UNIVERSITY OF CALIFORNIA
RIVERSIDE

Tomato-Potato Aphid Interactions: Insights into Plant Defense and the Aphid Pest

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Genetics, Genomics and Bioinformatics

by

Hagop Sarkis Atamian

December 2012

Dissertation Committee:
Dr. Isgouhi Kaloshian, Chairperson
Dr. Thomas Eulgem
Dr. Linda Walling

This Dissertation of Hagop Sarkis Atamian is approved:

_____

_____

_____
Committee Chairperson

University of California, Riverside

Acknowledgements

I would like to thank my PI and Dissertation Chair, Dr. Isgouhi Kaloshian, for her constant mentorship and valuable advices, for her generous financial support, for giving me the opportunity to work on several diverse projects that provided valuable experience, and for openhandedly giving me authorship in several publications in which I participated. I also want to thank my dissertation committee members Dr. Thomas Eulgem and Dr. Linda Walling for their kind guidance and advices, which always provided me encouragement and replenished my enthusiasm for research. The completion of my dissertation project would not have been possible without the input of the Kaloshian lab members. You created nice and warm working environment. I also would like to thank Dr. Thomas Girke for his generous help with bioinformatics analysis. I would also like to thank the administrative and technical staff of both Department of Nematology and Genetics, Genomics and Bioinformatics program.

Most importantly, I would like to thank my family and wife. I am grateful to you for all of your patience and support.

DEDICATIONS

To My Family and Wife

ABSTRACT OF THE DISSERTATION

Tomato-Potato Aphid Interactions: Insights into Plant Defense and the Aphid Pest

by

Hagop Sarkis Atamian

Doctor of Philosophy, Graduate Program in Genetics, Genomics and Bioinformatics
University of California, Riverside, December 2012
Dr. Isgouhi Kaloshian, Chairperson

Aphids cause extensive economic losses to cultivated crops worldwide. Their success as pests is in part due to their complex life cycle, wide host range, and the ability of a female aphid to contain not only the developing embryos of her daughters, but also those of her grand-daughters which develop within her daughters. The latter results in build up of immense populations very quickly. Resistant plants represent an environmental friendly approach to combat aphid pests. Better understanding of plant-aphid interaction will contribute to engineering durable plant resistance. In tomato (*Solanum lycopersicum*), the *Mi-1* gene confers resistance to potato aphid (*Macrosiphum euphorbiae*), root-knot nematode (RKN) (*Meloidogyne* sp.), whitefly (*Bemisia tabaci*), and tomato psyllid (*Bactericera cockerelli*). This incompatible tomato interaction with RKN is characterized by hypersensitive response and vast transcriptional reprogramming, including differential regulation of transcription

factors (TFs). Using gene knock-down approach in Chapter one, a role for *SlWRKY70* TF was identified in *Mi-1*-mediated tomato resistance against potato aphid and RKN. Gene expression analysis showed that the regulation of this TF by Salicylic acid and Jasmonic acid hormones is conserved between tomato and *Arabidopsis thaliana*. The study of *SlWRKY70* revealed that there is no consistent nomenclature for plant WRKY TF family. For this reason a phylogenetic analysis was conducted using sequences from 15 plant species. Chapter two presents the analysis and the established orthologous relationship of WRKY TFs among these plant species. Consequently, this analysis allowed the design of a systemic nomenclature for the WRKY TF family to include the inferred orthology relationships. Chapters three and four pursued another approach to understand plant-aphid interactions. These chapters focused on identifying the aphid effectors and putative lineage-specific set of genes through sequencing the aphid and its salivary gland transcriptomes. In Chapter three sequencing and annotation of the potato aphid transcriptome enabled us to conduct comparative sequence analysis with three other aphid species, as well as seven additional species of insects from different clades and a planktonic crustacean. This analysis identified a set of aphid-specific genes, which may contribute to aphid's unconventional biology. The transcripts of a subset of these aphid-specific genes were expressed in the salivary glands suggesting that they are involved in aphid-host interactions. To study this interaction in more detail, the potato aphid salivary gland transcriptome was sequenced in Chapter four. This enabled identification of secreted proteins based on prediction of secretion signal peptides. *In planta* functional characterization of eight of

these putative aphid secreted proteins identified roles for two, Me10 and Me23, in altering tomato responses to the aphid's advantage.

TABLE OF CONTENTS

List of Tables

Table S11 Sequences of the *Macrosiphum euphorbiae* salivary gland contigs less than 200 bp in length.

Table S12 Annotation of *Macrosiphum euphorbiae* salivary gland contigs.

List of Figures

General Introduction

Green plants, especially flowering plants, are more than just landscaping for the planet since they supply humanity with all the essentials of life: food and oxygen, as well as products that have shaped modern society (Levetin and McMahon 2012). Unable to move in response to changing conditions, plants are subjected to various biotic and abiotic stresses throughout their sedentary life cycle. These continuous stressful conditions have prompted plants to develop a range of responses to be able to cope with such adverse conditions. Plant defense against biotic stresses include physical barriers (Aist 1979; Hematy et al. 2009; Lai et al. 2000), chemical weapons (Bednarek and Osbourn 2009), and immune responses (Hammond-Kosak and Jones 1996).

Two principal immune responses operate against biotic stresses in plants. The first line of active defense is triggered by a class of immune receptors upon recognition of microbial-associated molecular patterns (MAMPs), signature motifs that are widely conserved among certain pathogen clades (Jones and Dangl 2006). This defense response is referred to as pattern-triggered immunity (PTI). As part of the continuous arms race between plants and pathogens, the later have evolved to acquire effector molecules to counteract the host PTI resulting in compatible interaction. This has prompted plants in turn to evolve R proteins that recognize directly or indirectly the pathogen effector(s) and initiate the second principle immune response termed effector-triggered immunity (ETI). This interaction is referred to as incompatible interaction. ETI is fast, effective and race-specific.

The conversion of the signal(s) triggered by the plant R proteins after pathogen recognition to actual resistant phenotype is the result of collective contribution of a multitude of reactions, studied extensively in the model plant *Arabidopsis thaliana* (Arabidopsis). These include rapid transcriptional reprogramming, production of pathogenesis-related (PR) proteins, reactive oxygen and nitrogen species (ROS and RNS), phytoalexins, antibiotic compounds, and hypersensitive response (HR) accompanied with physiological changes involving cell wall reinforcement around the infection site, lignification, deposition of callus, and differential regulation of photosynthesis and respiration (Ahuja et al. 2012; Berger et al. 2007; Coll et al. 2011; Dangl and Jones 2001; Torres et al. 2006).

Starting early nineties, extensive research led to the cloning of a large number of plant *R*-genes that were assigned to different classes based on encoding similar putative structural motifs such as leucine-rich repeat (LRR) shared among great majority of R proteins (Eitas and Dangl 2010; Liu et al. 2007). Interestingly, these R proteins with low structural diversity were shown to confer resistance to diverse pathogens and pests (Martin et al. 2003). Understanding the mode of action of R proteins has been of great interest for many years. Currently it is well known that R proteins detect the presence of effectors directly or indirectly leading to rapid activation of defense signaling pathways (van der Hoorn and Kamoun 2008).

Tomato (*Solanum lycopersicum* L., formerly *Lycopersicon esculentum* Miller) is the second most important vegetable crop in the world next to potato. About 145 million tons of tomatoes were produced in 2010 (FAOSTAT 2010). Due to the high

nutritional value of its fruit, high yield, short life cycle, and diverse varieties, tomato is widely grown year round under both outdoor and indoor conditions. Consequently, this widespread cultivation makes tomato plants vulnerable to diverse array of pathogenic agents and pests including fungi, bacteria, oomycetes, viruses, insects and nematodes (Blancard 2012).

Aphids (order Hemiptera, family Aphididae) are among the most destructive insect pests on cultivated plants. These soft-bodied insects use their piercing-sucking mouthpart to feed on plant sap. These phloem feeders damage the plants directly by sucking the plant sap and indirectly by vectoring plant viruses and excreting sugary honeydew on foliage, stems and fruits, which supports growth of the black sooty mold fungus (Harrington and Van Emden 2007).

Root-knot nematodes (RKNs; *Meloidogyne* spp.) are one of the three most economically damaging genera of plant-parasitic nematodes on horticultural and field crops. These plant parasites infect the root system of more than 2000 plant species worldwide. RKN feeding induces the formation of giant cells that drain the plant's nutrients and cause structural changes in the vascular element resulting in inefficient absorption of water and nutrients (Williamson and Gleason 2003). RKN infection is manifested by the formation of root galls and hence the common name of this group of nematodes: root-knot nematodes.

Tomato Resistance Genes

The use of resistant crop cultivars is an important component of a sustainable disease management strategy in modern agriculture. It is an environmentally benign method that can be used as an alternative to chemical pesticides, the applicability of which is becoming limited due to their adverse environmental effects and the emergence of resistant pathogens and pests strains (Walters 2011). The cultivated tomato, *S. lycopersicum,* has a narrow genetic base and is consequently vulnerable to many diseases and pests (Bai and Lindhout 2007; Sim et al. 2009). On the other hand, a repertoire of genetically diverse wild tomato species presents a rich source of *R*-genes (Bai and Lindhout 2007). Over the past 50 years, several race-specific *R*-genes have been identified in wild tomato species and extensive tomato breeding programs have been based on the transfer of these *R*-genes from accessions of wild origin into the cultivated tomato.

The cloning of *R*-genes was essential for studying the mechanisms of interaction between plants and pathogens/pests at the molecular and biochemical levels. An array of early recognition events in *R*-gene-mediated resistance has been documented depending on the particular *R*-gene and plant pathogen/pest combination (Martin et al. 2003; Soosaar et al. 2005).

The fungus *Cladosporium fulvum* (*Cf*)–tomato interaction is a well-established model system that complies with the gene-for-gene concept first described by Flor (Flor 1971). Elegant experiments demonstrated the involvement of pathogen effectors or avirulence factors (Avrs) in the induction of ETI post recognition by the tomato

resistant genes (*Cf*) against *Cladosporium fulvum*, resulting in incompatible interaction (Dixon et al. 1998; Dixon et al. 1996; Jones et al. 1994; Thomas et al. 1997). *Cf*-mediated resistance involves formation of cell wall appositions, callose deposition and phytoalexin accumulation together with culmination of the most typical defense response the hypersensitive response (HR). HR is a form of programmed cell death that results in localized necrotic tissue presumed to limit further growth of the fungal pathogen (Vossen et al. 2010). Thus *Cf*-mediated resistance phenotype is the combined result of HR and other defense responses. In contrast, the *I*-gene-mediated resistance against *Fusarium oxysporum* formae speciales *lycopersici* (Fol), a xylem-colonizing fungus, lacks the HR response and involves callose deposition, accumulation of phenolics and formation of tyloses (outgrowths of xylem contact cells) and gels in the infected vessels (Beckman 2000).

Differences in the incompatible responses to nematodes are also evident in tomato. The *Mi-1.2*-mediated resistance against the three RKN species (*M. arenaria, M. incognita and M. javanica*) involves HR, which results in the invading juvenile nematode not to be able to induce a visible feeding site (Bhattarai et al. 2008; Paulson and Webster 1972). Such an early and rapid cell death response around the nematode feeding site is known to be a common response observed during RKN infection in a number of host plants carrying resistance genes as is with *Meloidogyne exigua 1* (*Mex-1*)-mediated resistance in coffee (Anthony et al. 2005) and *Me3*-mediated resistance in pepper (Pegard et al. 2005). On the other hand, the *Hero*-mediated resistance against potato cyst nematodes (PCNs; *Globodera* spp.) is often described as a “hypersensitive-

like" or "delayed hypersensitive" response that appears after syncytium or the feeding structure induction and leads to slow deterioration or abnormal development of the feeding site (Holtmann et al. 2000). Consequently, PCNs and other cyst-forming nematodes usually are able to invade and develop on resistant plants but their reproduction is severely compromised (Sobczak et al. 2005). In some systems, non HR-mediated nematode resistance has been observed as in *Hsp1*-mediated resistance against the cyst nematode *Heterodera schachtii* in sugar beet (Holtmann et al. 2000), and *Rk*-mediated resistance against RKN in cowpea (Das et al. 2008). In the latter, the mechanism of resistance was found to be due to giant cell deterioration and arrested female nematode development leading to inability to reach maturity and initiate egg laying.

A set of distinct resistance mechanisms in tomato operate against plant viruses as well. The resistance to *Tomato mosaic virus* (ToMV) is due to the action of the resistant protein Tm-1, that binds to the replication proteins of ToMV and inhibits their function at a step before the viral replication complex is formed on the membrane surfaces (Ishibashi et al. 2007; Ishibashi et al. 2009).

Similarly, distinct resistance mechanisms associated with the *Ol*-genes against the powdery mildew fungus *Oidium neolycopersici* have been demonstrated using near-isogenic tomato lines (Bai et al. 2005). The dominant resistance genes (*Ol-1*, *Ol-3, Ol-4, Ol-5*, and *Ol-6*) hamper the fungal growth via classical HR of the host epidermal cells,

while the recessive gene *ol-2* confers resistance via papilla formation (Bai et al. 2005; Li et al. 2007).

Mi-1.2: Single Gene With Multiple Modes of Action

The tomato *Mi-1.2 R*-gene, originally identified in the wild tomato species *S. peruvianum,* was introgressed into cultivated tomato using embryo rescue (Smith 1944). Besides conferring resistance to three species of RKN, *Mi-1.2* confers resistance to potato aphid (*Macrosiphum euphorbiae*), whitefly (*Bemisia tabaci*) and tomato psyllid (*Bactericera cockerelli*) (Casteel et al. 2006; Dropkin 1969; Nombela et al. 2003; Roberts and Thomason 1986; Rossi et al. 1998). *Mi-1.2* encodes a coiled-coil (CC)-nucleotide-binding (NB)-leucine rich repeat (LRR) protein (Milligan et al. 1998). The NB domain is able to bind ATP and exert ATPase activity (Tameling et al. 2002).

The *Mi-1.2* gene, with its ability to recognize taxonomically divergent organisms, represents an interesting model among plant *R*-genes that typically confer resistance to a single species of a pathogen. Another *R*-gene with dual resistance includes the *Cf-2* gene conferring resistance against *Cladosporium fulvum* and *Globodera rostochiensis* (Lozano-Torres et al. 2012). Although the precise resistance mechanism(s) mediated by *Mi-1.2* remains unclear, available information suggests different resistance mechanisms operating against the different types of pests. The *Mi-1.2*-mediated resistance against RKN is active during all life stages of the tomato plant (Kaloshian et al. 1995). Its effect against aphid and whitefly is developmentally

regulated where tomato plants carrying the *Mi-1.2* gene show the resistant phenotype only after four true-leaf stage (Kaloshian et al. 1997; Nombela et al. 2003). This age-dependent development of resistance against aphids is regulated by mechanisms other than the transcriptional regulation of the *Mi-1.2* gene itself (Goggin et al. 2006; Martinez de Ilarduya and Kaloshian 2001). The possible regulation of the Mi-1.2 protein (in young vs old leaf) at the post-translational level or stability of the protein and/or the possible lack of factor(s) in young leaves other than *Mi-1.2* necessary for resistance needs further investigation.

The presence and absence of HR during *Mi-1.2*-mediated RKN and aphid/whitefly resistance, respectively, constitutes yet another difference in the mechanism of resistance. Although HR is typically associated with resistance in plants against pathogens, little is known about HR as a resistance mechanism against piercing-sucking insects, such as aphids and whiteflies. A strong, early HR response accompanied with significant oxidative burst has been observed in *Mi-1.2*-mediated resistance against RKN in tomato (Dropkin 1969; Mellilo et al. 2006). However, no HR was detected after aphid feeding on resistant tomato leaves (Martinez de Ilarduya et al. 2003). However oxidative burst was observed in tomato leaves 24 h after aphid infestation in both compatible and incompatible interactions indicating that ROS may contribute to basal defense against aphids and is not specific to *Mi-1.2*-mediated resistance.

In addition to the differences in HR production in the *Mi-1.2*-mediated resistance to RKN and aphids, variable mode of resistance has been reported for

potato aphids and whiteflies. The resistance against aphids is detected after stylet penetration of epidermis and mesophyl (epidermis/mesophyl level resistance) tissues as well as after stylet contacts the sieve element cells and initiate feeding (sieve element level resistance) resulting in inhibition of stylet penetration and shorter feeding on sieve element (Kaloshian et al. 2000; Pallipparambil et al. 2010). While the resistance to whitefly is associated with only inhibition of the stylet penetration (epidermis/mesophyl level resistance) since after reaching the sieve element whiteflies are able to feed continuously on the phloem sap (Nombela et al. 2003). These diverse mechanisms of *Mi-1.2*-mediated resistance suggest the presence of other factors necessary for the resistance phenotype that are different for each of the three organisms apparently resulting from different modes of action. A clear demonstration of this concept is seen following heterologous expression of the *Mi-1.2* gene in eggplant which has been shown to confer resistance against RKN but not potato aphids suggesting the requirements for *Mi-1.2*-mediated aphid and nematode resistance differ and that the additional factor(s) required for aphid resistance is not conserved between tomato and eggplant (Goggin et al. 2006).

Tomato *R*-Gene-Mediated Transcriptome Responses to Biotic Stresses

The *R*-gene-mediated defense responses or ETI are generally characterized by a vast transcriptional reprogramming after recognition of the pathogen/pest effector molecule(s) (Caplan et al. 2008; Dodds and Rathjen 2010; Eulgem 2005; Tsuda et al. 2009). High-throughput transcriptome analysis constitutes the first step in elucidating

the pathways operating during a given incompatible interaction through the identification of the differentially regulated genes and ultimately correlating expression to function. Functional characterization of genes requires the use of mutant lines or developing gene knock-down or knock-out mutants to assess their roles.

Recent tomato transcriptome profiling studies identified both similar and different global transcriptional responses and defense strategies against different biotic stresses initiated by distinct *R*-genes. Unlike the lack of dramatic changes in gene expression observed during tomato *OI*-gene-mediated response against the powdery mildew pathogen *O. neolycopersici* (Li et al. 2006a), massive cell reprogramming is evident during tomato incompatible defense responses against RKN, *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) race T3, *C. fulvum*, *Pseudomonas syringae* pv. *tomato* (*Pst)* and to some extend against *Verticillium dahliae* (Balaji et al. 2007; Bhattarai et al. 2010; Mysore et al. 2002; van Esse et al. 2009). Several families of plant transcription factors (TFs) have been extensively implicated in plant defense responses acting as both negative and positive regulators of defense (Moore et al. 2011). WRKY family of TFs are among the TFs implicated in plant defense and have been shown to be differentially regulated during *Mi-1.2*- and *Cf-9*-mediated tomato resistance to RKN and *C. fulvum*, but interestingly not *Ve*-mediated tomato responses to *V. dahliae* (Bhattarai et al. 2008; van Esse et al. 2009).

Ion fluxes through the plasma membrane is one of the earliest plant cellular responses mediated by *R*-genes, which also have a role in regulating myriad cellular responses in eukaryotes. An influx of calcium and hydrogen ions into the cell is

essential for the formation of HR and local lesions, a central feature of *R*-gene-mediated resistance that restricts the growth and spread of a pathogen/pest (Morel and Dangl 1997). Calmodulin, a highly conserved and well-characterized Ca^{+2} sensor involved in calcium- signaling (Bouche et al. 2005), was differentially upregulated during *Mi-1.2* defense against RKN (Bhattarai et al. 2008). The formation of HR is a complex process involving coordination with plant primary metabolism. In tobacco, it has been shown that shutdown of photosynthesis and increase in respiration precedes production of HR (Scharte et al. 2005). Consistent with this observation, downregulation of photosynthesis-related genes and upregulation of respiration related genes have been observed in tomato responses to *Pst*, *C. fulvum*, *Xcv* race T3 (Balaji et al. 2007; Mysore et al. 2002; van Esse et al. 2009), all involving HR, but not *V. dahliai* that does not induce HR (van Esse et al. 2009).

Other players of plant defense responses include the mitogen-activated protein kinase (MAPK) signaling pathway (Pitzschke et al. 2009; Rodriguez et al. 2010). This signaling cascade appears to be a highly conserved defense response against different biotic stresses, transferring extracellularly generated signals into the cell and subsequent activation of plant defense mechanisms (Rivas and Thomas 2005). Components of the MAPK signaling pathway has been shown to be differentially regulated in tomato after *Xcv* race T3, RKN and *Pst* infection (Balaji et al. 2007; Bhattarai et al. 2008; Mysore et al. 2002).

The plant response during incompatible interaction against biotic stresses also involves the cell's *de novo* protein synthesis machinery and/or the ubiquitination

pathways (Trujillo and Shirasu 2010). This is partly to keep up with the induction of the massive cell reprogramming and to act as another layer of regulation of the different signaling pathways. The ubiquitination pathway also contributes to the turnover of R proteins and control of the R protein signaling. It has been shown that overexpression of R proteins results in autoimmunity, which means that the activities of R proteins are normally under strict cellular control (Li et al. 2001; Shirano et al. 2002; Spoel and Dong 2012). In Arabidopsis mutants defective in the SKP1–CULLIN 1–F-box ubiquitin ligase complex, resulted in accumulation of higher levels of the R proteins SNC1 and RPS2, as well as in autoimmunity (Cheng et al. 2011). Moreover, mutation within one of the two E1 ubiquitin-activating enzymes in the *snc1* mutant background resulted in loss of the constitutive defense response phenotype, constitutive *PR* gene expression and resistance, seen in the *snc1* mutant plants. This suggests that the absence of the ubiquitin-activating enzyme results in failure of ubiquitin ligation to the proteins negatively regulating defense response to be catabolized by the proteasome (Goritschnig et al. 2007). Similarly, the turnover of the nuclear NPR1 (NONEXPRESSER of PR GENES 1) protein plays an important role in modulating transcription of its target genes and regulating plant defense (Spoel et al. 2009). The ubiquitin/proteasome pathway is upregulated during tomato-*C. fulvum* incompatible interaction suggesting the possible requirement for eliminating negative regulators for initiation of defense responses (Rowland et al. 2005).

Plant Transcription Factors

Transcription factors (TFs) are key regulators of gene expression controlling myriad biological processes including development, reproduction and immunity. The mechanism by which TFs achieve gene regulation is by binding to specific DNA sequences (*cis*-regulatory elements) in the promoters or enhancers of their target genes to promote or repress their transcription by RNA polymerase (Maston et al. 2006).

Some plant TFs are encoded by members of multigene families that expanded much more dramatically during land plant evolution than during the evolution of animals and fungi (Melzer and Theissen 2011). The Arabidopsis genome encodes about 1600 TFs, accounting for about 6% of the estimated 26,000 protein-coding genes, of which 45% belong to families common to *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* (Riechmann et al. 2000). The Arabidopsis multimember TF families include the MYB, MADS, basic helix-loop-helix (bHLH), APETALA2 (AP2)/ETHYLENE RESPONSE ELEMENT BINDING PROTEIN (EREBP) and WRKY families (Qu and Zhu 2006). Extensive expression and reverse genetic analyses in Arabidopsis identified the MYB superfamily (largest TF superfamily in plants) to be involved in a multitude of physiological processes (Yanhui et al. 2006), BHLH TF family in controlling cell proliferation and development of specific cell lineages (Toledo-Ortiz et al. 2003), MADS TFs in the regulation of flower-related physiological and developmental processes, and AP2/EREBP and WRKY TF families in responses to different biotic

and abiotic stresses (Eulgem and Somssich 2007; Feng et al. 2005; Gutterson and Reuber 2004; Rushton et al. 2010).

