Valsecchi, I., Guittard-Crilat, E., Maldiney, R., Habricot, Y., Lignon, S., Lebrun, R., Miginiac, E., Ruelland, E., Jeannette, E., and Lebreton, S. (2013). The intrinsically disordered C-terminal region of *Arabidopsis thaliana* TCP8 transcription factor acts both as a transactivation and self-assembly domain. Mol. Biosyst. **9**: 2282–2295.

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G.F., Chater, K.F., and Sinderen, D. van (2007). Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum. Microbiol. Mol. Biol. Rev. **71**: 495–548.

Viola, I.L., Reinheimers, R., Ripoll, R., Uberti Manassero, N.G., and Gonzalez, D.H. (2012). Determinants of the DNA binding specificity of class I and class II TCP transcription factors. J. Biol. Chem. **287**: 347–356.

Viola, I.L., Uberti Manassero, N.G., Ripoll, R., and Gonzalez, D.H. (2011). The *Arabidopsis* class I TCP transcription factor AtTCP11 is a developmental regulator with distinct DNA-binding properties due to the presence of a threonine residue at position 15 of the TCP domain. Biochem. J. **435**: 143–155.

Wang, G. et al. (2015a). The Decoy Substrate of a Pathogen Effector and a Pseudokinase Specify Pathogen-Induced Modified-Self Recognition and Immunity in Plants. Cell Host Microbe **18**: 285–295.

Wang, J., Song, L., Jiao, Q., Yang, S., Rui Gao, X.L., and Zhou, G. (2018a). Comparative genome analysis of jujube witches'-broom Phytoplasma, an obligate pathogen that causes jujube witches'-broom disease. BMC Genomics **19**: 1–12.

Wang, N., Yang, H., Yin, Z., Liu, W., Sun, L., and Wu, Y. (2018b). Phytoplasma effector SWP1 induces witches' broom symptom by destabilizing the TCP transcription factor BRANCHED1. Mol. Plant Pathol. **19**: 2623–2634.

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci. **9**: 244–252.

Wang, X., Gao, J., Zhu, Z., Dong, X., Wang, X., Ren, G., Zhou, X., and Kuai, B. (2015b). TCP transcription factors are critical for the coordinated regulation of *ISOCHORISMATE SYNTHASE 1* expression in *Arabidopsis thaliana*. Plant J. **82**: 151–162.

Webb, D.R., Bonfiglioli, R.G., Carraro, L., Osler, R., and Symons, R.H. (1999). Oligonucleotides as hybridization probes to localize phytoplasmas in host plants and insect vectors. Phytopathology **89**: 894–901.

Weintraub, P.G. and Beanland, L. (2006). Insect Vectors of Phytoplasmas. Annu. Rev. Entomol. **51**: 91–111.

Weisburg, W.G., Tully, J.G., Rose, D.L., Petzel, J.P., Oyaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T.G., and Van Etten, J. (1989). A phylogenetic analysis of the mycoplasmas: basis for their classification. J. Bacteriol. **171**: 6455–6467.

Weßling, R. et al. (2014). Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. Cell Host Microbe **16**: 364–375.

Woese, C.R. (1987). Bacterial evolution. Microbiol. Rev. **51**: 221–271.

Worley, C.K., Zenser, N., Ramos, J., Rouse, D., Leyser, O., Theologis, A., and Callis, J. (2000). Degradation of Aux / IAA proteins is essential for normal auxin signalling. Plant J. **21**: 553–562.

Wu, Y., Wood, M.D., Tao, Y., and Katagiri, F. (2003). Direct delivery of bacterial avirulence proteins into resistant Arabidopsis protoplasts leads to hypersensitive cell death. Plant J. **33**: 131–137.

Xin, X.-F. and He, S.Y. (2013). *Pseudomonas syringae* pv. tomato DC3000: A Model Pathogen for Probing Disease Susceptibility and Hormone Signaling in Plants . Annu. Rev. Phytopathol. **51**: 473–498.

Yang, L., Teixeira, P.J.P.L., Biswas, S., Finkel, O.M., He, Y., Salas-Gonzalez, I., English, M.E., Epple, P., Mieczkowski, P., and Dangl, J.L. (2017). *Pseudomonas syringae* Type III Effector HopBB1 Promotes Host Transcriptional Repressor Degradation to Regulate Phytohormone Responses and Virulence. Cell Host Microbe **21**: 156–168.

Yilmaz, A., Nishiyama, M.Y., Fuentes, B.G., Souza, G.M., Janies, D., Gray, J., and Grotewold, E. (2009). GRASSIUS: A platform for comparative regulatory genomics across the grasses. Plant Physiol. **149**: 171–180.

Yoo, S.-D., Cho, Y.-H., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat. Protoc. **2**: 1565–1572.

Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A

Molecular Timeline for the Origin of Photosynthetic Eukaryotes. Mol. Biol. Evol. **21**: 809–818.

Zamorano, A. and Fiore, N. (2016). Draft Genome Sequence of 16SrIII-J Phytoplasma , a Plant Pathogenic Bacterium with a Broad Spectrum of Hosts. Genome Announc. **4**: 6–7.

Zess, E.K., Białas, A., and Kamoun, S. (2019). Old fungus, new trick. Nat. Microbiol. **4**: 210–211.

Zhang, J., Hogenhout, S. a, Nault, L.R., Hoy, C.W., and Miller, S. a (2004). Molecular and symptom analyses of phytoplasma strains from lettuce reveal a diverse population. Phytopathology **94**: 842–849.

Zhu, Y., He, Y., Zheng, Z., Chen, J., Wang, Z., and Zhou, G. (2017). Draft Genome Sequence of Rice Orange Leaf Phytoplasma from Guangdong, China. Genome Announc. **5**: 9–10.

Zimmermann, M.R., Schneider, B., Mithöfer, A., Reichelt, M., Seemüller, E., and Furch, A.C.. (2015). Implications of *Candidatus* Phytoplasma mali infection on phloem function of apple trees. J. Endocytobiosis Cell Res. **26**: 67–75.

Zreik, L., Carle, P., Bové, J.M., and Garnier, M. (1995). Characterization of the mycoplasmalike organism associated with witches'-broom disease of lime and proposition of a *Candidatus* taxon for the organism, "*Candidatus* phytoplasma aurantifolia". Int. J. Syst. Bacteriol. **45**: 449–453.