WRKY TFs belong to the WRKY-GCM1 superfamily of zinc finger TFs that evolved from Mutator or Mutatorlike (Mule) transposases (Babu et al. 2006). This TF family has been long thought to be plant specific until a single *WRKY* member was reported in the slime mold *Dictyostelium discoideum* and the intestinal parasite *Giardia lamblia* (Glockner et al. 2002; Pan et al. 2009). The WRKY protein family is characterized by a highly conserved ~60 amino acid long stretch containing WRKY domain and zinc-finger-like motif (Eulgem et al. 2000). The WRKYGQK amino acids at the N-termini define the WRKY domain although slight variations within this heptapeptide have been reported in some WRKY proteins of Arabidopsis (Zhang and Wang 2005), rice (*Oryza sativa*) (Xie et al. 2005), tobacco (*Nicotiana tabacum*) (van Verk et al. 2008), barley (*Hordeum vulgare*) (Mangelsen et al. 2008) and canola (*Brassica napus*) (Yang et al. 2009). Both of these two motifs are vital for the high binding affinity of WRKY transcription factors to the consensus cis-acting elements termed the W-box (TTGACT/C), although alternative binding sites have been identified (Cai et al. 2008; Sun et al. 2003; van Verk et al. 2008). WRKY proteins with highly identical DNA binding sequences exhibit different binding specificities. This is partly attributed to additional adjacent DNA sequences flanking the TTGACT/C-core motif (Ciolkowski et al. 2008).

The Arabidopsis WRKY family is divided into three main groups based on both the number of WRKY domains and the features of their zinc-finger motif

(Eulgem et al. 2000). The group I members are mostly characterized by the presence of two WRKY domains. Group II has one WRKY domain containing the same Cys2-His2 zinc-finger motif and group III has one WRKY domain containing the different Cys2-His/Cys Cys2-His2 zinc-finger motif. The group II WRKY proteins are further divided into subgroups A–E based on additional conserved structural motifs outside the WRKY domain (Eulgem et al. 2000).

Since its discovery almost two decades ago, extensive research in Arabidopsis and rice demonstrated pivotal roles for the different WRKY proteins in plant responses to biotic (Knoth et al. 2007; Tao et al. 2009) and abiotic (Jiang and Deyholos 2009; Wu et al. 2009) stresses, seed development (Luo et al. 2005; Zhang et al. 2004), and leaf senescence (Jing et al. 2009; Robatzek and Somssich 2002). Moreover, with the advent of new sequencing technologies, genome-wide identification characterization or gene expression analyses of the WRKY gene family has been performed in several additional plant species including castor bean (*Ricinus communis)* (Li et al. 2012), tomato (Huang et al. 2012), poplar (*Populus trichocarpa*) (He et al. 2012), maize (*Zea mays*) (Wei et al. 2012), cucumber (*Cucumis sativus*) (Ling et al. 2011), western white pine (*Pinus monticola*) (Liu and Ekramoddoullah 2009) and barley (Mangelsen et al. 2008).

Roles for TFs in diverse functions in additional plant species have recently been elucidated. In *Catharanthus roseus Cr*WRKY1 positively regulates the terpenoid indole alkaloid biosynthesis (Suttipanta et al. 2011). In *Capsicum annuum Ca*WRKYb is required for resistance against Tobacco mosaic virus (TMV) by inducing the

expression of *CaPR-10*, *CaPR-1*, and *CaPR-5* (Lim et al. 2011). WRKY72-type TFs contribute to basal immunity in tomato and Arabidopsis as well as to gene-for-gene resistance mediated by *Mi-1.2* (Bhattarai et al. 2010). Roles have been suggested for potato WRKY TFs in arbuscular mycorrhizal establishment possibly by controlling plant defense genes (Gallou et al., 2012). In *Vitis pseudoreticulata Vp*WRKY1 and *Vp*WRKY2 enhanced salt and cold tolerance as well as resistance to powdery mildew *Erysiphe cichoracearum* when expressed in Arabidopsis (Li et al. 2010). In *Artemisia annua Aa*WRKY1 binds Amorpha-4,11-diene synthase (ADS) that catalyzes the conversion of farnesyl diphosphate into amorpha-4,11-diene, the first committed step in the biosynthesis of the antimalarial drug artemisinin suggesting a role for this TF in the regulation of artemisinin biosynthesis (Ma et al. 2009). *Nicotiana attenuata* WRKY3 and WRKY6 coordinate responses to herbivory (Skibbe et al. 2008). In cotton (*Gossypium hirsutum) Gh*WRKY3 was suggested to play an important role in plant defense responses and fulfill a pivotal role in plant development (Guo et al. 2011). In banana transient overexpression of the *WRKY71* led to the induction of several genes, homologs of which are involved in diverse stress responses suggesting a major regulatory role for this TF in banana (Shekhawat et al. 2011).

Interestingly WRKY function is conserved in diverse plant species, which was demonstrated by heterologous expression of grape (*Vitis vinifera)* and *Thlaspi caerulescens* WRKY members in tobacco (Marchive et al. 2007; Mzid et al. 2007; Wei et al. 2008), and expression of soybean (*Glycine max)*, strawberry, rice, wheat WRKY members in Arabidopsis (Encinas-Villarejo et al. 2009; Hwang et al. 2011;

Proietti et al. 2011; Yu et al. 2010; Zhou et al. 2008). Moreover, this conservation of function enabled the characterization of TFs in species that are not amenable to gene silencing or overexpression.

Despite the explosion of the number of published research papers involving characterization of WRKY TFs in diverse plant species (Rushton et al. 2010), this family lacks universal and consistent nomenclature. Consequently the naming of the *WRKY* genes in the different plant species is not based on putative orthologous relationship. As part of whole genome analysis, the rice and maize WRKY TFs were annotated according to the order of the *WRKY* genes appearing on the chromosomes (Wei et al. 2012; Wu et al. 2005). Similarly, some individually characterized *WRKY* genes in strawberry (Encinas-Villarejo *et al.,* 2009), hot pepper (Park *et al.,* 2006) Madagascar periwinkle (Suttipanta et al. 2011) were annotated as WRKY1 based on their order of identification. While, WRKY TFs from tomato (Atamian et al. 2012; Bhattarai et al. 2010), populus (Levee et al. 2009), grape (Liu et al. 2010), canola (Yang et al. 2009) and cotton (Yao et al. 2011) were annotated by simply blasting against the Arabidopsis TFs and using the corresponding Arabidopsis nomenclature. It is worth mentioning these studies lack phylogenetic analysis, which is necessary for accurate assignment of orthologous relationship. Taken together, the names of hundreds of WRKYs are inconsistent and do not indicate orthologous relationships complicating communication among researchers working on this family of TFs in different plant species. One of the goals of this thesis is to create a systemic nomenclature.

Role of Phytohormones in Plant Immunity

High level of regulation of plant defense responses is mediated by three phytohormones salicylic acid (SA), jasmonic acid (JA), and ethylene (ET) during the different layers of plant defense. A central role for these hormones in plant resistance was demonstrated using mutants that are defective in different steps of biosynthesis, perception, and/or signaling of these phytohormones. Together, these three major phytohormones fine-tune the temporal and spatial regulation of various plant defense cascades and propagate the resistance signal(s) originated from the recognition step for optimum resistance output. Roles for hormone-signaling pathways have long been identified in *R*-gene-mediated responses, with major steps of the different pathways being genetically defined. More than two decades of extensive research elucidated the molecular basis for hormone perception and signal transduction (Erb et al. 2012; Robert-Seilaniantz et al. 2011; Shan et al. 2012).

The interaction of 7-iso-jasmonoyl-L-isoleucine (JA-Ile) with its receptor (the SCF^{COI1} complex) triggers the SCF-26S proteasome-mediated proteolysis of negative regulators, such as JAZ repressor proteins, resulting in activation of downstream JA-responsive genes (Chini et al. 2007; Thines et al. 2007). Unlike the JA receptors that are positive regulators, the ET receptors are negative regulators of ET signaling. Consequently, ET binds to these receptors and inactivates them resulting in the accumulation of positive regulators (Shan et al. 2012). The SA signaling is more complex that includes both *NPR1*-dependent and independent pathways. In

Arabidopsis, NPR1 represents a key node in signaling downstream from SA where its degradation acts as a molecular switch (Durrant and Dong 2004; Spoel et al. 2009). After decades of search for the SA receptor in plants, recently it was shown that the Arabidopsis NPR1 paralogs NPR3 and NPR4 are the SA receptors and function as adaptors of Cullin 3 ubiquitin E3 ligase to mediate NPR1 degradation in a SA-regulated manner (Fu et al. 2012).

Generally in Arabidopsis, SA is involved in plant defense responses against biotrophic and hemi-biotrophic pathogens as well as the establishment of systemic acquired resistance. By contrast, JA and ET are usually associated with defense against necrotrophic pathogens and herbivorous insects (Bari and Jones 2009; Glazebrook 2005). While this model is generally correct, several exceptions have been demonstrated. The SA-signaling suppresses the Arabidopsis defense responses against the silverleaf whitefly (*B. tabaci*) pest (Zarate et al. 2007) and promote Arabidopsis susceptibility to wilt disease caused by the necrotrophic root-infecting fungal pathogen *F. oxysporum* (Thatcher et al. 2009).

Complex interactions among the different hormone-signaling pathways have also been demonstrated. Although most reports of SA and JA/ET defense pathways are antagonistic, synergistic interactions also exist (Mur et al. 2006; Robert-Seilaniantz et al. 2011). Since in nature different types of pathogens or pests simultaneously attack plants, their ability to fine-tune various signaling pathways for optimal defense responses with a minimal fitness cost is crucial for their survival (Koornneef and Pieterse 2008). Upon infection with the biotroph *Pseudomonas*

syringae, which induces SA-mediated defense, Arabidopsis plants become more susceptible to the necrotrophic pathogen *Alternaria brassicicola* as a result of suppression of the JA-signaling pathway (Spoel et al. 2007). An example of synergistic interaction includes the enhanced Arabidopsis resistance against virulent *P. syringae* upon simultaneous activation of SA- and JA-dependent defense pathways (Mur et al. 2006; van Wees et al. 2000).

Interestingly, some pathogens produce hormone mimics to manipulate host hormone-signaling pathways to their advantage. A well known example is the phytotoxin coronatine (COR), a JA-Ile analogue, produced by *P. syringae* to induce the JA signaling in plants, consequently resulting in the suppression of SA signaling, which is the effective defense signaling pathway against this pathogen (Uppalapati et al. 2007; Zheng et al. 2012). In addition to producing hormones themselves, some pathogens can induce hormone production by their host. This phenomenon is exemplified by *F. oxysporum* that hijacks the JA-signaling pathway to cause wilt-disease symptoms that lead to plant death in Arabidopsis (Thatcher et al. 2009).

JA-regulated genes are differentially upregulated at higher levels in *Mi-1.2* incompatible interaction compared to compatible interaction in both tomato roots and leaves, however JA signaling in not required for *Mi-1.2*-mediated resistance to both aphid and RKN (Bhattarai et al. 2008). In contrast, an intact JA signaling pathway is required for efficient RKN infection (Bhattarai et al. 2008). Moreover JA-signaling pathway is required for *Medicago truncatula* resistance against the bluegreen aphid

Ac*yrthosiphon kondoi)* (Gao et al. 2007) and activation of the JA-signaling pathway enhances Arabidopsis resistance against *M. persicae* (Ellis et al. 2002).

In spite of the role of SA in defense against biotrophic and hemibiotrophic pathogens, most *R*-gene mediated resistance in tomato do not require SA. In tomato, SA is not required for *Pto*-mediated resistance to *P. syringae* or *Cf-2*- and *Cf-9*-mediated resistance *C. fulvum* (Brading et al. 2000; Oldroyd and Staskawicz 1998). In contrast, a role for the SA has been demonstrated in *Mi-1.2*-mediated resistance to aphids. In transgenic tomato plants with *Mi-1.2* expressing *NahG*, a gene from *Pseudomonas putida* that encodes SA hydrolase an enzyme that metabolizes SA to catechol, reduction in potato aphid resistance was observed (Gaffney et al. 1993; Li et al. 2006b). However, the role for SA in *Mi-1.2*-mediated resistance against RKN remains unclear. *Mi-1.2*-mediated resistance was compromised in tomato transgenic hairy roots but not in transgenic tomato plants expressing *NahG* (Bhattarai et al. 2008; Branch et al. 2004). This lack of consistency in RKN resistance could be due to residual levels of SA in transgenic tomato plants that are sufficient for resistance.

Functional Genomics in Tomato

The most direct and decisive approach to identify a gene function in plants involves studying plants with altered expression of the respective gene. Over the last 2 decades, the application of several gene knockout/knockdown and overexpression approaches in *planta* provided unprecedented information regarding function and mechanism of actions of genes in the model plant species Arabidopsis and rice.

Tomato, an economically important crop worldwide, has emerged as a model system for genetic studies in Solanaceous species mainly due to its simple diploid genetics, short generation time, routine transformation technology, and availability of rich genetic and genomic resources (Barone et al. 2008). Tomato is well adapted to traditional genomic analysis (forward genetics) that relies on random introduction of mutations into the genome and screening for mutant phenotype. Screening for loss-of-function mutants has been a primary tool for dissecting genetic pathways in many organisms, including Arabidopsis. Such random mutagenized tomato populations have been generated through chemical mutagenesis (ethylmethane sulfonate [EMS]), insertional mutagenesis (maize *Ac/D* transposon) (Meissner et al. 2000) and fast-neutron bombardment (Martinez de Ilarduya et al. 2001; David-Schwartz et al. 2001). Although the recent release of the tomato genome sequence will accelerate the discovery of various regulatory and biochemical pathways operating within this family, efficient tools for reverse genetics unfortunately are not yet routinely available for tomato (Emmanuel and Levy 2002). Although transfer-DNA (T-DNA) insertional mutagenesis is a highly effective reverse genetics tool for genome-wide mutagenesis in Arabidopsis (Alonso and Ecker 2006), T-DNA insertion tomato populations are not yet available.

In plant species where no T-DNA insertion mutants are available alternate approaches to generate knockdown plants have been developed. This is achieved by using RNA interference [RNAi, also known as post-transcriptional gene silencing (PTGS) or co-suppression], which is another powerful technology currently available

for analysis of gene function. RNAi is triggered by either endogenous or exogenous dsRNA, and silences endogenous genes carrying homologous sequences at both the transcriptional and post-transcriptional levels (Tomoyasu et al. 2008). Some of the advantages of this reverse genetics approach over gene disruption by T-DNA insertion are the ability to silence multiple gene family members with a single RNAi-inducing transgene and that gene knockdowns due to RNAi are dominant (Preuss and Pikaard 2003). Several families of RNAi vectors that make use of *Agrobacterium tumefaciens*-mediated delivery into plants have been developed and made available to the public (Preuss and Pikaard 2003). However, RNAi cannot be used with several plant species including some agricultural important crops, as they are not yet amenable to stable genetic transformation. Tomato stable plant transformation is performed routinely however, the length of time required to produce transgenic lines limits the analysis of gene function in this species (Shibata 2005).

As an alternative to stable plant transformation to generate gene knockdown plants, virus-induced gene silencing (VIGS), a rapid virus-mediated transient gene knockdown approach, is increasingly being used for characterization of gene function in tomato. As apparent from its name, VIGS approach involves modification of viruses that naturally infect certain plants to carry portions of plant genes. The recombinant virus vector carrying a plant sequence is placed between right border (RB) and left border (LB) sites of the T-DNA and transformed into *A. tumefaciens*. This enables the rapid generation of the recombinant virus and its delivery into the plant via *A. tumefaciens* infiltration (agroinfiltration), which results in rapid gene

silencing effect (Benedito et al. 2004). VIGS has been performed in numerous plant species including monocots, legumes, cucurbits and *Rosaceae* fruit tree species using different viral vectors (Sasaki et al. 2011).

In plants agroinfiltrated with this recombinant virus vector, dsRNA are produced as an intermediate step during virus replication, which triggers RNAi that leads to degradation of the endogenous mRNA homologous to the inserted plant sequences. In general the effectiveness of VIGS depends on the virus vector, the host plant species and the targeted gene sequence.

VIGS has many advantages compared to other loss-of-gene function mutation approaches. These include the rapid generation of phenotype and no need for plant transformation, characterization of lethal phenotypes, potential to silence either individual or multiple members of a gene family, low cost of experiments and the possibility of conducting large-scale screening studies (Burch-Smith et al. 2004; Unver and Budak 2009).

A major disadvantage of VIGS is that incomplete silencing that results in plants that consist of a mosaic of silenced and non-silenced tissues evident by the patchy photobleaching symptoms observed by silencing the *phytoene desaturase* (*PDS*) gene involved in carotenoid biosynthesis (Bhattarai et al. 2007). This VIGS effect, which has been observed with all plants tested, is an impediment to the broad application of this technique (Liu and Page 2008). This means that in the absence of a visual phenotype, it is impossible to know which part of a tissue is efficiently silenced requiring the use of a large number of plants in a single experiment and multiple

replication of experiments before drawing definite conclusions. Orzaez et al. developed an anthocyanin-guided VIGS for tomato fruit in order to overcome this limitation of efficiency and patchiness (Orzaez et al. 2009). However this approach did not gain popularity due to the fact that disturbing anthocyanin production can interfere with function of certain genes and pathways not related directly to anthocyanin production.

Tobacco mosaic virus (TMV) was the first viral vector used to successfully elicit VIGS in plants (Kumagai et al. 1995). Another early VIGS vector was based on *Potato virus X* (PVX) (Ruiz et al. 1998). Although more efficient than TMV, its use was not extensive due to inability to infect meristematic tissue and a more limited host range (Burch-Smith et al. 2004). Tobacco rattle virus (TRV), a single-stranded RNA virus with a bipartite genome, was developed as a very successful plant VIGS vector (Liu et al. 2002; Ratcliff et al. 2001). Advantages of this vector include efficient spreading over the entire plant tissues including the meristem, ability to infect a large number of plant species and to induce very mild host symptoms. TRV has a reported host range of over 60 plant species from 12 families including monocots. To date, TRV has been successfully used for gene silencing in the lower eudicots *Aquilegia vulgaris* (Gould and Kramer 2007) and opium poppy (Hileman et al. 2005), petunia (Chen et al. 2005), Arabidopsis (Burch-Smith et al. 2006) and Solanaceae species including tomato (Liu et al. 2002), potato (Brigneti et al. 2004), *Nicotiana benthamiana* (Liu et al. 2004), eggplant (Liu et al. 2012) and chili pepper (Chung et al. 2004).

Additional Defense Genes Characterized in Tomato

Common molecular signaling pathways have been shown to operate in plant defense responses against various invading pathogens/pests, with some being conserved among evolutionary diverse plant species. For example, recently, a conserved response against the damping off fungus Pythium was demonstrated in Arabidopsis and the non-vascular plant *Physcomitrella patens* that has diverged from flowering plants at least 450 million years ago (Oliver et al. 2009). Although significant advances are made in understanding pathogen perception and identification of signal transduction components that link the perception of pathogens with downstream responses in the model plant species Arabidopsis and rice (Chen and Ronald 2011; Ryan et al. 2007), limited such information exists in tomato (Pedley and Martin 2004).

Similar to Arabidopsis, some of the identified defense signaling components in tomato has been shown to be required for more than one gene-for-gene interaction (Ekengren et al. 2003; Lu et al. 2003b; Peart et al. 2002). The successful application of VIGS in tomato provided a fast and high-throughput approach for gene functional analysis (Lu et al. 2003a). However the moderate efficiency of VIGS in tomato, prompted for the development of a heterologous system in a relative of tobacco, *Nicotiana benthamiana*. In addition to more efficient silencing observed in *N. benthamiana*, this plant species is highly amenable for transient expression that is not efficiently performed in other plant species including tomato. Thus, *N. benthamiana* is used to perform high-throughput functional screens to identify signal-transduction components of solanaceae *R*-genes that produce HR by transiently co-expressed with

their matching pathogen effector or expressing their constitutive active forms (Mantelin et al. 2011; Rowland et al. 2005; Wangdi et al. 2010). In addition to high-throughput screens, *N. benthamiana* is also used to analyze gene function and identify mechanisms of pathogen/pest recognition (Du et al. 2012; Lukasik-Shreepaathy et al. 2012; Tameling et al. 2006).

Several independent VIGS screens have been conducted to identify genes that function within the complex signaling networks of tomato *R*-gene mediated defense responses. High-throughput VIGS in *N. benthamiana* was conducted using the potato virus X (PVX) vector. A normalized library of *N. benthamiana* cDNA was generated in a vector-derived from PVX. Plants were infected with individual PVX constructs from the library and screened based on a hypersensitive cell death response that is elicited by the bacterial AvrPto protein in the presence of Pto. Those genes affecting the Pto-induced HR were assayed for *Pto*-mediated resistance against *Pseudomonas syringae*. Some of the cDNAs, that affected the Pto-induced HR upon silencing, correspond to *Heat-Shock Protein 90* (*HSP90). HSP90* has been shown to be required for *Pto*-mediated resistance against *P. syringae*. Moreover, silencing *HSP90* compromised *Rx*-mediated resistance against potato virus X and *N*-mediated tobacco mosaic virus (TMV) resistance (Lu et al. 2003b). A similar normalized *N. benthamiana* cDNA library cloned into PVX vector was used for VIGS on *N*-transgenic *N. benthamiana* plants. The plants were assayed for attenuation of *N* resistance three weeks after VIGS by inoculating each plant with a GFP-tagged strain of TMV (TMV:GFP). A role for *N requirement gene 1* (*NRG1*) in *N*-mediated TMV

resistance was identified (Peart et al. 2005). Another large-scale forward-genetics screen using VIGS and a cell death-based assay identified 14 *N. benthamiana* genes involved in pathogen-associated molecular pattern-triggered immunity (Chakravarthy et al. 2010).

The role of additional genes, orthologs of which are upregulated during incompatible interactions or characterized in model plant species to be involved in defense, were identified in tomato and other non-model species using VIGS (Ekengren et al. 2003; Scofield et al. 2005).

HSP90, suppressor of the G2 allele of SKP1 (SGT1), and required for *Mla 12* resistance (RAR1) are among the proteins that have been shown to be required for *R*-gene mediated resistance responses in diverse plant species (reviewed by Shirasu 2009). HSP90 is a highly conserved molecular chaperone implicated in the assembly, stabilization and maturation of key signaling proteins in eukaryotic cells (Kadota et al. 2010; Pearl and Prodromou 2006), whereas its interacting proteins SGT1 and RAR1 are thought to serve as HSP90 co-chaperones (Fu et al. 2009). Using VIGS in tomato, HSP90 has been demonstrated to be an integral component of *Mi-1.2*-, *Pto*-, *Ve*-, *Cf-9*-, and *Bs4*-mediated defense signaling pathways (Bhattarai et al. 2007; Ekengren et al. 2003; Fradin et al. 2009; Shirasu 2009). The role of RAR1 on the other hand seems to be not highly conserved, as it is not part of the *Mi-1.2*- and *Ve*-signaling pathways (Bhattarai et al. 2007; Fradin et al. 2009). The HSP90 co-chaperone SGT1 seem to be required for majority of *R*-gene-mediated resistances including *Mi-1.2*- and *Pto*-mediated resistance and HR induced by *Cf-4* and *Cf-9* after infiltration with Avr4 and

Avr9, respectively (Bhattarai et al. 2007; Ekengren et al. 2003; Gabriels et al. 2007). SGT1 has been also shown to be required for the function of SCF ubiquitin ligases and other multiprotein complexes. Similarly, Avr9/Cf-9–INDUCED F-BOX1 (ACIF1), a conserved protein closely related to F-box proteins regulating plant hormone signaling, encodes an F-box protein with a LRR domain that is recruited to SCF complexes. Silencing tobacco ACIF1 has been shown to result in HR attenuation triggered by various elicitors from distinct classes of pathogens including Avr9, and Avr4 from the fungus *C. fulvum*, AvrPto from *P. syringae*, Inf1 from *Phytophthora infestans* and P50 helicase of TMV (van den Burg et al. 2008). Interestingly, *ACIF1* silencing attenuated the *Cf-9*–dependent HR but not *Cf-9* resistance to *C. fulvum* indicating that the development of HR is distinct from resistance (van den Burg et al. 2008). However, resistance conferred by the *Cf-9* homolog *Cf-9B* to *C. fulvum* and *Ve* to *V. dahliae* was compromised in *ACIF1*-silenced tomato supporting a role of *ACIF1* in these two tomato defense responses (Fradin et al. 2009; van den Burg et al. 2008).

The MAPKs, one of the largest group of plant kinases, comprise another important signaling cascade that function in the regulation of plant defense reactions by altering the activity of the different signal transduction pathways through phosphorylation/dephosphorylation of its proteins (Taj et al. 2010). The MAPKs are a linear cascade of three consecutively acting protein kinases that are involved in various plant processes (reviewed by Mishra et al. 2006). The tomato *Le*MPK1, *Le*MPK2, and *Le*MPK3 activity have been shown to be induced by a number of biotic and abiotic elicitors (Stulemeijer et al. 2007). Moreover roles for tomato *LeMKK2* and

three MAPKs, *LeMPK1*, *LeMPK2*, and *LeMPK3* have been demonstrated in *Mi-1.2*–mediated aphid resistance (Li et al. 2006b) and in tomato cultivar "Hawaii 7996" with stable resistance against bacterial wilt caused by *Ralstonia solanacearum* (Chen et al. 2009). Moreover, VIGS of *LeMPK1* and *LeMPK3* revealed a role in *Cf*-mediated resistance against *C. fulvum* (Stulemeijer et al. 2007) with the later having a minor role in *Ve*-mediated resistance as well (Fradin et al. 2009). Furthermore, *NTF6*, *WIPK*, *MEK1,* and *MEK2* have been shown to play an important role in the *Pto*-mediated resistance in tomato (Ekengren et al. 2003), the latter being involved in *Ve*-mediated resistance as well (Fradin et al. 2009). These results suggest that one or more MAPK cascades operate downstream of distinct *R*-genes and that common defense pathways might be activated in resistance to diverse pests and pathogens (Li et al. 2006b; Pitzschke et al. 2009; Taj et al. 2010).

Protein kinases are one of the largest gene families in plants. They play an important role in controlling protein activity and cellular signaling. A member of a subfamily of protein kinase ACIK1 is required for *Cf-9* and *Cf-4*-mediated resistance and HR in tomato but not for HR or resistance mediated by, *Rx, N* or *Pto* R-genes (Rowland et al. 2005).

The Cf-9-interacting thioredoxin (CITRX) identified through yeast two-hybrid screen has been shown to be a negative regulator of *Cf-9* but not *Cf-2*-mediated HR and resistance, indicating a distinct requirement for these two gene family members (Rivas et al. 2004). Moreover, CITRX has been shown to mediate physical association between the cytoplasmic domain of Cf-9 and the ACIK1 (Nekrasov et al. 2006).

Two broad-spectrum signaling mediators, NON-RACE-SPECIFIC DISEASE RESISTANCE1 (NDR1) and ENHANCED DISEASE SUSCEPTIBILITY 1 (EDS1), were initially identified in Arabidopsis as essential components of resistance mediated by CC-NB-LRR and Toll and interleukin-1 receptor-like (TIR)-NB-LRR encoding *R*-genes, respectively. However, several CC-NB-LRR R-proteins were later found to function independent of NDR1. NDR1, a plasma membrane-localized protein, is believed to play a role in signaling from within the apoplast, transducing the signal from the extracellular space to within the cell (Knepper et al. 2011). In tomato, NDR1 has been shown to be a component of the extracellular receptor-like *R* gene *Ve*-mediated resistance (Fradin et al. 2009). EDS1, having homology to eukaryotic lipases and serving as a central regulatory protein involved in both biotic and oxidative stress signaling (Wiermer et al. 2005), is required for the resistance mediated by the TIR-NB-LRR genes *N, BS4*, and *I* as well as to the extracellular *R* gene *Ve* (Fradin et al. 2009; Hu et al. 2005). In contrast, *EDS1* is not required for *Mi-1.2*-mediated resistance (Hu et al. 2005).

The *NRC1*-gene encodes a CC-NB-LRR resistance protein analogue and has been shown to be an essential downstream signaling component in *Cf-4*- and *Ve*-mediated resistances suggesting that these extracellular receptor-like proteins require cytoplasmic NB-LRR protein for their function (Gabriels et al. 2007). NRC1 is also required in cell death signaling pathways triggered *by Cf-9, LeEix2, Pto, Mi-1.2 R*-genes but not for the *Pto*-mediated resistance against *P. syringae*, *Rx*-mediated resistance against PVX and *N*-mediated resistance against TMV (Gabriels et al. 2007).

NPR1, a key component of SA-mediated signaling leading to SAR, is a master regulator downstream of SA. It is constitutively expressed and translocated to the nucleus after induction by SA consequently promoting the activation of pathogenesis-related (*PR*) genes through interaction with TGA transcription factors (Dong 2004). In Arabidopsis, *NPR1* is required for *RPP5*-mediated resistance against *Hyaloperonospora arabidopsidis (Hpa)* but not *RPP8*-mediated *Hpa* resistance. Moreover, *NPR1* is not required for *RPM1*-, *RPS2*- and *RPS4*-mediated resistance against *P. syringae* (Rairdan and Delaney 2002). Taken together, these results suggest that NPR1 might have a limited role in Arabidopsis *R*-gene mediated resistances.

A role for NPR1 and its interacting proteins TGA1a and TGA2.2 have been demonstrated in *Pto*-mediated resistance against *P. syrinage* (Ekengren et al. 2003) and in "Hawaii 7996" tomato resistance against the bacterial wilt *R. solanacearum* (Chen et al. 2009). A minor role for NPR1 has also been implicated in *Ve*-mediated resistance but not been shown to be involved in other tomato resistances. NPR1 does not seem to be required for *Mi-1.2*-mediated resistance (Kaloshian, unpublished result). Moreover roles for WRKY TFs *SlWRKY72a* and *SlWRKY72b*, functioning in a *NPR1*-independent pathway, in *Mi-1.2*-mediated defense against potato aphids and nematodes have been demonstrated (Bhattarai et al. 2010). Tomato genome encodes for a total of six putative *At*NPR1-like proteins. Recent studies in Arabidopsis showed that *At*NPR3 and *At*NPR4 are SA receptors (Fu et al. 2012). It is possible that one or more of the other tomato *Sl*Npr1 homologs are involved in the *Mi-1.2*-mediated resistance.

Receptor-like kinases (RLKs) are surface localized, transmembrane receptors that have been shown to recognize distinct ligands of microbial origin or ligands derived from intracellular protein/carbohydrate signals (Greeff et al. 2012). The somatic embryogenesis receptor kinases (SERKs) belong to a large family of LRR receptor like kinases (LRR-RLKs). Arabidopsis encodes for five SERK members that have been shown to be involved in various processes including embryogenic competence, cell adhesion during organ abscission, male sporogenesis and immunity (Albrecht et al. 2005; Heese et al. 2007; Lewis et al. 2010). Recently it was shown that the tomato Somatic Embryogenesis Receptor Kinase 1 *(SlSERK1)*, the ortholog of the Arabidopsis *SERK1* and *SERK2* is crucial for *Mi-1*-dependent resistance to potato aphids (Mantelin et al. 2011). The *AtSERK1* has also been shown to be required for *Ve1*-mediated Arabidopsis resistance against race 1 strains of both *V. dahliae* and *V. albo-atrum* (Fradin et al. 2011).

The second part of my thesis investigates the tomato interaction with its aphid pest from the aphid's perspective. In the following two sections I want to introduce aphid biology and what is known regarding its interaction with host plants.

Aphids have a Unique Biology

Aphids have somewhat complex life cycle, comprising of both sexual and asexual (parthenogenetic) modes of reproduction and host alternation (Blackman and Eastop 2000). Asexual mode of reproduction occurs during most of the year with sexual reproduction happening only before winter where eggs are laid on a perennial plant for

overwintering. The mechanisms and biochemical processes behind this vast change in behavior due to perception of environmental cues is poorly understood.

Aphid species have also high diversity in terms of host range and host plant specialization. Moreover, they possess a diverse symbiont community that includes the obligate bacterial symbiont *Buchnera aphidicola* (Buchner 1965), as well as several facultative symbionts that vary among aphid populations of the same aphid species (Tsuchida et al. 2002). Current data raise the hypotheses that symbionts may both positively and negatively contribute to the insect's ability to utilize different plants as hosts as well as positively affect its fitness (Ferrari et al. 2007; Leonardo and Muiru 2003; Tsuchida et al. 2004). This tritrophic interaction among host plant, insect and endosymbiont makes aphids one of the organisms to study the evolution of tritrophic interactions.

Wing dimorphism is another exciting characteristic of aphids, since winged and wingless phenotypes differ in a range of morphological, physiological, life history and behavioral features representing a clear example of adaptive phenotypic plasticity (Braendle et al. 2006; Brisson 2010). This dimorphism can be environmentally induced known as polyphenism, or genetically determined known as polymorphism (Braendle et al. 2006). The aphid's phenotypic plasticity, via epigenetic control (Srinivasan and Brisson 2012), provides an opportunity to address the relationship between environmental and genetic induction of alternative phenotypes that is also thought to contribute to aphid's adaptation to diverse conditions, as aphid's genetic

polymorphism does not provide a complete explanation of the observed trait variability (Lombaert et al. 2009).

Understanding the complex relationship of aphids with their hosts has been a long-standing scientific quest. During feeding, aphids penetrate plant tissues while moving their slender stylets primarily intercellularly towards the phloem sieve elements where they feed and secrete two types of saliva during the process. This highly specialized mode of feeding causes little apparent damage to the plant enabling aphids to evade a wide range of plant biochemical defenses (Walling 2008). Consequently aphids are able to ingest phloem sap continuously for many hours or even days from a single sieve element (Tjallingii 1995).

The genome sequence of the pea aphid (*Acyrthosiphon pisum)* provided a valuable resource to investigate aphid biology at the genome level (International Aphid Genomic Consortium 2010). Several publications reported genome-wide detailed analyses of specific aspects of major gene families or genes involved in certain traits associated with aphids (International Aphid Genomics Consortium 2010, and references within). The pea aphid genome is predicted to encode more than 30,000 genes. Extensive gene duplication in more than 2000 gene families like those involved in chromatin modification, miRNA synthesis, and sugar transport exists. In contrast, loss of evolutionarily conserved genes central to the selenoprotein utilization, purine salvage, and the entire urea cycle was identified. An extensive annotation of the immune and stress gene repertoire, by looking for the presence of homologs in pea aphid for 155 genes present in insect genomes characterized to date, showed that pea

aphid lacks several genes thought to be critical for recognition, signaling and killing of microbes (Gerardo et al. 2010). These include: (1) peptidoglycan receptor proteins (PGRPs), which recognize peptidoglycans present in cell walls of Gram-positive and Gram-negative bacteria and activate both the Toll and IMD/JNK pathways (Steiner 2004); (2) Class C scavenger receptors that facilitate phagocytosis and contributes to the suppression of bacterial infection in Drosophila (Lazzaro 2005); (3) members of the Nimrod superfamily that appear to function as receptors in phagocytosis and bacterial binding (Kurucz et al. 2007); and (4) many crucial components of the immune deficiency (IMD) signaling pathway that is mainly critical for fighting Gram-negative bacteria in Drosophila (Boutros et al. 2002). The lack of aphid homologs to many immune genes in insects like flies, mosquitoes and bees could be the result of the large evolutionary distance between aphids and these insects. The ancestors of aphids and these insects diverged approximately 350 million years ago (Gaunt and Miles 2002). An alternative explanation for the lack of known immune-related genes in pea aphids is that aphids mount an alternative, but equal, immune response (Gerardo et al. 2010). Finally symbiont-mediated host protection may explain why aphids have a reduced (or specialized) antimicrobial defense. Aphids coevolved with primary and secondary symbionts for millions of years (Baumann 2005; Tamas et al. 2002). The secondary symbiont, *Regiella insecticola*, has been shown to protect pea aphids against fungal pathogens (Scarborough et al. 2005), while another secondary symbiont, *Hamiltonella defensa*, provides protection against the parasitoid wasp *Aphidius ervi* (Oliver et al. 2005).

Pea aphid genome analysis also identified 12 novel genes belonging to the dynamin superfamily involved in the fission and fusion of membranes (Nakabachi and Miyagishima 2010). Among pea aphid neuropeptides and neurohormones that are important for perceiving environmental signals, homologs for corazonin, vasopressin and sulfakinin are missing, while the presence of 10 different genes coding insulin-related peptides was demonstrated (Huybrechts et al. 2010). Currently a wealth of EST sequences exists for a few additional aphid species such as *Myzus persicae* and *Aphis gossypii* providing the opportunity to conduct comparative studies within aphids.

Aphid-Host Interactions

Upon landing on a plant, aphids probe the leaf surface very briefly with their modified mouthpart called stylet. It is believed that via these first probes aphids are able to differentiate between host and non-host plants (Powell and Hardie 2000). Following the location of a suitable host, aphids insert their stylets, which move intercellularly until they reach the sieve elements. During the penetration process, gelling saliva is continuously secreted, which forms a lubricating and hardening sheath around the stylets (Tjallingii and Hogen Esch 1993). Also during the path to the sieve elements, the stylets briefly puncture cells (but do not run through them) and are withdrawn within a few seconds. During this process a small quantity of watery saliva is injected in the cytosol and a minute quantity of saliva/cytoplasm mixture is ingested (Tjallingii 2006). One of the purposes of these intracellular probes, which become more frequent near the phloem vessels, is to locate the position of the stylets within the plant tissues

by assessing the internal chemistry of the punctured cells (Hewer et al. 2010). After gathering the required information and stylet withdrawal, the punctured site is readily sealed by gelling saliva (Tjallingii and Hogen Esch 1993). At the end of the road when the stylet reaches the sieve elements aphids inject watery saliva before phloem uptake (Will et al. 2009). This watery saliva injection that occurs frequently during the feeding period is believed to counteract plant defense mechanisms (Prado and Tjallingii 2007).

It is speculated that the composition of the gelling saliva might be common among the different aphid species (Miles 1999). In contrast, the composition of the watery saliva has been shown to differ considerably (Baumann and Baumann 1995; Cherqui and Tjallingii 2000). Moreover it has been shown that the composition of the watery saliva differs within the same aphid species according to diet, suggesting that host plant range depends on variations of watery saliva composition (Carolan et al. 2009; Cherqui and Tjallingii 2000). Independent reports demonstrated enzymatic activity in the aphid watery saliva including pectinase and cellulase activities, which are thought to facilitate stylet penetration, although the actual mechanism of action for these enzymes have not been demonstrated (Giordanengo et al. 2010). Two oxidoreductases, polyphenol oxidase (PPO) and peroxidase (Px), were identified in the saliva of the grain aphid (*Sitobion avenae*) showing no high substrate specificity, as they could oxidize a wide range of phenolic compounds (Urbanska et al. 1998). Oxidoreductases have also been identified in the saliva of the spotted alfalfa aphid *Therioaphis maculata, A. pisum* and *Megoura viciae* (Harmel et al. 2008;

Madhusudhan and Miles 1998). It is believed that the presence of these enzymes in the saliva enable the aphid to neutralize the detrimental effect of a wide range of phenolic compounds before they are ingested (Urbanska et al. 1998).

Phytopathogens secrete proteins known as effectors to manipulate their hosts for effective colonization (Deslandes and Rivas 2012; Oliva et al. 2010; Schornack et al. 2009). However, as part of the continuous arms race between plants and their pathogens some of these elicitors or avirulence factors are recognized by the plant and result in activation of defense responses. Unlike the vast information available about pathogen effectors and elicitors, little is known about insect effectors and elicitors (Hogenhout and Bos 2011). Studies conducted on chewing insects have shown that regurgitants act as both elicitors inducing plant defense reactions and as effectors repressing plant immunity similar to pathogenic microbial effectors (Consales et al. 2011; Halitschke et al. 2001; Major and Constabel 2007; Musser et al. 2002).

Aphid saliva also contains proteins that counteract sieve-tube occlusion, a mechanism involved in plant defense against phloem-feeding insects that results in blockage of nutrient supply (Caillaud and Niemeyer 1996). In *Vicia faba*, the mechanism of counteraction is through the inhibition of Ca^{+2}-induced dispersion of forisomes (giant protein bodies that plug the sieve tubes) possibly by binding to Ca^{+2} (Will et al. 2007). Using proteomic approach (LC-MS/MS) two members of metalloprotease family (angiotensin-converting enzyme and M1 zinc-dependant metalloprotease), glucose-methanol-choline (GMC)-oxidoreductase, regucalcin and five other uncharacterized proteins have been identified in the *A. pisum* saliva

(Carolan et al. 2009). The metalloproteases might be involved in the degradation of the phloem proteins functioning in plant immunity resulting in suppression of plant defense responses. Moreover the phloem protein degradation might provide a supplementary source of nitrogen to the aphid by increasing/recycling available free amino acids (Carolan et al. 2009). A similar proteomic approach identified glucose oxidase, glucose dehydrogenase, NADH dehydrogenase, α-glucosidase and α-amylase in the *M. persicae* saliva (Harmel et al. 2008). Moreover, proteinaceous component(s) with a size between 3 and 10 kD in the *M. persicae* saliva has been shown to act as elicitor of Arabidopsis defense responses (De Vos and Jander 2009).

As an alternative approach for identification of putative salivary proteins, the salivary gland transcriptomes of *M. persicae* and *A. pisum* were sequenced (Carolan et al. 2011; Ramsey et al. 2007). Using bioinformatics and the signalP algorithm for prediction of secretion signal peptides at the N-terminal of proteins, 48 and 262 candidate proteins have been identified to be presumably secreted in the saliva of *M. persicae* and *A. pisum*, respectively (Bos et al. 2010; Carolan et al. 2011). Functional characterization of 48 candidate *M. persicae*-secreted proteins identified roles for Mp10 and Mp42 as elicitors of plant defense and a role for MpC002 in suppression of plant defense (Bos et al. 2010). Previously, it was shown that *A. pisum* C002 is injected into the host plant during aphid feeding in agreement with its role *in planta*. Moreover through RNAi-based transcript knockdown, it was shown that the *C002*

gene is important for the survival of the aphid on the host plant and that the knockdown of this gene impairs aphid's foraging and feeding abilities (Mutti et al. 2008).

Objectives of the dissertation research

In depth understanding of plant-pest interaction and pest biology at the molecular level will enable us to exploit the weakness of the pest and the strengths of the plant, the application of which will improve crop production by reducing the damage caused by the pests.

During *R* gene-mediated defense, different classes of transcription factors (TF) are implicated in regulation of downstream defense genes. Previous microarray analysis in our lab has shown that WRKY TFs are differentially regulated during *Mi-1.2*-mediated resistance against root-knot nematode (RKN). Moreover, roles for two WRKY TFs *SlWRKY72a and SlWRKY72b* have been demonstrated in this resistance as well as in basal defense against RKN and potato aphid. The first objective of this dissertation research is to characterize the function of additional WRKY TFs during *Mi-1.2*-mediated resistance against RKN and potato aphid. The WRKY TF family has expanded during evolution of plants. It is believed that this expansion is associated with development of highly sophisticated defense mechanisms in higher plants as part of the arms race between co-evolving plants and their pathogens/pests. Over the past decade, enormous progress has been achieved in characterization of roles for the WRKY TFs in plant defense and development. Moreover WRKY TFs have been

identified and characterized in diverse plant species. However the naming of these TFs was done in a non-systematic manner and consequently the names do not reflect true phylogenetic relationships. Therefore, the second objective of my dissertation was to identify criteria to clearly define orthologous relationships among the WRKY TFs from different plant species.

Aphids have unusual biology including two modes of reproduction, winged and wingless forms and intricate association with their host plants. We hypothesized that some of these aphid characteristics could be due to presence of a unique set of genes among aphids. The third objective of my dissertation was to sequence the transcriptome of the potato aphid and conduct a comparative analysis with other aphid and insect species to demonstrate the presence of such a unique set of genes in aphids. Finally, the intricate association between aphids and their hosts has long been speculated to be mediated by the aphid salivary secretions. Recently it has been shown that aphid salivary proteins can alter plant responses and negatively or positively affect aphid performance. The Forth and last objective of my research was to sequence the potato aphid salivary gland transcriptome, identify putative secreted proteins and functionally characterize the role of a selected few proteins on aphid fecundity.

References

Ahuja, I., Kissen, R., and Bones, A. M. 2012. Phytoalexins in defense against pathogens. Trends Plant Sci. 17:73-90.

Aist, J. 1979. Papillae and related wound plugs of plant cells. Annu. Rev. Phytopathol. 14:145-163.

Albrecht, C., Russinova, E., Hecht, V., Baaijens, E., and de Vries, S. 2005. The *Arabidopsis thaliana* SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis. Plant Cell 17:3337-3349.

Alonso, J. M., and Ecker, J. R. 2006. Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. Nat. Rev. Genet. 7:524-536.

Anthony, F., Topart, P., Martinez, A., Silva, M., and Nicole, M. 2005. Hypersensitive-like reaction conferred by the *Mex-1* resistance gene against *Meloidogyne exigua* in coffee. Plant Pathol. 54:476-482.

Atamian, H. S., Eulgem, T., and Kaloshian, I. 2012. *SlWRKY70* is required for *Mi-1* mediated resistance to aphids and nematodes in tomato. Planta 235:299-309.

Babu, M. M., Iyer, L. M., Balaji, S., and Aravind, L. 2006. The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. Nucleic Acids Res. 34:6505-6520.

Bai, Y., and Lindhout, P. 2007. Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? Ann. Bot. 100:1085-1094.

Bai, Y., van der Hulst, R., Bonnema, G., Marcel, T. C., Meijer-Dekens, F., Niks, R. E., and Lindhout, P. 2005. Tomato defense to *Oidium neolycopersici*: dominant *Ol* genes confer isolate-dependent resistance via a different mechanism than recessive *ol-2*. Mol. Plant-Microbe Interact. 18:354-362.

Balaji, V., Gibly, A., Debbie, P., and Sessa, G. 2007. Transcriptional analysis of the tomato resistance response triggered by recognition of the *Xanthomonas* type III effector AvrXv3. Funct. Integr. Genomics 7:305-316.

Bari, R., and Jones, J. D. 2009. Role of plant hormones in plant defence responses. Plant Mol. Biol. 69:473-488.

Barone, A., Chiusano, M. L., Ercolano, M. R., Giuliano, G., Grandillo, S., and Frusciante, L. 2008. Structural and functional genomics of tomato. Int. J. Plant Genomics 2008:820274.

Baumann, L., and Baumann, P. 1995. Soluble salivary proteins secreted by *Schizaphis graminum*. Entomol. Exp. Appl. 77:56-60.

Baumann, P. 2005. Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annu. Rev. Microbiol. 59:155-189.

Beckman, C. H. 2000. Phenolic-storing cells: keys to programmed cell death and periderm formation in wilt disease resistance and in general defence responses in plants. Physiol. Mol. Plant Pathol. 57:101-110.

Bednarek, P., and Osbourn, A. 2009. Plant-microbe interactions: chemical diversity in plant defense. Science 324:746-748.

Benedito, V. A., Visser, P. B., Angenent, G. C., and Krens, F. A. 2004. The potential of virus-induced gene silencing for speeding up functional characterization of plant genes. Genet. Mol. Res. 3:323-341.

Berger, S., Sinha, A. K., and Roitsch, T. 2007. Plant physiology meets phytopathology: plant primary metabolism and plant-pathogen interactions. J. Exp. Bot. 58:4019-4026.

Bhattarai, K. K., Atamian, H. S., Kaloshian, I., and Eulgem, T. 2010. WRKY72-type transcription factors contribute to basal immunity in tomato and Arabidopsis as well as gene-for-gene resistance mediated by the tomato *R*-gene *Mi-1*. Plant J. 63:229-240.

Bhattarai, K. K., Li, Q., Liu, Y., Dinesh-Kumar, S. P., and Kaloshian, I. 2007. The *Mi-1*-mediated pest resistance requires *Hsp90* and *Sgt1*. Plant Physiol. 144:312-323.

Bhattarai, K. K., Xie, Q. G., Mantelin, S., Bishnoi, U., Girke, T., Navarre, D. A., and Kaloshian, I. 2008. Tomato susceptibility to root-knot nematodes requires an intact jasmonic acid signaling pathway. Mol. Plant-Microbe Interact. 21:1205-1214.

Blackman, R. L., and Eastop, V. F. 2000. Aphids on the World's Crops. John Wiley & Sons, Ltd, New York.

Blancard, D. 2012. Tomato diseases. Academic Press, London.

Bos, J. I., Prince, D., Pitino, M., Maffei, M. E., Win, J., and Hogenhout, S. A. 2010. A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). PLoS Genet. 6:e1001216.

Bouche, N., Yellin, A., Snedden, W. A., and Fromm, H. 2005. Plant-specific calmodulin-binding proteins. Annu. Rev. Plant Biol. 56:435-466.

Boutros, M., Agaisse, H., and Perrimon, N. 2002. Sequential activation of signaling pathways during innate immune responses in Drosophila. Dev. Cell 3:711-722.

Brading, P. A., Hammond-Kosack, K. E., Parr, A., and Jones, J. D. 2000. Salicylic acid is not required for *Cf-2*- and *Cf-9*-dependent resistance of tomato to *Cladosporium fulvum*. Plant J. 23:305-318.

Braendle, C., Davis, G. K., Brisson, J. A., and Stern, D. L. 2006. Wing dimorphism in aphids. Heredity 97:192-199.

Branch, C., Hwang, C. F., Navarre, D. A., and Williamson, V. M. 2004. Salicylic acid is part of the *Mi-1*-mediated defense response to root-knot nematode in tomato. Mol. Plant-Microbe Interact. 17:351-356.

Brigneti, G., Martin-Hernandez, A. M., Jin, H., Chen, J., Baulcombe, D. C., Baker, B., and Jones, J. D. 2004. Virus-induced gene silencing in Solanum species. Plant J. 39:264-272.

Brisson, J. A. 2010. Aphid wing dimorphisms: linking environmental and genetic control of trait variation. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365:605-616.

Buchner, P. 1965. Endosymbiosis of animals with plant microorganisms. John Wiley, New York.

Burch-Smith, T. M., Anderson, J. C., Martin, G. B., and Dinesh-Kumar, S. P. 2004. Applications and advantages of virus-induced gene silencing for gene function studies in plants. Plant J. 39:734-746.

Burch-Smith, T. M., Schiff, M., Liu, Y., and Dinesh-Kumar, S. P. 2006. Efficient virus-induced gene silencing in Arabidopsis. Plant Physiol. 142:21-27.

Cai, M., Qiu, D., Yuan, T., Ding, X., Li, H., Duan, L., Xu, C., Li, X., and Wang, S. 2008. Identification of novel pathogen-responsive cis-elements and their binding proteins in the promoter of *OsWRKY13*, a gene regulating rice disease resistance. Plant Cell Environ. 31:86-96.

Caillaud, C. M., and Niemeyer, H. M. 1996. Possible involvement of the phloem sealing system in the acceptance of a plant as host by an aphid. Cell Mol. Life Sci. 52:927-931.

Caplan, J., Padmanabhan, M., and Dinesh-Kumar, S. P. 2008. Plant NB-LRR immune receptors: from recognition to transcriptional reprogramming. Cell Host Microbe 3:126-135.

Carolan, J. C., Fitzroy, C. I., Ashton, P. D., Douglas, A. E., and Wilkinson, T. L. 2009. The secreted salivary proteome of the pea aphid *Acyrthosiphon pisum* characterised by mass spectrometry. Proteomics 9:2457-2467.

Carolan, J. C., Caragea, D., Reardon, K. T., Mutti, N. S., Dittmer, N., Pappan, K., Cui, F., Castaneto, M., Poulain, J., Dossat, C., Tagu, D., Reese, J. C., Reeck, G. R., Wilkinson, T. L., and Edwards, O. R. 2011. Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrthosiphon pisum*): a dual transcriptomic/proteomic approach. J. Proteome Res. 10:1505-1518.

Casteel, C. L., Walling, L. L., and Paine, T. D. 2006. Behavior and biology of the tomato psyllid, *Bactericerca cockerelli*, in response to the *Mi-1.2* gene. Entomol. Exp. Appl. 121:67-72.

Chakravarthy, S., Velasquez, A. C., Ekengren, S. K., Collmer, A., and Martin, G. B. 2010. Identification of *Nicotiana benthamiana* genes involved in pathogen-associated molecular pattern-triggered immunity. Mol. Plant-Microbe Interact. 23:715-726.

Chen, J. C., Jiang, C. Z., and Reid, M. S. 2005. Silencing a prohibitin alters plant development and senescence. Plant J. 44:16-24.

Chen, X., and Ronald, P. C. 2011. Innate immunity in rice. Trends Plant Sci. 16:451-459.

Chen, Y. Y., Lin, Y. M., Chao, T. C., Wang, J. F., Liu, A. C., Ho, F. I., and Cheng, C. P. 2009. Virus-induced gene silencing reveals the involvement of ethylene-, salicylic acid- and mitogen-activated protein kinase-related defense pathways in the resistance of tomato to bacterial wilt. Physiol. Plant. 136:324-335.

Cheng, Y. T., Li, Y., Huang, S., Huang, Y., Dong, X., Zhang, Y., and Li, X. 2011. Stability of plant immune-receptor resistance proteins is controlled by SKP1-CULLIN1-F-box (SCF)-mediated protein degradation. Proc. Natl. Acad. Sci. U.S.A. 108:14694-14699.

Cherqui, A., and Tjallingii, W. F. 2000. Salivary proteins of aphids, a pilot study on identification, separation and immunolocalisation. J. Insect Physiol. 46:1177-1186.

Chini, A., Fonseca, S., Fernandez, G., Adie, B., Chico, J. M., Lorenzo, O., Garcia-Casado, G., Lopez-Vidriero, I., Lozano, F. M., Ponce, M. R., Micol, J. L., and Solano, R. 2007. The JAZ family of repressors is the missing link in jasmonate signalling. Nature 448:666-671.

Chung, E., Seong, E., Kim, Y. C., Chung, E. J., Oh, S. K., Lee, S., Park, J. M., Joung, Y. H., and Choi, D. 2004. A method of high frequency virus-induced gene silencing in chili pepper (*Capsicum annuum* L. cv. Bukang). Mol. Cells 17:377-380.

Ciolkowski, I., Wanke, D., Birkenbihl, R. P., and Somssich, I. E. 2008. Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. Plant Mol. Biol. 68:81-92.

Coll, N. S., Epple, P., and Dangl, J. L. 2011. Programmed cell death in the plant immune system. Cell Death Differ. 18:1247-1256.

Consales, F., Schweizer, F., Erb, M., Gouhier-Darimont, C., Bodenhausen, N., Bruessow, F., Sobhy, I., and Reymond, P. 2011. Insect oral secretions suppress wound-induced responses in Arabidopsis. J. Exp. Bot. 63:727-737.

Dangl, J. L., and Jones, J. D. G. 2001. Plant pathogens and integrated defence responses to infection. Nature 411:826-833.

Das, S., DeMason, D. A., Ehlers, J. D., Close, T. J., and Roberts, P. A. 2008. Histological characterization of root-knot nematode resistance in cowpea and its relation to reactive oxygen species modulation. J. Exp. Bot. 59:1305-1313.

David-Schwartz, R., Badani, H., Smadar, W., Levy, A. A., Galili, G., and Kapulnik, Y. 2001. Identification of a novel genetically controlled step in mycorrhizal colonization: plant resistance to infection by fungal spores but not extra-radical hyphae. Plant J. 27:561-569.

De Vos, M., and Jander, G. 2009. *Myzus persicae* (green peach aphid) salivary components induce defence responses in *Arabidopsis thaliana*. Plant Cell Environ. 32:1548-1560.

Deslandes, L., and Rivas, S. 2012. Catch me if you can: bacterial effectors and plant targets. Trends Plant Sci. doi:10.1016/j.tplants.2012.06.011.

Dixon, M., Hatzixanthis, K., Jones, D. A., Harrison, K., and Jones, J. D. G. 1998. The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. Plant Cell 10:1915-1925.

Dixon, M. S., Jones, D. A., Keddie, J. S., Thomas, C. M., Harrison, K., and Jones, J. D. G. 1996. The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 84:451-460.

Dodds, P. N., and Rathjen, J. P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. Nat. Rev. Genet. 11:539-548.

Dong, X. 2004. NPR1, all things considered. Curr. Opin. Plant Biol. 7:547-552.

Dropkin, V. H. 1969. Cellular responses of plants to nematode infections. Annu. Rev. Phytopathol. 7:101-122.

Du, X., Miao, M., Ma, X., Liu, Y., Kuhl, J. C., Martin, G. B., and Xiao, F. 2012. Plant programmed cell death caused by an autoactive form of Prf is suppressed by co-expression of the Prf LRR domain. doi: 10.1093/mp/sss014.

Durrant, W. E., and Dong, X. 2004. Systemic acquired resistance. Annu. Rev. Phytopathol. 42:185-209.

Eitas, T. K., and Dangl, J. L. 2010. NB-LRR proteins: pairs, pieces, perception, partners, and pathways. Curr. Opin. Plant Biol. 13:472-477.

Ekengren, S. K., Liu, Y., Schiff, M., Dinesh-Kumar, S. P., and Martin, G. B. 2003. Two MAPK cascades, NPR1, and TGA transcription factors play a role in Pto-mediated disease resistance in tomato. Plant J. 36:905-917.

Ellis, C., Karafyllidis, I., and Turner, J. G. 2002. Constitutive activation of jasmonate signaling in an Arabidopsis mutant correlates with enhanced resistance to *Erysiphe cichoracearum*, *Pseudomonas syringae*, and *Myzus persicae*. Mol. Plant-Microbe Interact. 15:1025-1030.

Emmanuel, E., and Levy, A. A. 2002. Tomato mutants as tools for functional genomics. Curr. Opin. Plant Biol. 5:112-117.

Encinas-Villarejo, S., Maldonado, A. M., Amil-Ruiz, F., de los Santos, B., Romero, F., Pliego-Alfaro, F., Munoz-Blanco, J., and Caballero, J. L. 2009. Evidence for a positive regulatory role of strawberry (*Fragaria x ananassa*) *Fa*WRKY1 and Arabidopsis *At*WRKY75 proteins in resistance. J. Exp. Bot. 60:3043-3065.

Eulgem, T. 2005. Regulation of the Arabidopsis defense transcriptome. Trends Plant Sci. 10:71-78.

Eulgem, T., and Somssich, I. E. 2007. Networks of WRKY transcription factors in defense signaling. Curr. Opin. Plant Biol. 10:366-371.

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. 2000. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 5:199-206.

Erb, M., Meldau, S., and Howe, G. A. 2012. Role of phytohormones in insect-specific plant reactions. Trends Plant Sci. 17:250-259.

Feng, J. X., Liu, D., Pan, Y., Gong, W., Ma, L. G., Luo, J. C., Deng, X. W., and Zhu, Y. X. 2005. An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the Arabidopsis AP2/EREBP transcription factor gene family. Plant Mol. Biol. 59:853-868.

Ferrari, J., Scarborough, C. L., and Godfray, H. C. 2007. Genetic variation in the effect of a facultative symbiont on host-plant use by pea aphids. Oecologia 153:323-329.

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

Fradin, E. F., Abd-El-Haliem, A., Masini, L., van den Berg, G. C., Joosten, M. H., and Thomma, B. P. 2011. Interfamily transfer of tomato *Ve1* mediates Verticillium resistance in Arabidopsis. Plant Physiol. 156:2255-2265.

Fradin, E. F., Zhang, Z., Juarez Ayala, J. C., Castroverde, C. D., Nazar, R. N., Robb, J., Liu, C. M., and Thomma, B. P. 2009. Genetic dissection of *Verticillium wilt* resistance mediated by tomato *Ve1*. Plant Physiol. 150:320-332.

Fu, D. Q., Ghabrial, S., and Kachroo, A. 2009. *GmRAR1* and *GmSGT1* are required for basal, *R* gene-mediated and systemic acquired resistance in soybean. Mol. Plant-Microbe Interact. 22:86-95.

Fu, Z. Q., Yan, S., Saleh, A., Wang, W., Ruble, J., Oka, N., Mohan, R., Spoel, S. H., Tada, Y., Zheng, N., and Dong, X. 2012. NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. Nature 486:228-232.

Gabriels, S. H., Vossen, J. H., Ekengren, S. K., van Ooijen, G., Abd-El-Haliem, A. M., van den Berg, G. C., Rainey, D. Y., Martin, G. B., Takken, F. L., de Wit, P. J., and Joosten, M. H. 2007. An NB-LRR protein required for HR signalling mediated by both extra- and intracellular resistance proteins. Plant J. 50:14-28.

Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S., Ward, E., and Ryals, J. 1993. Requirement for salicylic acid for the induction of systemic acquired resistance. Science 261:754-756.

Gallou, A., Declerck, S., and Cranenbrouck, S. 2012. Transcriptional regulation of defence genes and involvement of the WRKY transcription factor in arbuscular mycorrhizal potato root colonization. Funct. Integr. Genomics 12:183-198.

Gao, L. L., Anderson, J. P., Klingler, J. P., Nair, R. M., Edwards, O. R., and Singh, K. B. 2007. Involvement of the octadecanoid pathway in bluegreen aphid resistance in *Medicago truncatula*. Mol. Plant-Microbe Interact. 20:82-93.

Gaunt, M. W., and Miles, M. A. 2002. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Mol. Biol. Evol. 19:748-761.

Gerardo, N. M., Altincicek, B., Anselme, C., Atamian, H., Barribeau, S. M., de Vos, M., Duncan, E. J., Evans, J. D., Gabaldon, T., Ghanim, M., Heddi, A., Kaloshian, I., Latorre, A., Moya, A., Nakabachi, A., Parker, B. J., Perez-Brocal, V., Pignatelli, M., Rahbe, Y., Ramsey, J. S., Spragg, C. J., Tamames, J., Tamarit, D., Tamborindeguy, C., Vincent-Monegat, C., and Vilcinskas, A. 2010. Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome Biol. 11:R21.

Giordanengo, P., Brunissen, L., Rusterucci, C., Vincent, C., van Bel, A., Dinant, S., Girousse, C., Faucher, M., and Bonnemain, J. L. 2010. Compatible plant-aphid interactions: how aphids manipulate plant responses. C. R. Biol. 333:516-523.

Glazebrook, J. 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu. Rev. Phytopathol. 43:205-227.

Glockner, G., Eichinger, L., Szafranski, K., Pachebat, J. A., Bankier, A. T., Dear, P. H., Lehmann, R., Baumgart, C., Parra, G., Abril, J. F., Guigo, R., Kumpf, K., Tunggal, B., Cox, E., Quail, M. A., Platzer, M., Rosenthal, A., and Noegel, A. A. 2002. Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. Nature 418:79-85.

Goggin, F. L., Jia, L., Shah, G., Hebert, S., Williamson, V. M., and Ullman, D. E. 2006. Heterologous expression of the *Mi-1.2* gene from tomato confers resistance against nematodes but not aphids in eggplant. Mol. Plant-Microbe Interact. 19:383-388.

Goritschnig, S., Zhang, Y., and Li, X. 2007. The ubiquitin pathway is required for innate immunity in Arabidopsis. Plant J. 49:540-551.

Gould, B., and Kramer, E. M. 2007. Virus-induced gene silencing as a tool for functional analyses in the emerging model plant *Aquilegia* (columbine, Ranunculaceae). Plant Methods 3:6.

Greeff, C., Roux, M., Mundy, J., and Petersen, M. 2012. Receptor-like kinase complexes in plant innate immunity. Front. Plant Sci. 3:209

Guo, R., Yu, F., Gao, Z., An, H., Cao, X., and Guo, X. 2011. *GhWRKY3*, a novel cotton (*Gossypium hirsutum* L.) *WRKY* gene, is involved in diverse stress responses. Mol. Biol. Rep. 38:49-58.

Gutterson, N., and Reuber, T. L. 2004. Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr. Opin. Plant Biol. 7:465-471.

Halitschke, R., Schittko, U., Pohnert, G., Boland, W., and Baldwin, I. T. 2001. Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. III. Fatty acid-amino acid conjugates in herbivore oral secretions are necessary and sufficient for herbivore-specific plant responses. Plant Physiol. 125:711-717.

Hammond-Kosak, K. E., and Jones, J. D. G. 1996. Resistance-gene dependent plant defense mechanisms. Plant Cell 8:1773-1791.

Harmel, N., Letocart, E., Cherqui, A., Giordanengo, P., Mazzucchelli, G., Guillonneau, F., De Pauw, E., Haubruge, E., and Francis, F. 2008. Identification of aphid salivary proteins: a proteomic investigation of *Myzus persicae*. Insect Mol. Biol. 17:165-174.

Harrington, R., and Van Emden, H. F. 2007. Aphids as Crop Pests. CABI, Wallingford.

He, H., Dong, Q., Shao, Y., Jiang, H., Zhu, S., Cheng, B., and Xiang, Y. 2012. Genome-wide survey and characterization of the *WRKY* gene family in *Populus trichocarpa*. Plant Cell Rep. 7:1199-1217.

Heese, A., Hann, D. R., Gimenez-Ibanez, S., Jones, A. M., He, K., Li, J., Schroeder, J. I., Peck, S. C., and Rathjen, J. P. 2007. The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. Proc. Natl. Acad. Sci. U.S.A. 104:12217-12222.

Hematy, K., Cherk, C., and Somerville, S. 2009. Host-pathogen warfare at the plant cell wall. Curr. Opin. Plant Biol. 12:406-413.

Hewer, A., Will, T., and van Bel, A. J. 2010. Plant cues for aphid navigation in vascular tissues. J. Exp. Biol. 213:4030-4042.

Hileman, L. C., Drea, S., Martino, G., Litt, A., and Irish, V. F. 2005. Virus-induced gene silencing is an effective tool for assaying gene function in the basal eudicot species *Papaver somniferum* (opium poppy). Plant J. 44:334-341.

Hogenhout, S. A., and Bos, J. I. 2011. Effector proteins that modulate plant-insect interactions. Curr. Opin. Plant Biol. 14:422-428.

Holtmann, B., Kleine, M., and Grundler, F. M. W. 2000. Ultrastructure and anatomy of nematode-induced syncytia in roots of susceptible and resistant sugar beet. Protoplasma 211:39-50.

Hu, G., deHart, A. K., Li, Y., Ustach, C., Handley, V., Navarre, R., Hwang, C. F., Aegerter, B. J., Williamson, V. M., and Baker, B. 2005. *EDS1* in tomato is required for resistance mediated by TIR-class R genes and the receptor-like *R* gene *Ve*. Plant J. 42:376-391.

Huang, S., Gao, Y., Liu, J., Peng, X., Niu, X., Fei, Z., Cao, S., and Liu, Y. 2012. Genome-wide analysis of WRKY transcription factors in *Solanum lycopersicum*. Mol. Genet. Genomics. 6:495-513.

Huybrechts, J., Bonhomme, J., Minoli, S., Prunier-Leterme, N., Dombrovsky, A., Abdel-Latief, M., Robichon, A., Veenstra, J. A., and Tagu, D. 2010. Neuropeptide and neurohormone precursors in the pea aphid, *Acyrthosiphon pisum*. Insect Mol. Biol. 19:87-95.

Hwang, S. H., Yie, S. W., and Hwang, D. J. 2011. Heterologous expression of *OsWRKY6* gene in Arabidopsis activates the expression of defense related genes and enhances resistance to pathogens. Plant Sci. 181:316-323.

International Aphid Genomic Consortium. 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biol. 8:e1000313.

Ishibashi, K., Naitob, S., Meshia, T., and Ishikawaa, M. 2009. An inhibitory interaction between viral and cellular proteins underlies the resistance of tomato to nonadapted tobamoviruses. Proc. Natl. Acad. Sci. U.S.A. 106:8778-8783.

Ishibashi, K., Masuda, K., Naito, S., Meshi, T., and Ishikawa, M. 2007. An inhibitor of viral RNA replication is encoded by a plant resistance gene. Proc. Natl. Acad. Sci. U.S.A. 104:13833-13838.

Jiang, Y., and Deyholos, M. K. 2009. Functional characterization of Arabidopsis NaCl-inducible *WRKY25* and *WRKY33* transcription factors in abiotic stresses. Plant Mol. Biol. 69:91-105.

Jing, S., Zhou, X., Song, Y., and Yu, D. 2009. Heterologous expression of *OsWRKY23* gene enhances pathogen defense and dark-induced leaf senescence in Arabidopsis. Plant Growth Regul. 58:181–190.

Jones, D. A., Thomas, C. M., Hammond-Kosack, K. E., Balint-Kurti, P. J., and Jones, J. D. G. 1994. Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science 266:789-793.

Jones, J. D., and Dangl, J. L. 2006. The plant immune system. Nature 444:323-329.

Kadota, Y., Shirasu, K., and Guerois, R. 2010. NLR sensors meet at the SGT1-HSP90 crossroad. Trends Biochem. Sci. 35:199-207.

Kaloshian, I., Lange, W. H., and Williamson, V. M. 1995. An aphid-resistance locus is tightly linked to the nematode-resistance gene, *Mi*, in tomato. Proc. Natl. Acad. Sci. U.S.A. 92:622-625.

Kaloshian, I., Kinsey, M. G., Ullman, D. E., and Wiliamson, V. M. 1997. The impact of *Meu1*-mediated resistance in tomato on longevity, fecundity and behavior of the potato aphid, *Macrosiphum euphorbiae*. Entomol. Exp. Appl. 83:181–187.

Kaloshian, I., Kinsey, M. G., Williamson, V. M., and Ullman, D. E. 2000. *Mi*-mediated resistance against the potato aphid *Macrosiphum euphorbiae* (Hemiptera: Aphididae) limits sieve element ingestion. Environ. Entomol. 29:690–695.

Knepper, C., Savory, E. A., and Day, B. 2011. Arabidopsis NDR1 is an integrin-like protein with a role in fluid loss and plasma membrane-cell wall adhesion. Plant Physiol. 156:286-300.

Knoth, C., Ringler, J., Dangl, J. L., and Eulgem, T. 2007. Arabidopsis *WRKY70* is required for full *RPP4*-mediated disease resistance and basal defense against *Hyaloperonospora parasitica*. Mol. Plant-Microbe Interact. 20:120-128.

Koornneef, A., and Pieterse, C. M. 2008. Cross talk in defense signaling. Plant Physiol. 146:839-844.

Kumagai, M. H., Donson, J., della-Cioppa, G., Harvey, D., Hanley, K., and Grill, L. K. 1995. Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. Proc. Natl. Acad. Sci. U.S.A. 92:1679-1683.

Kurucz, E., Markus, R., Zsamboki, J., Folkl-Medzihradszky, K., Darula, Z., Vilmos, P., Udvardy, A., Krausz, I., Lukacsovich, T., Gateff, E., Zettervall, C. J., Hultmark, D., and Ando, I. 2007. Nimrod, a putative phagocytosis receptor with EGF repeats in Drosophila plasmatocytes. Curr. Biol. 17:649-654.

Lai, A., Cianciolo, V., Chiavarinin, S., and Sonninol, A. 2000. Effects of glandular trichomes on the development of *Phytophthora infestans* infection in potato (*S. tuberosum*). Euphytica 114:165-174.

Lazzaro, B. P. 2005. Elevated polymorphism and divergence in the class C scavenger receptors of *Drosophila melanogaster* and *D. simulans*. Genetics 169:2023-2034.

Leonardo, T. E., and Muiru, G. T. 2003. Facultative symbionts are associated with host plant specialization in pea aphid populations. Proc. R. Soc. Lond. B Biol. Sci. 270:209-212.

Levee, V., Major, I., Levasseur, C., Tremblay, L., MacKay, J., and Seguin, A. 2009. Expression profiling and functional analysis of *Populus WRKY23* reveals a regulatory role in defense. New Phytol. 184:48-70.

Levetin, E., and McMahon, K. 2012. Plants and Society. McGraw-Hill, London.

Lewis, M. W., Leslie, M. E., Fulcher, E. H., Darnielle, L., Healy, P. N., Youn, J. Y., and Liljegren, S. J. 2010. The SERK1 receptor-like kinase regulates organ separation in Arabidopsis flowers. Plant J. 62:817-828.

Li, C., Bai, Y., Jacobsen, E., Visser, R., Lindhout, P., and Bonnema, G. 2006a. Tomato defense to the powdery mildew fungus: differences in expression of genes in susceptible, monogenic- and polygenic resistance responses are mainly in timing. Plant Mol. Biol. 62:127-140.

Li, C., Bonnema, G., Che, D., Dong, L., Lindhout, P., Visser, R., and Bai, Y. 2007. Biochemical and molecular mechanisms involved in monogenic resistance responses to tomato powdery mildew. Mol. Plant-Microbe Interact. 20:1161-1172.

Li, H., Xu, Y., Xiao, Y., Zhu, Z., Xie, X., Zhao, H., and Wang, Y. 2010. Expression and functional analysis of two genes encoding transcription factors, *VpWRKY1* and *VpWRKY2*, isolated from Chinese wild *Vitis pseudoreticulata*. Planta 232:1325-1337.

Li, H. L., Zhang, L. B., Guo, D., Li, C. Z., and Peng, S. Q. 2012. Identification and expression profiles of the WRKY transcription factor family in *Ricinus communis*. Gene 503:248-253.

Li, Q., Xie, Q. G., Smith-Becker, J., Navarre, D. A., and Kaloshian, I. 2006b. *Mi-1*-mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling cascades. Mol. Plant-Microbe Interact. 19:655-664.

Li, X., Clarke, J. D., Zhang, Y., and Dong, X. 2001. Activation of an EDS1-mediated *R*-gene pathway in the *snc1* mutant leads to constitutive, NPR1-independent pathogen resistance. Mol. Plant-Microbe Interact. 14:1131-1139.

Lim, J. H., Park, C. J., Huh, S. U., Choi, L. M., Lee, G. J., Kim, Y. J., and Paek, K. H. 2011. *Capsicum annuum* WRKYb transcription factor that binds to the *CaPR-10* promoter functions as a positive regulator in innate immunity upon TMV infection. Biochem. Biophys. Res. Commun. 411:613-619.

Ling, J., Jiang, W., Zhang, Y., Yu, H., Mao, Z., Gu, X., Huang, S., and Xie, B. 2011. Genome-wide analysis of WRKY gene family in *Cucumis sativus*. BMC Genomics 12:471.

Liu, E., and Page, J. E. 2008. Optimized cDNA libraries for virus-induced gene silencing (VIGS) using tobacco rattle virus. Plant Methods 4:5.

Liu, H., Yang, W., Liu, D., Han, Y., Zhang, A., and Li, S. 2010. Ectopic expression of a grapevine transcription factor *VvWRKY11* contributes to osmotic stress tolerance in Arabidopsis. Mol. Biol. Rep. 38:417-427.

Liu, H., Fu, D., Zhu, B., Yan, H., Shen, X., Zuo, J., Zhu, Y., and Luo, Y. 2012. Virus-induced Gene Silencing in Eggplant (*Solanum melongena*). J. Integr. Plant Biol. 54: 422-429.

Liu, J., Liu, X., Dai, L., and Wang, G. 2007. Recent progress in elucidating the structure, function and evolution of disease resistance genes in plants. J. Genet. Genomics 34:765-776.

Liu, J. J., and Ekramoddoullah, A. K. 2009. Identification and characterization of the WRKY transcription factor family in *Pinus monticola*. Genome 52:77-88.

Liu, Y., Schiff, M., and Dinesh-Kumar, S. P. 2002. Virus-induced gene silencing in tomato. Plant J. 31:777-786.

Liu, Y., Nakayama, N., Schiff, M., Litt, A., Irish, V. F., and Dinesh-Kumar, S. P. 2004. Virus induced gene silencing of a DEFICIENS ortholog in *Nicotiana benthamiana*. Plant Mol. Biol. 54:701-711.

Lombaert, E., Carletto, J., Piotte, C., Fauvergue, X., Lecoq, H., Vanlerberghe-Masutti, F., and Lapchin, L. 2009. Response of the melon aphid, *Aphis gossypii*, to host-plant resistance: evidence for high adaptive potential despite low genetic variability. Entomol. Exp. Appl. 133:46-56.

Lozano-Torres, J. L., Wilbers, R. H., Gawronski, P., Boshoven, J. C., Finkers-Tomczak, A., Cordewener, J. H., America, A. H., Overmars, H. A., Van 't Klooster, J. W., Baranowski, L., Sobczak, M., Ilyas, M., van der Hoorn, R. A., Schots, A., de Wit, P. J., Bakker, J., Goverse, A., and Smant, G. 2012. Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. Proc. Natl. Acad. Sci. U.S.A. 109:10119-10124.

Lu, R., Martin-Hernandez, A. M., Peart, J. R., Malcuit, I., and Baulcombe, D. C. 2003a. Virus-induced gene silencing in plants. Methods 30:296-303.

Lu, R., Malcuit, I., Moffett, P., Ruiz, M. T., Peart, J., Wu, A. J., Rathjen, J. P., Bendahmane, A., Day, L., and Baulcombe, D. C. 2003b. High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. EMBO J. 22:5690-5699.

Lukasik-Shreepaathy, E., Slootweg, E., Richter, H., Goverse, A., Cornelissen, B. J., and Takken, F. L. 2012. Dual regulatory roles of the extended N terminus for activation of the tomato Mi-1.2 resistance protein. Mol. Plant-Microbe Interact. 25:1045-1057.

Luo, M., Dennis, E. S., Berger, F., Peacock, W. J., and Chaudhury, A. 2005. *MINISEED3* (*MINI3*), a *WRKY* family gene, and *HAIKU2* (*IKU2*), a *leucine-rich repeat* (*LRR*) *KINASE* gene, are regulators of seed size in Arabidopsis. Proc. Natl. Acad. Sci. U.S.A. 102:17531-17536.

Ma, D., Pu, G., Lei, C., Ma, L., Wang, H., Guo, Y., Chen, J., Du, Z., Wang, H., Li, G., Ye, H., and Liu, B. 2009. Isolation and characterization of *Aa*WRKY1, an *Artemisia annua* transcription factor that regulates the amorpha-4,11-diene synthase gene, a key gene of artemisinin biosynthesis. Plant Cell Physiol. 50:2146-2161.

Madhusudhan, V. V., and Miles, P. W. 1998. Mobility of salivary components as a possible reason for differences in response of alfalfa to the spotted alfalfa aphid and pea aphid. Entomol. Exp. Appl. 86:25-39.

Major, I. T., and Constabel, C. P. 2007. Insect regurgitant and wounding elicit similar defense responses in poplar leaves: not something to spit at? Plant Signal. Behav. 2:1-3.

Mangelsen, E., Kilian, J., Berendzen, K. W., Kolukisaoglu, U. H., Harter, K., Jansson, C., and Wanke, D. 2008. Phylogenetic and comparative gene expression analysis of barley (*Hordeum vulgare*) WRKY transcription factor family reveals putatively retained functions between monocots and dicots. BMC Genomics 9:194.

Mantelin, S., Peng, H. C., Li, B., Atamian, H. S., Takken, F. L., and Kaloshian, I. 2011. The receptor-like kinase *SlSERK1* is required for *Mi-1*-mediated resistance to potato aphids in tomato. Plant J. 67:459-471.

Marchive, C., Mzid, R., Deluc, L., Barrieu, F., Pirrello, J., Gauthier, A., Corio-Costet, M. F., Regad, F., Cailleteau, B., Hamdi, S., and Lauvergeat, V. 2007. Isolation and characterization of a *Vitis vinifera* transcription factor, *Vv*WRKY1, and its effect on responses to fungal pathogens in transgenic tobacco plants. J. Exp. Bot. 58:1999-2010.

Martin, G., Bogdanove, A., and Sessa, G. 2003. Understanding the functions of plant disease resistance proteins. Annu. Rev. Plant Biol. 54:23-61.

Martinez de Ilarduya, O., and Kaloshian, I. 2001. *Mi-1.2* transcripts accumulate ubiquitously in root-knot nematode resistant *Lycopersicon esculentum*. J. Nematol. 33:116-120.

Martinez-de IIarduya, O., Moore, A. E., and Kaloshian, I. 2001. The tomato *Rme1* locus is required for *Mi*-mediated resistance to root knot nematodes and the potato aphid. Plant J. 27:417-425.

Martinez de Ilarduya, O., Xie, Q., and Kaloshian, I. 2003. Aphid-induced defense responses in *Mi-1*-mediated compatible and incompatible tomato interactions. Mol. Plant-Microbe Interact. 16:699-708.

Maston, G. A., Evans, S. K., and Green, M. R. 2006. Transcriptional regulatory elements in the human genome. Annu. Rev. Genomics Hum. Genet. 7:29-59.

Meissner, R., Chague, V., Zhu, Q., Emmanuel, E., Elkind, Y., and Levy, A. A. 2000. Technical advance: a high throughput system for transposon tagging and promoter trapping in tomato. Plant J. 22:265-274.

Mellilo, M. T., Leonetti, P., Bongiovanni, M., Castagnone-Sereno, P., and Bleve-Zacheo, T. 2006. Modulation of reactive oxygen species activity and H2O2 accumulation during compatible and incompatible tomato-root-knot nematode interactions. New Phytol. 170:501-512.

Melzer, R., and Theissen, G. 2011. MADS and more: transcription factors that shape the plant. Methods Mol. Biol. 754:3-18.

Miles, P. 1999. Aphid saliva. Biol. Rev. 74:41-85.

Milligan, S. B., Bodeau, J., Yaghoobi, J., Kaloshian, I., Zabel, P., and Williamson, V. M. 1998. The root knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. Plant Cell 10:1330-1332.

Mishra, N. S., Tuteja, R., and Tuteja, N. 2006. Signaling through MAP kinase networks in plants. Arch. Biochem. Biophys. 452:55-68.

Moore, J. W., Loake, G. J., and Spoel, S. H. 2011. Transcription dynamics in plant immunity. Plant Cell 23:2809-2820.

Morel, J. B., and Dangl, J. L. 1997. The hypersensitive response and the induction of cell death in plants. Cell Death Differ. 19:17-24.

Mur, L. A., Kenton, P., Atzorn, R., Miersch, O., and Wasternack, C. 2006. The outcomes of concentration-specific interactions between salicylate and jasmonate signaling include synergy, antagonism, and oxidative stress leading to cell death. Plant Physiol. 140:249-262.

Musser, R. O., Hum-Musser, S. M., Eichenseer, H., Peiffer, M., Ervin, G., Murphy, J. B., and Felton, G. W. 2002. Herbivory: caterpillar saliva beats plant defences. Nature 416:599-600.

Mutti, N. S., Louis, J., Pappan, L. K., Pappan, K., Begum, K., Chen, M. S., Park, Y., Dittmer, N., Marshall, J., Reese, J. C., and Reeck, G. R. 2008. A protein from the salivary glands of the pea aphid, *Acyrthosiphon pisum*, is essential in feeding on a host plant. Proc. Natl. Acad. Sci. U.S.A. 105:9965-9969.

Mysore, K. S., Crasta, O. R., Tuori, R. P., Folkerts, O., Swirsky, P. B., and Martin, G. B. 2002. Comprehensive transcript profiling of Pto- and Prf-mediated host defense responses to infection by *Pseudomonas syringae* pv. *tomato*. Plant J. 32:299-315.

Mzid, R., Marchive, C., Blancard, D., Deluc, L., Barrieu, F., Corio-Costet, M. F., Drira, N., Hamdi, S., and Lauvergeat, V. 2007. Overexpression of *Vv*WRKY2 in tobacco enhances broad resistance to necrotrophic fungal pathogens. Physiol. Plant. 131:434-447.

Nakabachi, A., and Miyagishima, S. 2010. Expansion of genes encoding a novel type of dynamin in the genome of the pea aphid, *Acyrthosiphon pisum*. Insect Mol. Biol. 19:165-173.

Nekrasov, V., Ludwig, A. A., and Jones, J. D. 2006. CITRX thioredoxin is a putative adaptor protein connecting Cf-9 and the ACIK1 protein kinase during the Cf-9/Avr9- induced defence response. FEBS Lett. 580:4236-4241.

Nombela, G., Williamson, V., M, and Muñiz, M. 2003. The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. Mol. Plant-Microbe Interact. 16:645-649.

Oldroyd, G. E. D., and Staskawicz, B. J. 1998. Genetically engineered broad-spectrum disease resistance in tomato. Proc. Natl. Acad. Sci. U.S.A. 95:10300-10305.

Oliva, R., Win, J., Raffaele, S., Boutemy, L., Bozkurt, T. O., Chaparro-Garcia, A., Segretin, M. E., Stam, R., Schornack, S., Cano, L. M., van Damme, M., Huitema, E., Thines, M., Banfield, M. J., and Kamoun, S. 2010. Recent developments in effector biology of filamentous plant pathogens. Cell. Microbiol. 12:705-715.

Oliver, J. P., Castro, A., Gaggero, C., Cascon, T., Schmelz, E. A., Castresana, C., and Ponce de Leon, I. 2009. Pythium infection activates conserved plant defense responses in mosses. Planta 230:569-579.

Oliver, K. M., Moran, N. A., and Hunter, M. S. 2005. Variation in resistance to parasitism in aphids is due to symbionts not host genotype. Proc. Natl. Acad. Sci. U.S.A. 102:12795-12800.

Orzaez, D., Medina, A., Torre, S., Fernandez-Moreno, J. P., Rambla, J. L., Fernandez-Del-Carmen, A., Butelli, E., Martin, C., and Granell, A. 2009. A visual reporter system for virus-induced gene silencing in tomato fruit based on anthocyanin accumulation. Plant Physiol. 150:1122-1134.

Pallipparambil, G. R., Reese, J. C., Avila, C. A., Louis, J. M., and Goggin, F. L. 2010. *Mi*-mediated aphid resistance in tomato: tissue localization and impact on the feeding behavior of two potato aphid clones with differing levels of virulence. Entomol. Exp. Appl. 135:295–307.

Pan, Y. J., Cho, C. C., Kao, Y. Y., and Sun, C. H. 2009. A novel WRKY-like protein involved in transcriptional activation of cyst wall protein genes in *Giardia lamblia*. J. Biol. Chem. 284:17975-17988.

Park, C. J., Shin, Y. C., Lee, B. J., Kim, K. J., Kim, J. K., and Paek, K. H. 2006. A hot pepper gene encoding WRKY transcription factor is induced during hypersensitive response to Tobacco mosaic virus and *Xanthomonas campestris*. Planta 223:168-179.

Paulson, R. E., and Webster, J. M. 1972. Ultrastructure of the hypersensitive reaction in roots of tomato, *Lycopersicon esculentum* L., to infection by the root-knot nematode, *Meloidogyne incognita*. Physiol. Plant Pathol. 2:227-234.

Pearl, L. H., and Prodromou, C. 2006. Structure and mechanism of the Hsp90 molecular chaperone machinery. Annu. Rev. Biochem. 75:271-294.

Peart, J. R., Mestre, P., Lu, R., Malcuit, I., and Baulcombe, D. C. 2005. NRG1, a CC-NB-LRR protein, together with N, a TIR-NB-LRR protein, mediates resistance against tobacco mosaic virus. Curr. Biol. 15:968-973.

Peart, J. R., Lu, R., Sadanandom, A., Malcuit, I., Moffett, P., Brice, D. C., Schauser, L., Jaggard, D. A., Xiao, S., Coleman, M. J., Dow, M., Jones, J. D., Shirasu, K., and Baulcombe, D. C. 2002. Ubiquitin ligase-associated protein SGT1 is required for host and nonhost disease resistance in plants. Proc. Natl. Acad. Sci. U.S.A. 99:10865-10869.

Pedley, K. F., and Martin, G. B. 2004. Identification of MAPKs and their possible MAPK kinase activators involved in the Pto-mediated defense response of tomato. J. Biol. Chem. 279:49229-49235.

Pegard, A., Brizzard, G., Fazari, A., Soucaze, O., Abad, P., and Djian-Caporalino, C. 2005. Histological characterization of resistance to different Root-knot nematode species related to phenolics accumulation in *Capsicum annuum*. Phytopathology 95:158-165.

Pitzschke, A., Schikora, A., and Hirt, H. 2009. MAPK cascade signalling networks in plant defence. Curr. Opin. Plant Biol. 12:421-426.

Powell, G., and Hardie, J. 2000. Host-selection behaviour by genetically identical aphids with different plant preferences. Physiol. Entomol. 25:54-62.

Prado, E., and Tjallingii, W. F. 2007. Behavioral evidence for local reduction of aphid-induced resistance. J. Insect. Sci. 7:1-8.

Preuss, S., and Pikaard, C. S. 2003. Targeted gene silencing in plants using RNA interference Pages 23-36 in: RNA Interference (RNAi): Nuts & Bolts of RNAi Technology, D. Engelke, ed. DNA Press, LLC.

Proietti, S., Bertini, L., Van der Ent, S., Leon-Reyes, A., Pieterse, C. M., Tucci, M., Caporale, C., and Caruso, C. 2011. Cross activity of orthologous WRKY transcription factors in wheat and Arabidopsis. J. Exp. Bot. 62:1975-1990.

Qu, L. J., and Zhu, Y. X. 2006. Transcription factor families in Arabidopsis: major progress and outstanding issues for future research. Curr. Opin. Plant Biol. 9:544-549.

Rairdan, G. J., and Delaney, T. P. 2002. Role of salicylic acid and NIM1/NPR1 in race-specific resistance in arabidopsis. Genetics 161:803-811.

Ramsey, J. S., Wilson, A. C., de Vos, M., Sun, Q., Tamborindeguy, C., Winfield, A., Malloch, G., Smith, D. M., Fenton, B., Gray, S. M., and Jander, G. 2007. Genomic resources for *Myzus persicae*: EST sequencing, SNP identification, and microarray design. BMC Genomics 8:423.

Ratcliff, F., Martin-Hernandez, A. M., and Baulcombe, D. C. 2001. Technical Advance. Tobacco rattle virus as a vector for analysis of gene function by silencing. Plant J. 25:237-245.

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C.-J., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J. Z., Ghandahari, D., Sherman, B. K., and Yu, G.-L. 2000. Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. Science 290:2105-2110.

Rivas, S., and Thomas, C. M. 2005. Molecular interactions between tomato and the leaf mold pathogen *Cladosporium fulvum*. Annu. Rev. Phytopathol. 43:395-436.

Rivas, S., Rougon-Cardoso, A., Smoker, M., Schauser, L., Yoshioka, H., and Jones, J. D. 2004. CITRX thioredoxin interacts with the tomato Cf-9 resistance protein and negatively regulates defence. EMBO J. 23:2156-2165.

Robatzek, S., and Somssich, I. E. 2002. Targets of *At*WRKY6 regulation during plant senescence and pathogen defense. Genes Dev. 16:1139-1149.

Robert-Seilaniantz, A., Grant, M., and Jones, J. D. 2011. Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. Annu. Rev. Phytopathol. 49:317-343.

Roberts, P. A., and Thomason, I. J. 1986. Variability in reproduction of isolates of *Meloidogyne incognita* and *M. javanica* on resistant tomato genotypes. Plant Disease 70:547-551.

Rodriguez, M. C., Petersen, M., and Mundy, J. 2010. Mitogen-activated protein kinase signaling in plants. Annu. Rev. Plant Biol. 61:621-649.

Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E., and Williamson, V. M. 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc. Natl. Acad. Sci. U.S.A. 95:9750-9754.

Rowland, O., Ludwig, A. A., Merrick, C. J., Baillieul, F., Tracy, F. E., Durrant, W. E., Fritz-Laylin, L., Nekrasov, V., Sjolander, K., Yoshioka, H., and Jones, J. D. 2005. Functional analysis of *Avr9*/*Cf-9* rapidly elicited genes identifies a protein kinase, ACIK1, that is essential for full Cf-9-dependent disease resistance in tomato. Plant Cell 17:295-310.

Ruiz, M. T., Voinnet, O., and Baulcombe, D. C. 1998. Initiation and maintenance of virus-induced gene silencing. Plant Cell 10:937-946.

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. 2010. WRKY transcription factors. Trends Plant Sci. 15:247-258.

Ryan, C. A., Huffaker, A., and Yamaguchi, Y. 2007. New insights into innate immunity in Arabidopsis. Cell. Microbiol. 9:1902-1908.

Sasaki, S., Yamagishi, N., and Yoshikawa, N. 2011. Efficient virus-induced gene silencing in apple, pear and Japanese pear using *Apple latent spherical virus* vectors. Plant Methods 7:15.

Scarborough, C. L., Ferrari, J., and Godfray, H. C. 2005. Aphid protected from pathogen by endosymbiont. Science 310:1781.

Scharte, J., Schön, H., and Weis, E. 2005. Photosynthesis and carbohydrate metabolism in tobacco leaves during an incompatible interaction with *Phytophthora nicotianae*. Plant, Cell and Environ. 28:1421-1435.

Schornack, S., Huitema, E., Cano, L. M., Bozkurt, T. O., Oliva, R., Van Damme, M., Schwizer, S., Raffaele, S., Chaparro-Garcia, A., Farrer, R., Segretin, M. E., Bos, J., Haas, B. J., Zody, M. C., Nusbaum, C., Win, J., Thines, M., and Kamoun, S. 2009. Ten things to know about oomycete effectors. Mol. Plant Pathol. 10:795-803.

Scofield, S. R., Huang, L., Brandt, A. S., and Gill, B. S. 2005. Development of a virus-induced gene-silencing system for hexaploid wheat and its use in functional analysis of the *Lr21*-mediated leaf rust resistance pathway. Plant Physiol. 138:2165-2173.

Shan, X., Yan, J., and Xie, D. 2012. Comparison of phytohormone signaling mechanisms. Curr. Opin. Plant Biol. 15:84-91.

Shekhawat, U. K., Ganapathi, T. R., and Srinivas, L. 2011. Cloning and characterization of a novel stress-responsive WRKY transcription factor gene (*MusaWRKY71*) from Musa spp. cv. Karibale Monthan (ABB group) using transformed banana cells. Mol. Biol. Rep. 38:4023-4035.

Shibata, D. 2005. Genome sequencing and functional genomics approaches in tomato. J. Gen. Plant Pathol. 71:1-7.

Shirano, Y., Kachroo, P., Shah, J., and Klessig, D. F. 2002. A gain-of-function mutation in an Arabidopsis Toll Interleukin1 receptor-nucleotide binding site-leucine-rich repeat type *R* gene triggers defense responses and results in enhanced disease resistance. Plant Cell 14:3149-3162.

Shirasu, K. 2009. The HSP90-SGT1 chaperone complex for NLR immune sensors. Annu. Rev. Plant Biol. 60:139-164.

Sim, S. C., Robbins, M. D., Chilcott, C., Zhu, T., and Francis, D. M. 2009. Oligonucleotide array discovery of polymorphisms in cultivated tomato (*Solanum lycopersicum* L.) reveals patterns of SNP variation associated with breeding. BMC Genomics 10:466.

Skibbe, M., Qu, N., Galis, I., and Baldwin, I. T. 2008. Induced plant defenses in the natural environment: *Nicotiana attenuata* WRKY3 and WRKY6 coordinate responses to herbivory. Plant Cell 20:1984-2000.

Smith, P. G. 1944. Embryo culture of a tomato species hybrid. Am. Soc. Hortic. Sci. 44:413-416.

Sobczak, M., Avrova, A., Jupowicz, J., Phillips, M. S., Ernst, K., and Kumar, A. 2005. Characterization of susceptibility and resistance responses to potato cyst nematode (Globodera spp.) infection of tomato lines in the absence and presence of the broad-spectrum nematode resistance *Hero* gene. Mol. Plant-Microbe Interact. 18:158-168.

Soosaar, J. L., Burch-Smith, T. M., and Dinesh-Kumar, S. P. 2005. Mechanisms of plant resistance to viruses. Nat. Rev. Microbiol. 3:789-798.

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

Spoel, S. H., Johnson, J. S., and Dong, X. 2007. Regulation of tradeoffs between plant defenses against pathogens with different lifestyles. Proc. Natl. Acad. Sci. U.S.A. 104:18842-18847.

Spoel, S. H., Mou, Z., Tada, Y., Spivey, N. W., Genschik, P., and Dong, X. 2009. Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. Cell 137:860-872.

Srinivasan, D. G., and Brisson, J. A. 2012. Aphids: A model for polyphenism and epigenetics. Genetics Res. Int. 2012:12.

Steiner, H. 2004. Peptidoglycan recognition proteins: on and off switches for innate immunity. Immunol. Rev. 198:83-96.

Stulemeijer, I. J., Stratmann, J. W., and Joosten, M. H. 2007. Tomato mitogen-activated protein kinases *Le*MPK1, *Le*MPK2, and *Le*MPK3 are activated during the Cf-4/Avr4-induced hypersensitive response and have distinct phosphorylation specificities. Plant Physiol. 144:1481-1494.

Sun, C., Palmqvist, S., Olsson, H., Boren, M., Ahlandsberg, S., and Jansson, C. 2003. A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the *iso1* promoter. Plant Cell 15:2076-2092.

Suttipanta, N., Pattanaik, S., Kulshrestha, M., Patra, B., Singh, S. K., and Yuan, L. 2011. The transcription factor *Cr*WRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. Plant Physiol. 157:2081-2093.

Taj, G., Agarwal, P., Grant, M., and Kumar, A. 2010. MAPK machinery in plants: recognition and response to different stresses through multiple signal transduction pathways. Plant Signal. Behav. 5:1370-1378.

Tamas, I., Klasson, L., Canback, B., Naslund, A. K., Eriksson, A. S., Wernegreen, J. J., Sandstrom, J. P., Moran, N. A., and Andersson, S. G. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science 296:2376-2379.

Tameling, W. I., Elzinga, S. D., Darmin, P. S., Vossen, J. H., Takken, F. L., Haring, M. A., and Cornelissen, B. J. 2002. The tomato *R* gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity. Plant Cell 14:2929-2939.

Tameling, W. I., Vossen, J. H., Albrecht, M., Lengauer, T., Berden, J. A., Haring, M. A., Cornelissen, B. J., and Takken, F. L. 2006. Mutations in the NB-ARC domain of I-2 that impair ATP hydrolysis cause autoactivation. Plant Physiol. 140:1233-1245.

Tao, Z., Liu, H., Qiu, D., Zhou, Y., Li, X., Xu, C., and Wang, S. 2009. A pair of allelic WRKY genes play opposite roles in rice-bacteria interactions. Plant Physiol. 151:936-948.

Thatcher, L. F., Manners, J. M., and Kazan, K. 2009. *Fusarium oxysporum* hijacks *COI1*-mediated jasmonate signaling to promote disease development in Arabidopsis. Plant J. 58:927-939.

Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G., Nomura, K., He, S. Y., Howe, G. A., and Browse, J. 2007. JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. Nature 448:661-665.

Thomas, C. M., Jones, D. A., Parniske, M., Harrison, K., Balint-Kurti, P. J., Hatzixanthis, K., and Jones, J. D. G. 1997. Characterization of the Tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. Plant Cell 9:2209-2224.

Tjallingii, W. F. 1995. Regulation of phloem sap feeding by aphids. Pages 190–209 in: Regulatory mechanisms in insect feeding, R.F. Chapman and G. De Boer, eds. Chapman and Hall, New York.

Tjallingii, W. F. 2006. Salivary secretions by aphids interacting with proteins of phloem wound responses. J. Exp. Bot. 57:739-745.

Tjallingii, W. F., and Hogen Esch, T. 1993. Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. Physiol. Entomol. 18:317–328.

Toledo-Ortiz, G., Huq, E., and Quail, P. H. 2003. The Arabidopsis basic/helix-loop-helix transcription factor family. Plant Cell 15:1749-1770.

Tomoyasu, Y., Miller, S. C., Tomita, S., Schoppmeier, M., Grossmann, D., and Bucher, G. 2008. Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in Tribolium. Genome Biol. 9:R10.

Torres, M. A., Jones, J. D., and Dangl, J. L. 2006. Reactive oxygen species signaling in response to pathogens. Plant Physiol. 141:373-378.

Trujillo, M., and Shirasu, K. 2010. Ubiquitination in plant immunity. Curr. Opin. Plant Biol. 13:402-408.

Tsuchida, T., Koga, R., and Fukatsu, T. 2004. Host plant specialization governed by facultative symbiont. Science 303:1989.

Tsuchida, T., Koga, R., Shibao, H., Matsumoto, T., and Fukatsu, T. 2002. Diversity and geographic distribution of secondary endosymbiotic bacteria in natural populations of the pea aphid, *Acyrthosiphon pisum*. Mol. Ecol. 11:2123-2135.

Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J., and Katagiri, F. 2009. Network properties of robust immunity in plants. PLoS Genet. 5:e1000772.

Unver, T., and Budak, H. 2009. Virus-induced gene silencing, a post transcriptional gene silencing method. Int. J. Plant Genomics 2009: doi:10.1155/2009/198680

Uppalapati, S. R., Ishiga, Y., Wangdi, T., Kunkel, B. N., Anand, A., Mysore, K. S., and Bender, C. L. 2007. The phytotoxin coronatine contributes to pathogen fitness and is required for suppression of salicylic acid accumulation in tomato inoculated with *Pseudomonas syringae* pv. *tomato DC3000*. Mol. Plant-Microbe Interact. 20:955-965.

Urbanska, A., Tjallingii, W. F., Dixon, A. F. G., and Leszczynski, B. 1998. Phenol oxidising enzymes in the grain aphid's saliva. Entomol. Exp. Appl. 86:197-203.

van den Burg, H. A., Tsitsigiannis, D. I., Rowland, O., Lo, J., Rallapalli, G., Maclean, D., Takken, F. L., and Jones, J. D. 2008. The F-box protein ACRE189/ACIF1 regulates cell death and defense responses activated during pathogen recognition in tobacco and tomato. Plant Cell 20:697-719.

van der Hoorn, R. A., and Kamoun, S. 2008. From Guard to Decoy: a new model for perception of plant pathogen effectors. Plant Cell 20:2009-2017.

van Esse, H. P., Fradin, E. F., de Groot, P. J., de Wit, P. J., and Thomma, B. P. 2009. Tomato transcriptional responses to a foliar and a vascular fungal pathogen are distinct. Mol. Plant-Microbe Interact. 22:245-258.

van Verk, M. C., Pappaioannou, D., Neeleman, L., Bol, J. F., and Linthorst, H. J. 2008. A novel WRKY transcription factor is required for induction of *PR-1a* gene expression by salicylic acid and bacterial elicitors. Plant Physiol. 146:1983-1995.

van Wees, S. C., de Swart, E. A., van Pelt, J. A., van Loon, L. C., and Pieterse, C. M. 2000. Enhancement of induced disease resistance by simultaneous activation of salicylate- and jasmonate-dependent defense pathways in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U.S.A. 97:8711-8716.

Vossen, J. H., Abd-El-Haliem, A., Fradin, E. F., van den Berg, G. C., Ekengren, S. K., Meijer, H. J., Seifi, A., Bai, Y., ten Have, A., Munnik, T., Thomma, B. P., and Joosten, M. H. 2010. Identification of tomato phosphatidylinositol-specific phospholipase-C (PI-PLC) family members and the role of PLC4 and PLC6 in HR and disease resistance. Plant J. 62:224-239.

Walling, L. L. 2008. Avoiding effective defenses: strategies employed by phloem-feeding insects. Plant Physiol. 146:859-866.

Walters, D. 2011. Plant Defense: Warding off attack by pathogens, herbivores and parasitic plants. Blackwell Publishing Ltd. London.

Wangdi, T., Uppalapati, S. R., Nagaraj, S., Ryu, C. M., Bender, C. L., and Mysore, K. S. 2010. A virus-induced gene silencing screen identifies a role for *Thylakoid Formation1* in *Pseudomonas syringae* pv *tomato* symptom development in tomato and Arabidopsis. Plant Physiol. 152:281-292.

Wei, K. F., Chen, J., Chen, Y. F., Wu, L. J., and Xie, D. X. 2012. Molecular Phylogenetic and Expression Analysis of the Complete WRKY Transcription Factor Family in Maize. DNA Res 19:153-164.

Wei, W., Zhang, Y., Han, L., Guan, Z., and Chai, T. 2008. A novel WRKY transcriptional factor from *Thlaspi caerulescens* negatively regulates the osmotic stress tolerance of transgenic tobacco. Plant Cell Reports 27:795-803.

Wiermer, M., Feys, B. J., and Parker, J. E. 2005. Plant immunity: the EDS1 regulatory node. Curr. Opin. Plant Biol. 8:383-389.

Will, T., Tjallingii, W. F., Thonnessen, A., and van Bel, A. J. 2007. Molecular sabotage of plant defense by aphid saliva. Proc. Natl. Acad. Sci. U.S.A. 104:10536-10541.

Will, T., Kornemann, S. R., Furch, A. C., Tjallingii, W. F., and van Bel, A. J. 2009. Aphid watery saliva counteracts sieve-tube occlusion: a universal phenomenon? J. Exp. Bot. 212:3305-3312.

Williamson, V. M., and Gleason, C. A. 2003. Plant-nematode interactions. Curr. Opin. Plant Biol. 6:327-333.

Wu, K. L., Guo, Z. J., Wang, H. H., and Li, J. 2005. The WRKY family of transcription factors in rice and Arabidopsis and their origins. DNA Res. 12:9-26.

Wu, X., Shiroto, Y., Kishitani, S., Ito, Y., and Toriyama, K. 2009. Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing *OsWRKY11* under the control of *HSP101* promoter. Plant Cell Reports 28:21-30.

Xie, Z., Zhang, Z. L., Zou, X., Huang, J., Ruas, P., Thompson, D., and Shen, Q. J. 2005. Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. Plant Physiol. 137:176-189.

Yang, B., Jiang, Y., Rahman, M. H., Deyholos, M. K., and Kav, N. N. 2009. Identification and expression analysis of WRKY transcription factor genes in canola (*Brassica napus* L.) in response to fungal pathogens and hormone treatments. BMC Plant Biol. 9:68.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li, Z., Xiaohui, D., Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G., and Li-Jia, Q. 2006. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol. 60:107-124.

Yao, D., Zhang, X., Zhao, X., Liu, C., Wang, C., Zhang, Z., Zhang, C., Wei, Q., Wang, Q., Yan, H., Li, F., and Su, Z. 2011. Transcriptome analysis reveals salt-stress-regulated biological processes and key pathways in roots of cotton (*Gossypium hirsutum* L.). Genomics 98:47-55.

Yu, S., Ligang, C., Liping, Z., and Diqiu, Y. 2010. Overexpression of *OsWRKY72* gene interferes in the abscisic acid signal and auxin transport pathway of Arabidopsis. J. Biosci. 35:459-471.

Zarate, S. I., Kempema, L. A., and Walling, L. L. 2007. Silverleaf whitefly induces salicylic acid defenses and suppresses effectual jasmonic acid defenses. Plant Physiol. 143:866-875.

Zhang, Y., and Wang, L. 2005. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol. Biol. 5:1.

Zhang, Z. L., Xie, Z., Zou, X., Casaretto, J., Ho, T. H., and Shen, Q. J. 2004. A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. Plant Physiol. 134:1500-1513.

Zheng, X. Y., Spivey, N. W., Zeng, W., Liu, P. P., Fu, Z. Q., Klessig, D. F., He, S. Y., and Dong, X. 2012. Coronatine promotes pseudomonas syringae virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. Cell Host Microbe 11:587-596.

Zhou, Q. Y., Tian, A. G., Zou, H. F., Xie, Z. M., Lei, G., Huang, J., Wang, C. M., Wang, H. W., Zhang, J. S., and Chen, S. Y. 2008. Soybean WRKY-type transcription factor genes, *GmWRKY13*, *GmWRKY21*, and *GmWRKY54*, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants. Plant Biotechnol. J. 6:486-503.

CHAPTER ONE

SlWRKY70 is required for *Mi-1*-mediated resistance to aphids and nematodes in tomato.

Abstract

Plant resistance (*R*) gene-mediated defense responses against biotic stresses include vast transcriptional reprogramming. In several plant-pathogen systems, members of the WRKY family of transcription factors have been demonstrated to act as both positive and negative regulators of plant defense transcriptional networks. To identify possible roles of tomato (*Solanum lycopersicum)* WRKY transcription factors in defense mediated by the *R* gene *Mi-1* against potato aphid, *Macrosiphum euphorbiae*, and root-knot nematode (RKN), *Meloidogyne javanica*, we used tobacco rattle virus (TRV)-based virus-induced gene silencing and transcriptionally suppressed *SlWRKY70,* a tomato ortholog of the *Arabidopsis thaliana WRKY70* gene. Silencing *SlWRKY70* attenuated *Mi-1*-mediated resistance against both potato aphid and RKN showing that *SlWRKY70* is required for *Mi-1* function. Furthermore, we found *SlWRKY70* transcripts to be inducible in response to aphid infestation and RKN inoculation. *Mi-1*-mediated recognition of these pests modulates this transcriptional response. As previously described for *AtWRKY70*, we found *SlWRKY70* transcript levels to be up-regulated by salicylic acid and suppressed by methyl jasmonate. This indicates that some aspects of WRKY70 regulation are conserved among distantly related eudicots.

Introduction

Plants are hosts to a wide range of pathogens and pests that utilize them as a source of energy and nutrients vital for their survival and reproduction. To protect themselves from pathogen and pest attack, plants employ various defense strategies. Besides preformed physical barriers and chemical defenses, plants utilize inducible immune responses that are regulated by complex signaling networks primarily at the level of transcription. Thus, transcription factors play an important part in regulating the temporal and spatial expression patterns of genes involved in plant defense responses (Eulgem 2005; Rushton and Somssich 1998; Singh et al. 2002).

Two classes of immune receptors trigger defense-associated transcriptional reprogramming and immunity in plants. Pattern recognition receptors (PRRs) mediate recognition of pathogen associated molecular patterns (PAMPs), chemical signatures that appear to be widely conserved among certain pathogen clades (Jones and Dangl 2006). The resulting PAMP-triggered immunity (PTI) is often counteracted by pathogen-derived effector molecules that are secreted into host cells (Abramovitch et al. 2006). Consequently, immunity of the host is weakened allowing for growth and propagation of the invading pathogen resulting in compatible interactions. Despite the virulence activity of effectors, plants are often able to respond to PAMP perception with a weakened immune reaction termed basal defense, which limits spread and growth of the pathogen. A second class of plant immune receptors, termed disease resistance (R) proteins, can recognize pathogen effectors and activate effector-triggered immunity (ETI), a strong immune response resulting in incompatible plant-

pathogen/pest interactions (Jones and Dangl 2006). ETI is a form of the well-described phenomenon of gene-for-gene resistance (Flor 1971), as it is triggered by a pair of complementary host *R*-genes and avirulence-conferring pathogen effector genes (*Avr* genes). Numerous studies have shown that PTI, basal defense and ETI utilize related signaling processes, which involve the defense hormones salicylic acid (SA) and jasmonic acid (JA) (Glazebrook et al. 2003; Nimchuk et al. 2003). Both synergistic and antagonistic effects on immunity between these two types of hormones have been described (Mur et al. 2006; Tsuda et al. 2009). Despite extensive efforts during the past two decades, molecular mechanisms connecting R-mediated effector recognition to regulatory processes involved in basal defense and PTI are largely elusive. Recent reports, however, are suggesting that R proteins directly interfere with transcriptional regulators to activate the transcriptional network controlling immunity (Cheng et al. 2009; Shen et al. 1992; Wirthmueller et al. 2007).

Several families of transcriptions factors (TFs) are known to regulate plant immune responses against pathogens and pests (Singh et al. 2002). The WRKY family of TFs, originally believed to be unique to plants, was recently shown to have much earlier evolutionary origins (Pan et al. 2009). In Arabidopsis (*Arabidopsis thaliana)* and rice (*Oryza sativa*) the WRKY family consists of 74 and 102 members, respectively (Ross et al. 2007). Members of this family contain either one or two copies of the conserved WRKY domain. Frequently, these ~60 amino acid comprising domains mediate binding to a pathogen-responsive promoter element called the W-

box (Eulgem et al. 2000). WRKY family members are divided into three major groups based on the number and variations of their WRKY domains (Eulgem et al. 2000).

In a wide range of plant-pathogen systems, loss- and gain-of-function studies have demonstrated the involvement of WRKY TFs as both positive and negative regulators of the plant defense network (Eulgem and Somssich 2007; Pandey and Somssich 2009). However, roles for WRKY TFs in plant immune responses against herbivore pests are not widely characterized. In *Nicotiana attenuata, NaWRKY3* and *NaWRKY6* have been shown to be required for resistance against larvae of the tobacco hornworm, *Menduca sexta*. Silencing of these WRKY TFs by stable transformation resulted in impaired JA accumulation suggesting that these TFs control plant immune responses by regulating the JA-signaling pathway (Skibbe et al. 2008). In tomato (*Solanum lycopersicum*), *SlWRKY72a* and *SlWRKY72b* were shown to be involved in *Mi-1*-mediated resistance as well as basal defense against potato aphid, *Macrosiphum euphorbiae* and root-knot nematodes (RKN), *Meloidogyne* species (Bhattarai et al. 2010). Recently, (Van Eck et al. 2010)(2010) reported a role for *TaWRKY53* in bread wheat (*Triticum aestivum*) resistance against Russian wheat aphid, *Diuraphis noxia* biotype RWA2.

In Arabidopsis, *AtWRKY70* encoding a group III WRKY protein has been shown to play a complex role in defense and to integrate signals from both JA- and SA-mediated defense pathways (Li et al. 2006a; Li et al. 2004). This gene promotes disease resistance to various pathogens as part of a SA-dependent inducible mechanism, while suppressing defense responses mediated by JA (AbuQamar et al.

2006; Knoth et al. 2007; Li et al. 2006a; Li et al. 2004). Overexpression of *At*WRKY70 enhanced resistance to the bacterial necrotroph *Erwinia carotovora* and the hemibiotroph *Pseudomonas syringae* pv. *tomato*, while resistance to the necrotrophic fungi *Alternaria brassicicola* which requires JA-mediated signal transduction pathway (Thomma et al. 1998), was reduced (Li et al. 2004). Furthermore, knock-down of *AtWRKY70* reduced SA-mediated basal defense to the fungal biotroph *Erysiphae cichoracearum* and enhanced susceptibility to the fungal necrotroph *Botrytis cinerea* (AbuQamar et al. 2006; Li et al. 2006a). Moreover, using T-DNA mutants *AtWRKY70* was found to contribute to SA-dependent basal defense and gene-for-gene resistance mediated by the *R*-gene *RPP4* against the biotrophic oomycete *Hyaloperonospora arabidopsidis* (Knoth et al. 2007).

SA mediates *AtWRKY70* transcript accumulation. Consistently *AtWRKY70* orthologs in tobacco (*Nicotiana tabacum*) are transcriptionally inducible by SA (Chen and Chen 2000). On the other hand, transcript levels of *AtWRKY70* are repressed by the stress hormone JA (Li et al. 2004). The effect of JA on *AtWRKY70* expression seems to be complex involving a mechanism dependent on the F-box protein COI1 as well as a COI-1-independent pathway (Li et al. 2004; Ren et al. 2008). Li et al. (2004) also showed that *AtWRKY70* activates expression of defense-related genes known to be inducible by SA, but suppresses expression of JA-responsive genes further supporting it may act as a node of convergence for integrating SA- and JA-signaling events during plant defense.

After nearly 20 years of intensive research defense signaling processes triggered by *R*-genes are still insufficiently understood. Only a small number of *R* genes from solanaceous species have so far been functionally characterized. We previously demonstrated by virus-induced gene silencing (VIGS) that the tomato *R*-gene *Mi-1* requires orthologs of the Arabidopsis defense regulators, SGT1b, HSP90, and a mitogen-activated protein kinase (MAPK) cascade including the MAPK kinase LeMKK2 and the MAPKs LeMAPK1, LeMAPK2 and LeMAPK3 (Bhattarai et al. 2007; Li et al. 2006b). In addition, we identified the somatic embrogenesis receptor kinase 1 (SERK1) to be required for *Mi-1*-mediated aphid resistance (Mantelin et al. 2011). We also found WRKY72-type transcription factors to contribute to *Mi-1*-mediated immunity (Bhattarai et al. 2010). In order to identify additional components of this pathway we have been testing other *WRKY* genes for their role in *Mi-1*-mediate pest resistance.

Here we report on the use of VIGS to transiently knockdown the *AtWRKY70* ortholog *SlWRKY70* in tomato to assess its role in *Mi-1*-mediated resistance. Moreover, we profile the expression of *SlWRKY70* after potato aphid infestation, RKN inoculation, and treatments with SA or methyl jasmonate (MeJA) hormones. Our data implicate a role for *SlWRKY70* in *Mi-1*-mediated resistance against potato aphids and RKN and show differential regulation after aphid infestation, RKN infection and hormone treatments. The present work solely focuses on the contribution of *SlWRKY70* to immunity mediated by *Mi-1*. The contribution of

WRKY70-type transcription factors to additional immune responses, such as basal defense, systemic immunity or SA-priming in tomato, is beyond the scope of this study.

Materials and Methods

Plant materials and growth conditions

Tomato cv. UC82B (*mi-1*/*mi-1*) (Lockhart Seeds Inc., Stockton, CA), cv. Castlemart (*mi-1*/*mi-1*), *jai1-1* mutant (cv. Castlemart background) and near isogenic lines cv. Motelle (*Mi-1*/*Mi-1*) and cv. Moneymaker (*mi-1*/*mi-1*) were used. Castlemart and *jai1-1* mutant seeds were obtained from G. Howe, Michigan State University, while the remaining tomato genotypes were bulked in our lab. Homozygous *jai1-1* mutant plants are sterile. A heterozygous population of *jai1-1* mutants was screened for MeJA sensitivity and genotyped for the presence of deletion in the *COI-1* gene as described previously (Bhattarai et al. 2007b).

Seedlings with a pair of newly emerged leaves were used in VIGS and maintained at 19ºC in growth chambers with a 16-h-light and 8-h-dark photoperiod and 200 μmol m^{-2} s^{-1} light intensity until bioassay (Bhattarai et al. 2007a). Plants were supplemented with Osmocote (17-6-10; Sierra Chemical Company) and fertilized biweekly with MiracleGro (18-18-21; Scotts MiracleGro Company, Marysville, OH).

Potato aphid and RKN rearing and inoculum preparation

A colony of a parthenogenetic *Mi-1*–avirulent potato aphid, and *Mi-1*-avirulent *M. javanica* were grown on susceptible tomato cv. UC82B plants. Potato aphids were

maintained inside an insect cage in a pesticide-free greenhouse at 22-26ºC and nematodes in a greenhouse at 23-30ºC. Nematode eggs were collected by bleach extracting the roots following established protocols (Hussey 1973). Eggs were allowed to hatch using modified Baermann funnels (Martinez-de IIarduya et al. 2001). After 2 days, infective-stage juveniles (J2) were collected and used directly in bioassays. For transcript profiling after RKN inoculation, J2 were cleaned using a sand column as described in Lambert et al. (1999).

VIGS experiments and potato aphid/RKN bioassays

The bipartite tobacco rattle virus (TRV) vector (pTRV1 and pTRV2) was used for VIGS (Hayward et al. 2011). Cultures of *Agrobacterium tumefaciens* strain GV3101 containing pTRV1 or a pTRV2 containing a *Nicotiana benthamiana* VIGS construct WRKY3 described previously were grown as described earlier (Li et al. 2006b). *A. tumefaciens* cultures were pelleted, resuspended in infiltration buffer, and adjusted to an OD_{600} of 1.0. Cells were incubated at room temperature for 3 h before use. Equal volume of pTRV1 Agrobacterium culture was mixed with *WRKY* sequence-containing pTRV2 or pTRV2 empty vector culture before infiltration. Leaflets of 2-3-week-old seedlings were infiltrated with Agrobacterium cultures (agroinfiltration) using a 1-mL needle less syringe.

Plants treated with TRV and maintained at 19ºC for 4-5 weeks were used in aphid bioassays. Around 50 mixed stages of potato aphids were caged onto four

individual leaflets per plant and aphid survival was recorded after 10-12 days once all the aphids were dead on the resistant genotype. For the RKN bioassays, two weeks after TRV infiltration, plants were individually inoculated with 5,000 J2. Plants were maintained at 24ºC in a growth chamber for three weeks and then transferred to a greenhouse at 23-28ºC for an additional four weeks. Later, roots were stained in 0.001% (w/v) erioglaucine (Sigma-Aldrich, St. Louis, MO) and nematode egg masses were counted. For each experiment, 10 and 20 plants per VIGS construct were used for aphid and nematode bioassays, respectively. Experiments were performed twice. For *SlWRKY70* transcript evaluation, individual leaflets were collected from agroinfiltrated resistant cv. Motelle showing aphid susceptibility and controls and instantly frozen in liquid nitrogen and stored at -80ºC until RNA extraction.

Potato aphid and RKN time-course infestation experiments

Time-course potato aphid infestation and RKN inoculation experiments were described previously (Bhattarai et al. 2007a; Bhattarai et al. 2008).

Hormone treatments

Five-week-old cv. Motelle tomato plants were sprayed with 1.5 mM SA (Sigma-Aldrich, St. Louis, MO) and 1.5 mM MeJA (Bedoukian Research, Inc., Danbury, CT) hormones using PREVAL 267-paint sprayers. Several leaflets from the top

growth were collected 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 24, and 48 h after treatment (hat), instantly frozen in liquid nitrogen and stored at -80ºC until RNA extraction.

RNA extraction and reverse transcription

For RNA extraction from leaves, leaflets were ground to powder in liquid nitrogen and RNA was extracted using hot phenol as described previously or using TRIzol (Invitrogen) according to manufacturer's recommendations (Martinez-de IIarduya et al. 2001). For RNA extraction from roots, hot phenol was used as described previously (Bhattarai et al. 2008).

For cDNA synthesis, 20 μg of RNA was treated with DNase I (New England Biolabs) followed by phenol-chloroform extraction and RNA precipitation with isopropanol. cDNA was synthesized from 5μg DNase treated RNA using Superscript III (Invitrogen) reverse transcriptase enzyme and Oligo-dT primers according to the manufacturer's recommendations.

Transcript level evaluation

Semi-quantitative and quantitative approaches were used to evaluate transcript levels. Semi-quantitative PCR was used to analyze the expression of *SlWRKY70* in silenced and control leaves and after hormone treatments using primers WRKY70F (5'-

AGAAGAAGAAG GAGAAGCAAGACCG-3') and WRKY70R (5'-TGTCCTTTGGATTCTTCCTCTT-3'). Ubiquitin (*Ubi3*-X58253) was used as internal control and amplified using primers Ubi3F (5'-GTGTGGGCTCACCTACGTTT-3') and Ubi3R (5'-ACAATCCCAAGGGTTGTCAC-3'). As control for the hormone treatments, the expression of two marker genes, *PR1b1* (Y08804) and *PinII* (AY129402), known to be induced by exogenous application of SA or JA, respectively, was assessed using primers PR1b1F (5'- TTATACTCAAGTAGTCTGGCGCA-3'), PR1b1R (5'-TTGCAAGAAATGAACCACCA-3') and PinIIF (5'-CACAGGGTACAAGGGTTGCT-3'), PinIIR (5'- TTTTGGGCAATCCAGAAGAT-3'), respectively.

PCR was performed in 25 μl with 1 μl of template cDNA, 2.5 μl of 10X PCR buffer, 2.5 μl of 25 mM $MgCl_2$, 0.4 μl of 25 mM dNTPs, 0.5 μl of *Taq* DNA polymerase, and 10 μM of forward and reverse primers. The PCR program was initialized at 94° C for 5 min, followed by 28 cycles of 94° C for 1 min, 58° C for 45 s, and 72° C for 1 min, with a final extension at 72° C for 10 min.

Quantitative PCR was used to analyze the temporal expression of *SlWRKY70* in leaves and roots of susceptible and resistant tomato after aphid or RKN attack, respectively. Similarly, using qPCR the expression of *SlWRKY70* was evaluated in *jai1-1* mutant and the wild type parent CM. Primers SlWRKYq70F 5'-CATGGATGAGAGAATCTGCA- 3' and SlWRKYq70R 5'-

GGATTTTCTTGGATTATTTGAAC- 3' were used. The iQ-SYBR Green Supermix (Bio-Rad) was used as intercalating dye to detect the amplification product in iCyclcler5 IQ (Bio-Rad). PCR was carried out in 20 μl using 2.5 μl of 5-times-diluted cDNA as template. The PCR program was initialized at 95° C for 5 min, followed by 45 cycles of 95° C for 30 s, 58° C for 30 s, and 72° C for 45 s, with a final extension at 72° C for 5 min. Three technical replicates were performed. The fold change expression was calculated using the statistical model described in (Fu et al. 2006). Briefly, ΔΔCt was calculated by subtracting ΔCt ($Ct_{target} - Ct_{Ub-3}$) and the ΔΔCt of the treated samples were divided by that of the 0 h time-point sample and the fold change calculated as $2^{\Delta\Delta Ct}$.

Phylogenetic analysis

Phylogenetic tree was built based on alignments of six Arabidopsis Group III and one tomato WRKY amino acid sequences. The sequences were manually aligned based on conserved motifs and domains. Trees were constructed with MrBayes (version 3.2-cvs) (Ronquist and Huelsenbeck 2003) under mixed protein substitution model running 4 chains and 3 runs for 100,000 generations, though the runs converged (average stddev of split frequencies < 0.01) after 30,000 generations. Maximum likelihood phylogenetic reconstruction with RAxML (7.0.4) (Stamatakis 2006) using rapid bootstrap were also computed and found to have identical topology.

Statistical analysis

For the VIGS bioassays statistical analysis was performed using one tailed Flinger-Policello test followed by Bonferroni adjustment (BenMamoun 2006).

Results

VIGS constructs and gene annotation

The tobacco rattle virus (TRV)-based VIGS construct WRKY3 described in Ekengren et al. (2003) was prepared from *N. benthamiana* cDNA fragments. In the absence of a full tomato and *N. benthamiana* genome sequences, these constructs were originally named based on limited partial sequence information. New genome sequence resources that have become available for tomato provided an opportunity to determine the tomato genes silenced by this VIGS construct. Since the whole genome of Arabidopsis is available and fully annotated, the Arabidopsis classification of the WRKY transcription factors was used as reference for the annotation of the tomato gene(s) silenced by this *N. benthamiana* VIGS construct. Sequence analyses using full genome tomato sequences showed that the WRKY3-VIGS construct (hereon referred as TRV-WRKY70) is predicted to specifically silence a tomato gene, with high similarity to *AtWRKY70,* a member of group III of the WRKY family. The phylogenetic analysis shows that this tomato gene is closely related to *AtWRKY70* and *AtWRKY54* that seem to represent a *WRKY* gene that has duplicated after divergence of tomato and Arabidopsis (Fig. 1). Since the protein encoded by this tomato gene has higher amino acid identity to *AtWRKY70* (e = 2e-21) compared to the closely related *AtWRKY54* (e = 3e-14) (Fig. 1), we presumed *AtWRKY70* likely to be the gene with retained orthologous function. Thus, we named it *SlWRKY70*. The amino acid sequence of SlWRKY70 is shown in figure 1b. As highlighted in this

figure *Sl*WRKY70 shares with *At*WRKY70 not only a group III-type WRKY DNA binding domain, but also a Q-rich motif in the N-terminal region as well as a putative nuclear localization signal. Q-rich motifs often function as transcriptional activation domains. Although neither the Q-rich region of *At*WRKY70, nor its putative nuclear localization signal have been shown to be functional, the sequence features conserved between *At*WRKY70 and *Sl*WRKY70 suggest that both of these proteins act as nuclear localized transcriptional activators that bind to promoters via their WRKY domains.

SlWRKY70 is required for *Mi-1*-mediated resistance

Tomato cv. Motelle (*Mi-1*/*Mi-1*) and Moneymaker (*mi-1*/*mi-1*) agroinfiltrated with TRV1 and TRV2 empty vector or TRV2 containing *SlWRKY70* (heron referred to as TRV-WRKY70) were used in potato aphid and RKN bioassays. Two weeks after aphid infestation, no live aphids were found on leaflets of the control resistant cv. Motelle plants agroinfiltrated with the TRV1 and the TRV2 empty vector (Fig. 2a). In contrast, numerous aphids were still alive on some leaflets of cv. Motelle agroinfiltrated with the TRV1 and the TRV-WRKY70 construct and on all infested leaflets of susceptible cv. Moneymaker plants agroinfiltrated with TRV1 and TRV2 empty vector (Fig. 2a), indicating that *SlWRKY70* is required for *Mi-1* mediated resistance against potato aphids. TRV silencing is known to be patchy in tomato,

which may account for the inconsistent attenuation of resistance in resistant cv. Motelle plants agroinfiltrated with TRV-WRKY70 construct (Bhattarai et al. 2007).

Root-knot nematodes were able to infect and reproduce on sectors of roots of resistant cv. Motelle agroinfiltrated with TRV1 and TRV-WRKY70 VIGS construct, while no RKN infection and reproduction was observed on Motelle plants agroinfiltrated with TRV1 and TRV2 empty vector (Fig. 2b). This shows that *SlWRKY70* also contributes to *Mi-1*-mediated resistance against RKN. In *SlWRKY70*-silenced Motelle plants the attenuation of resistance against RKN was less pronounced compared to resistance against potato aphids. Most likely this reflects the known phenomenon that VIGS-mediated gene silencing is less efficient in roots compared to leaves (Bhattarai et al. 2007). To confirm that *SlWRKY70* is efficiently silenced in TRV-WRKY70 agroinfiltrated Motelle plants, leaflets showing attenuated aphid resistance were used as a source of RNA for transcript evaluation. sqPCR analyses showed that *SlWRKY70* transcript levels are clearly reduced in silenced leaflets compared to TRV2 empty vector agroinfiltrated control leaflets though at variable levels (Fig. 3).

SlWRKY70 expression after exposure to potato aphids or RKN

Temporal expression of *SlWRKY70* was assessed after exposure to aphids or RKN using qPCR. Transcript levels of *SlWRKY70* was differentially regulated after exposure to these pests (Fig. 4). Potato aphid infestation resulted in accumulation of

SlWRKY70 transcripts starting 6 h after treatment (hat) with aphids in the resistant leaflets of cv. Motelle but not in the susceptible Moneymaker, reaching their peak at 24 hat and returning to basal levels at 48 hat (Fig. 4a). The induction level was considerably higher in the resistant Motelle genotype compared to the susceptible Moneymaker. On the other hand, *SlWRKY70* expression was induced to considerably higher levels 12 hat with RKN in the resistant roots compared to the susceptible. However with RKN, the maximum level of transcript accumulation was similar for both genotypes which was at 36 hat (Fig. 4b). Taken together these results indicate that *SlWRKY70* transcript levels are induced during both basal defense to aphids and RKN as well as *Mi-1*-mediated resistance to these pests. However, *Mi-1*-mediated recognition of aphids or RKN mediates an enhancement or acceleration of this response, respectively.

SlWRKY70 expression after hormone treatments

Being involved in plant defense against various pathogens and in diverse plant developmental processes, *WRKY* genes are regulated by multiple hormonal signaling pathways (Ramamoorthy et al. 2008; Yang et al. 2009). SqPCR analysis of the temporal expression of *SlWRKY70* transcripts in tomato seedlings treated with SA and MeJA hormones revealed that this gene is differentially regulated by these hormones (Fig. 5). *SlWRKY70* transcript levels were increased as early as 1 hat with SA, peaking around 6 hat and returned to basal levels starting 12 hat (Fig. 5a).

Temporal expression analysis of seedlings treated with MeJA revealed repression of *SlWRKY70* expression starting 3 hat with slight increase at 6 hat (Fig. 5b). No *SlWRKY70* transcripts were detected at 48 hat with MeJA suggesting that the hormone completely suppressed expression of this gene in tomato (Fig. 5b).

Basal expression of *SlWRKY70* in the *jai1-1* mutant

Untreated Arabidopsis *coi1* mutants exhibit elevated transcript levels of *AtWRKY70* indicating that *COI1* negatively regulates basal expression of this gene (AbuQamar et al. 2006; Li et al. 2004). We tested whether *SlWRKY70* is also negatively regulated by *SlCOI1* by evaluating *SlWRKY70* transcript levels in the tomato mutant line *jai1-1*, a functional null allele of the tomato *COI1* ortholog. As shown by qPCR the tomato *jai1-1* mutant exhibits similar *SlWRKY70* basal expression levels as its tomato wild type parent cv. Castlemart (Fig. 6).

Discussion

To expand our understanding of *Mi-1*-mediated tomato resistance against potato aphids and RKN, we characterized possible roles of a tomato WRKY transcription factor in this defense pathway. Using a VIGS gene knockdown approach, we showed that *SlWRKY70* is required for *Mi-1*-mediated resistance against potato aphids and RKN. In addition, we found that *Mi-1*-dependent and *Mi-1*-independent mechanisms up-regulate transcript levels of this *WRKY* gene. However, *SlWRKY70* transcript modulation in response to *Mi-1*-mediated recognition of aphids and RKN is enhanced and accelerated, respectively.

In Arabidopsis *At*WRKY70 is required for full immunity mediated by the *R*-gene *RPP4* (Knoth et al. 2007). This together with our new finding that a tomato orthologue of this transcription factor is required for *Mi-1*-function suggests that the role of WRKY70-type transcription factors in *R*-mediated immunity is conserved. Given the large evolutionary distance between Arabidopsis and tomato, the requirement of WRKY70 orthologs for *R*-gene functions seems to be universal within the clade of eudicots.

Recently, we found by microarray analysis that SA-, JA-, and ethylene (ET)-regulated pathways are activated during both basal defense and *Mi-1* triggered resistance to RKN (Bhattarai et al. 2008). In each case the level of activation was higher during incompatible interactions. This suggests a considerable overlap between basal defense responses and *Mi-1*-mediated resistance in tomato and is consistent with

observations made in Arabidopsis, where global transcript profiles during basal and *R*-mediated defense appear qualitatively similar, but are quantitatively distinct (Tao et al. 2003). In addition, transcriptional changes were found to be accelerated and more intense during *R*-mediated resistance (Tao et al. 2003). Thus, a large part of the difference between compatible and incompatible responses can be explained by quantitative differences in the behavior of the same signal transduction system.

The same concept seems also to apply to *Mi-1*-mediated regulation of *SlWRKY70* in tomato. While *SlWRKY70* transcript levels reached similar levels during both *Mi-1*-mediated resistance and basal defense to RKN, *Mi-1* accelerated this response mediating strong accumulation of this transcript already within 6 hat. Similarly at 24 hat with aphids, *SlWRKY70* transcripts were induced to considerably higher levels in resistant Motelle than in susceptible Moneymaker plants. Therefore, *Mi-1*-mediated signal amplification mechanisms account for the stronger induction of *SlWRKY70* transcript levels after potato aphid infestation in resistant plants and accelerated induction of *SlWRKY70* expression 6 hat with RKN. Such a scenario may imply an additional role of *Sl*WRKY70 in basal defense. Its transcriptional responses during compatible interactions with aphids and RKN and after SA-treatment as well as the fact that its Arabidopsis ortholog promotes basal defense (Li et al. 2004, 2006a; Knoth et al. 2007) provide support for this additional role. Given that most genes that are required for *R*-mediated immunity also contribute to basal defense (Nimchuk et al. 2003), a basal defense function of *Sl*WRKY70 would not be surprising and is rather predictable. Future experiments should address this possibility.

While *Mi-1* appears to utilize a similar set of downstream components in mediating immunity to RKN and potato aphids, there are some notable differences regarding its function in both processes (Bhattarai et al. 2007). *Mi-1* mediated resistance to potato aphids is developmentally regulated and does not involve hypersensitive response (HR), while *Mi-1*-mediated resistance to RKN is active at all growth stages and does include a HR. In both resistant and susceptible tomato genotypes, expression of the *SlWRKY70* gene characterized here partially differed after exposure to potato aphids or RKN. This may reflect tissue-specific differences of defense-regulatory processes. A recent study comparing global transcriptional changes in tomato roots and foliage during incompatible interactions with the vascular fungal pathogen *Verticillium dahliae* Klebahn revealed substantial tissue-specific differences, as more genes were induced in the roots than in the foliage (van Esse et al. 2009). Thus, at least some of the differences observed in *Mi-1*-mediated responses against potato aphids and RKN could result from tissue-specific differences.

WRKY transcription factors are regulated by multiple hormonal signaling pathways and complex crosstalk among these pathways, along with regulatory interactions between individual *WRKY* genes, are known to contribute to the proper expression and function of this large family of TFs (Eulgem and Somssich 2007). We found *SlWRKY70* transcripts to be up-regulated by exogenous application of SA and down-regulated by MeJA in tomato. In Arabidopsis, *AtWRKY70* shows the same response pattern and seems to act as a node of convergence integrating SA- and JA-dependent signals. Our data suggest that mechanisms regulating *WRKY70* expression

are largely conserved between Arabidopsis and tomato. However, unlike the situation on Arabidopsis, we did not observe any *SlCOI1*-mediated suppression of *SlWRKY70* basal expression in tomato, as untreated tomato *jai1-1* mutants and wild type plants accumulated similar transcript levels of this gene.

The observation that *SlWRKY70* transcript levels increase in response to SA is consistent with the observation that *Mi-1* mediated immunity against aphids is SA-dependent (Li et al. 2006b). Thus, *Mi-1*-mediated aphid recognition may activate *WRKY70* function by triggering SA signaling. Besides having a direct role in *R*-mediated immunity and basal defense, SA has been implicated in the phenomenon of defense priming (Conrath et al. 2006; Ahmad et al. 2010*)*. Priming of defense responses in plants involves a primary stimulus, such as exogenous application of SA, which enhances the plant's responsiveness to a secondary defense-related signal, such as pathogen-recognition. This phenomenon has been defined as an "augmented capacity to express basal defense mechanisms" (Ahmad et al. 2010*)*. Broad-spectrum, systemic pathogen defense responses, such as systemic acquired resistance (SAR; Ryals et al. 1996) or induced systemic resistance (ISR, Van Loon et al. 1998), as well as immunity against herbivores, have been linked to priming-related mechanisms (Conrath et al. 2006). Recently, post-translational histone modifications likely affecting chromatin states were shown to be possibly causal for priming of defense-related expression of the Arabidopsis *WRKY* members *AtWRKY6*, *AtWRKY29* and *AtWRKY53* mediated by the SA-analog BTH (Jaskiewicz et al. 2011). While it is

formally possible that *SlWRKY70* is also subject to priming-related regulation, we think that the role of this gene in *Mi-1*-mediated defense is unlikely to involve such an indirect mode of expression control. Accelerated or enhanced transcriptional induction of *SlWRKY70* in response to *Mi-1*-mediated aphid/RKN recognition appears not to involve any "priming stimulus" and is a direct and immediate response to an initial *Mi-1*-generated signal. Nevertheless, we cannot exclude that priming-related processes could further enhance *Mi-1*-mediated responses after prior application of a primary defense signal. Considering that *Mi-1*-mediated resistance to aphids and RKN is extremely tight, however, priming is unlikely to be necessary for resistance mediated by this *R*-gene.

We recently found WRKY72-related transcription factors to have conserved roles in basal defense of tomato and Arabidopsis (Bhattarai et al. 2010). However, *Mi-1* is the only *R*-gene known to signal through WRKYs of this type. Thus, *Mi-1* may utilize a conserved WRKY72-dependent basal defense functions for *R*-mediated immunity. WRKY70 and WRKY72 differ in various structural and functional aspects. While WRKY70 belongs to group III, WRKY72-type transcription factors are members of subgroup IIb of the WRKY family. In contrast to WRKY70 orthologs, which appear to be conserved components of SA-dependent defense mechanisms (this paper; Knoth et al. 2007; Li et al. 2006a; Li et al. 2004), WRKY72-type transcription factors appear to control SA-independent defense responses (Bhattarai et al. 2010). Therefore, *Mi-1* most likely triggers at least two independent

defense signaling routes in parallel: A SA and WRKY70-dependent pathway and a second SA-independent WRKY72-dependent mechanism.

References

Abramovitch, R. B., Anderson, J. C., and Martin, G. B. 2006. Bacterial elicitation and evasion of plant innate immunity. Nat. Rev. Mol. Cell Biol. 7:601-611.

AbuQamar, S., Chen, X., Dhawan, R., Bluhm, B., Salmeron, J., Lam, S., Dietrich, R. A., and Mengiste, T. 2006. Expression profiling and mutant analysis reveals complex regulatory networks involved in Arabidopsis response to Botrytis infection. Plant J. 48:28-44.

Ahmad, S., Gordon-Weeks, R., Pickett, J., and Ton, J. 2010. Natural variation in priming of basal resistance: from evolutionary origin to agricultural exploitation. Mol. Plant Pathol. 11:817–827.

BenMamoun, M. 2006. FPRANK: Stata module to compute Two-Sample Fligner-Policello Robust Rank Order Test.

Bhattarai, K. K., Atamian, H. S., Kaloshian, I., and Eulgem, T. 2010. *WRKY72*-type transcription factors contribute to basal immunity in tomato and Arabidopsis as well as gene-for-gene resistance mediated by the tomato *R*-gene *Mi-1*. Plant J. 63:229-240.

Bhattarai, K. K., Li, Q., Liu, Y., Dinesh-Kumar, S. P., and Kaloshian, I. 2007a. The *Mi-1*-mediated pest resistance requires *Hsp90* and *Sgt1*. Plant Physiol. 144:312-323.

Bhattarai, K. K., Xie, Q. G., Pourshalimi, D., Younglove, T., and Kaloshian, I. 2007b. *Coi1*-dependent signaling pathway is not required for *Mi-1*-mediated potato aphid resistance. Mol. Plant-Microbe Interact. 20:276–282

Bhattarai, K. K., Xie, Q. G., Mantelin, S., Bishnoi, U., Girke, T., Navarre, D. A., and Kaloshian, I. 2008. Tomato susceptibility to root-knot nematodes requires an intact jasmonic acid signaling pathway. Mol. Plant-Microbe Interact. 21:1205-1214.

Chen, C., and Chen, Z. 2000. Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco. Plant Mol. Biol. 42:387-396.

Cheng, Y. T., Germain, H., Wiermer, M., Bi, D., Xu, F., Garcia, A. V., Wirthmueller, L., Despres, C., Parker, J. E., Zhang, Y., and Li, X. 2009. Nuclear pore complex component MOS7/Nup88 is required for innate immunity and nuclear accumulation of defense regulators in Arabidopsis. Plant Cell 21:2503-2516.

Ekengren, S. K., Liu, Y., Schiff, M., Dinesh-Kumar, S. P., and Martin, G. B. 2003. Two MAPK cascades, NPR1, and TGA transcription factors play a role in Pto-mediated disease resistance in tomato. Plant J. 36:905-917.

Eulgem, T. 2005. Regulation of the Arabidopsis defense transcriptome. Trends Plant Sci. 10:71-78.

Eulgem, T., and Somssich, I. E. 2007. Networks of WRKY transcription factors in defense signaling. Curr. Opin. Plant Biol. 10:366-371.

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. 2000. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 5:199-206.

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

Glazebrook, J., Chen, W., Estes, B., Chang, H. S., Nawrath, C., Metraux, J. P., Zhu, T., and Katagiri, F. 2003. Topology of the network integrating salicylate and jasmonate signal transduction derived from global expression phenotyping. Plant J. 34:217-228.

Hayward, A., Padmanabhan, M., and Dinesh-Kumar, S. P. 2011. Virus-induced gene silencing in *Nicotiana benthamiana* and other plant species. Methods Mol. Biol. 678:55-63.

Hussey, K. L. 1973. Effects of microsporidan infection on larval trematodes: infection with *Nosema strigeoideae* or *N. echinostomi*. J. Invertebr. Pathol. 22:193–198.

Jaskiewicz, M., Conrath, U., and Peterhansel, C. 2011. Chromatin modification acts as a memory for systemic acquired resistance in the plant stress response. EMBO Rep. 12:50–55.

Jones, J. D., and Dangl, J. L. 2006. The plant immune system. Nature 444:323-329.

Knoth, C., Ringler, J., Dangl, J. L., and Eulgem, T. 2007. Arabidopsis WRKY70 is required for full *RPP4*-mediated disease resistance and basal defense against *Hyaloperonospora parasitica*. Mol. Plant-Microbe Interact. 20:120-128.

Lambert, K. N., Ferrie, B. J., Nombela, G., Brenner, E. D., and Williamson, V. M. 1999. Identification of genes whose transcripts accumulate rapidly in tomato after root-knot nematode infection. Physiol. Mol. Plant Pathol. 55:341-348.

Li, J., Brader, G., and Palva, E. T. 2004. The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell 16:319-331.

Li, J., Brader, G., Kariola, T., and Palva, E. T. 2006a. WRKY70 modulates the selection of signaling pathways in plant defense. Plant J. 46:477-491.

Li, Q., Xie, Q. G., Smith-Becker, J., Navarre, D. A., and Kaloshian, I. 2006b. *Mi-1*-Mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling cascades. Mol. Plant-Microbe Interact. 19:655-664.

Mantelin, S., Peng, H. C., Li, B., Atamian, H. S., Takken, F. L., and Kaloshian, I. 2011. The receptor-like kinase *SlSERK1* is required for *Mi-1*-mediated resistance to potato aphids in tomato. Plant J. 67:459-471.

Martinez-de IIarduya, O., Moore, A. E., and Kaloshian, I. 2001. The tomato *Rme1* locus is required for *Mi*-mediated resistance to root knot nematodes and the potato aphid. Plant J. 27:417-425.

Mur, L. A., Kenton, P., Atzorn, R., Miersch, O., and Wasternack, C. 2006. The outcomes of concentration-specific interactions between salicylate and jasmonate signaling include synergy, antagonism, and oxidative stress leading to cell death. Plant Physiol. 140:249-262.

Nimchuk, Z., Eulgem, T., Holt, B. F., 3rd, and Dangl, J. L. 2003. Recognition and response in the plant immune system. Annu. Rev. Genet. 37:579-609.

Pan, Y. J., Cho, C. C., Kao, Y. Y., and Sun, C. H. 2009. A Novel WRKY-like Protein Involved in Transcriptional Activation of Cyst Wall Protein Genes in Giardia lamblia. J. Biol. Chem. 284:17975-17988.

Pandey, S. P., and Somssich, I. E. 2009. The role of WRKY transcription factors in plant immunity. Plant Physiol. 150:1648-1655.

Ramamoorthy, R., Jiang, S. Y., Kumar, N., Venkatesh, P. N., and Ramachandran, S. 2008. A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. Plant Cell Physiol. 49:865-879.

Ren, C. M., Zhu, Q., Gao, B. D., Ke, S. Y., Yu, W. C., Xie, D. X., and Peng, W. 2008. Transcription factor WRKY70 displays important but no indispensable roles in jasmonate and salicylic acid signaling. J. Integr. Plant Biol. 50:630-637.

Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Ross, C. A., Liu, Y., and Shen, Q. J. 2007. The *WRKY* gene family in rice (*Oryza sativa*). J. Integr. Plant Biol. 49:827-842.

Rushton, P. J., and Somssich, I. E. 1998. Transcriptional control of plant genes responsive to pathogens. Curr. Opin. Plant Biol. 1:311-315.

Ryals, J. L., Neuenschwander, U. H., Willits, M. C, Molina, A., Steiner, H-Y., and Hunt, M. D. 1996. Systemic acquired resistance. Plant Cell 8:1809–1819

Shen, H., Gold, S. E., Tamaki, S. J., and Keen, N. T. 1992. Construction of a Tn7-lux system for gene expression studies in gram-negative bacteria. Gene 122:27-34.

Singh, K., Foley, R. C., and Onate-Sanchez, L. 2002. Transcription factors in plant defense and stress responses. Curr. Opin. Plant Biol. 5:430-436.

Skibbe, M., Qu, N., Galis, I., and Baldwin, I. T. 2008. Induced plant defenses in the natural environment: *Nicotiana attenuata WRKY3* and *WRKY6* coordinate responses to herbivory. Plant Cell 20:1984-2000.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690.

Tao, Y., Xie, Z., Chen, W., Glazebrook, J., Chang, H. S., Han, B., Zhu, T., Zou, G., and Katagiri, F. 2003. Quantitative nature of Arabidopsis responses during compatible and incompatible interactions with the bacterial pathogen Pseudomonas syringae. Plant Cell 15:317-330.

Thomma, B. P., Eggermont, K., Penninckx, I. A., Mauch-Mani, B., Vogelsang, R., Cammue, B. P., and Broekaert, W. F. 1998. Separate jasmonate-dependent and salicylate-dependent defense-response pathways in Arabidopsis are essential for resistance to distinct microbial pathogens. Proc. Natl. Acad. Sci. U. S. A. 95:15107-15111.

Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J., and Katagiri, F. 2009. Network properties of robust immunity in plants. PLoS Genet. 5:e1000772.

Van Eck, L., Schultz, T., Leach, J. E., Scofield, S. R., Peairs, F. B., Botha, A. M., and Lapitan, N. L. 2010. Virus-induced gene silencing of *WRKY53* and an inducible phenylalanine ammonia-lyase in wheat reduces aphid resistance. Plant Biotechnol. J. 8:1023-1032.

van Esse, H. P., Fradin, E. F., de Groot, P. J., de Wit, P. J., and Thomma, B. P. 2009. Tomato transcriptional responses to a foliar and a vascular fungal pathogen are distinct. Mol. Plant-Microbe Interact. 22:245-258.

VanLoon, L. C., Bakker, P, A. H. M. and Pieterse, C. M. J. 1998. Systemic resistance induced by rhizosphere bacteria. Annu. Rev. Phytopathol. 36: 453–483.

Wirthmueller, L., Zhang, Y., Jones, J. D., and Parker, J. E. 2007. Nuclear accumulation of the Arabidopsis immune receptor RPS4 is necessary for triggering EDS1-dependent defense. Curr. Biol. 17:2023-2029.

Yang, B., Jiang, Y., Rahman, M. H., Deyholos, M. K., and Kav, N. N. 2009. Identification and expression analysis of WRKY transcription factor genes in canola (Brassica napus L.) in response to fungal pathogens and hormone treatments. BMC Plant Biol. 9:68.

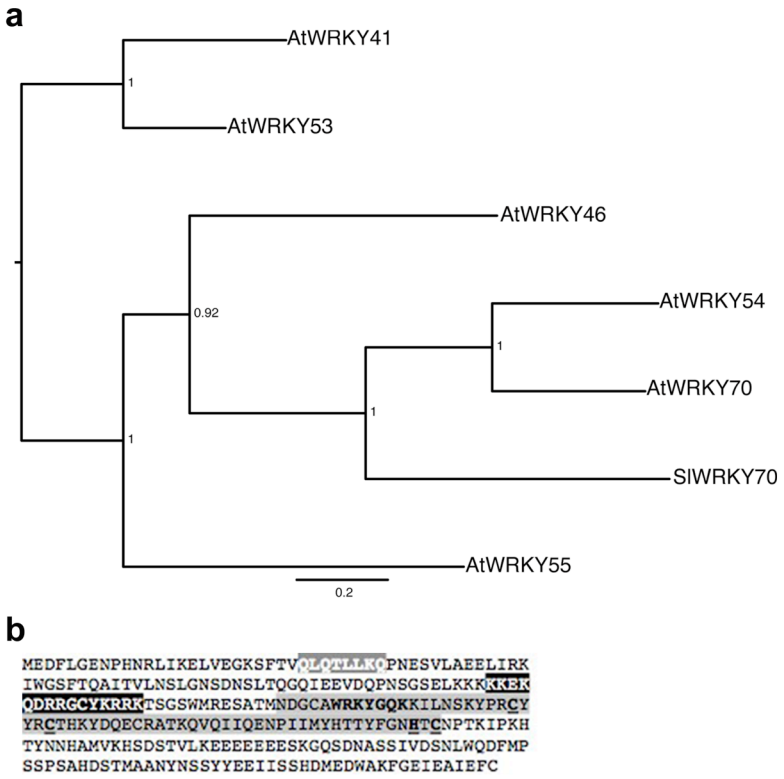


Fig. 1.1A-B (A) Phylogenetic tree of WRKY group III. Sequences of the Arabidopsis WRKY proteins (AT4G11070, AT4G23810, AT2G46400, AT2G40750, AT3G56400, and AT2G40740) and the tomato WRKY unigene (SGN-U582610) were manually aligned. Trees were constructed with MrBayes under mixed protein substitution model running 4 chains and 3 runs for 100,000 generations, though the runs converged (average stddev of split frequencies <0.01) after 30,000 generations. Scale bar indicates the number of substitution per site. (B) Amino acid sequence of *Sl*WRKY70. Features conserved between SlWRKY70 and AtWRKY70 (Knoth et al. 2007) are highlighted and include the WRKY domain (highlighted in light grey), a Q-rich region (printed in white and highlighted in dark grey) as well as a putative bipartite nuclear localization signal identified by prosite (http://www.expasy.org/cgi-bin/scanprosite), printed in white and highlighted in black). The WRKYGQK motif, that is nearly invariant in all WRKY domains, as well as cysteine and histidine residues of the zinc-finger motif conserved in group III WRKYs are bold and underlined.

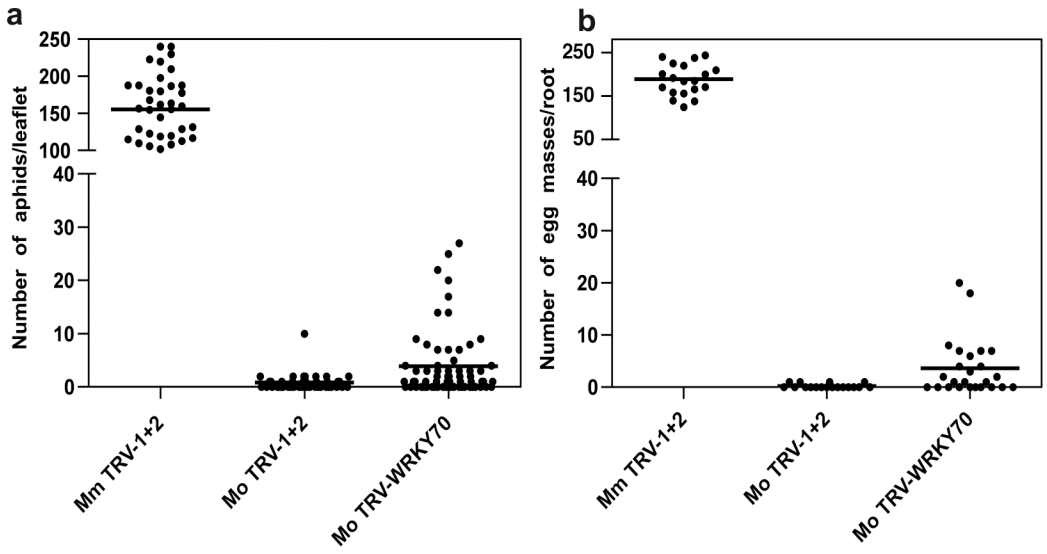


Fig. 1.2A-B TRV-based virus-induced gene silencing of tomato *WRKY70* and assessment of its role in *Mi-1*-mediated resistance against potato aphid and RKN. Tomato cv. Moneymaker (Mm; *mi-1*/*mi-1*) or Motelle (Mo; *Mi-1*/*Mi-1*) agroinfiltrated with TRV1 plus TRV2 empty vector (TRV1+2), and cv. Motelle agroinfiltrated with TRV1 plus TRV-WRKY70 silencing *SlWRKY70*. (A) Potato aphid survival on leaves of control and silenced plants. Circles represent the number of live aphids on a single leaflet. (B) Nematode reproduction on roots of control and silenced plants. Circles represent the number of egg masses on a single root. Experiments in (A) and (B) were repeated once with similar results. Data from a single experiment is presented. The number of aphids per leaflet or egg masses between *Sl*WRKY70-silenced and non-silenced Motelle plants was highly significantly different ($P < 0.005$) for each of the replicates.

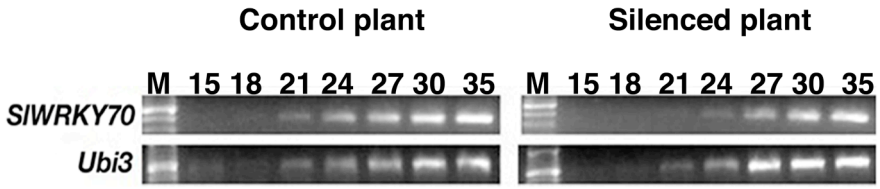


Fig. 1.3 Effect of TRV-VIGS on transcript level of *SlWRKY70* in control and silenced tomato cv. Motelle (*Mi-1*/*Mi-1*) leaves. Leaflets from Motelle plants, agroinfiltrated with TRV-WRKY70 construct showing attenuation in *Mi-1* resistance were collected for cDNA synthesis and *SlWRKY70* transcript levels were evaluated using semi-quantitative PCR. cDNAs from leaflets of Motelle plants agroinfiltrated with TRV2 empty vector were used as control. PCR amplification from cDNA from a single representative sample is presented. Amplification of the tomato ubiquitin *Ubi3* gene was used as an internal control for equal cDNA use from control and silenced plants. PCR cycles are indicated on the top of ethidium bromide stained 1.5% agarose gels. Lane M indicates DNA ladder.

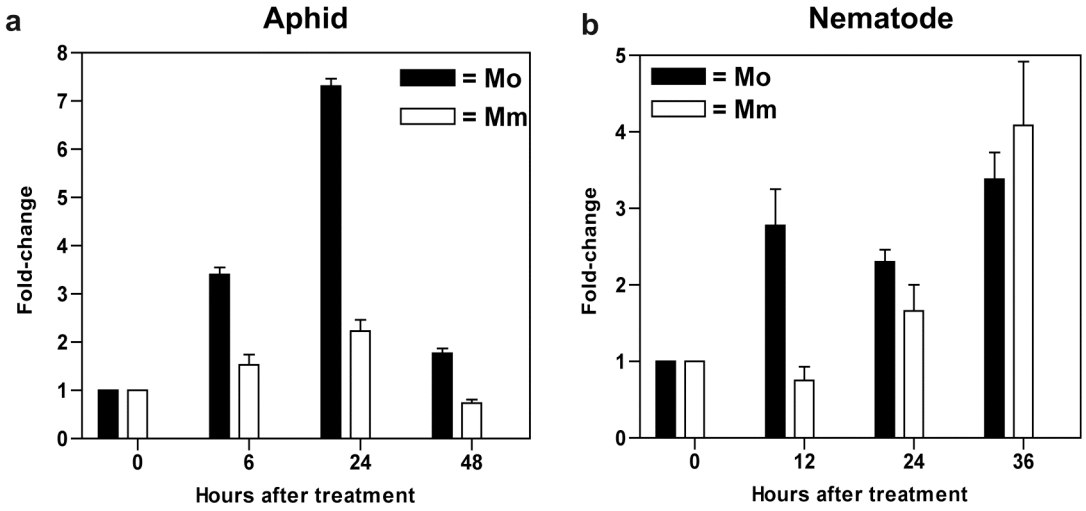


Fig. 1.4A-B *SlWRKY70* temporal expression in tomato leaves and roots after (A) potato aphid infestation or (B) RKN inoculation, respectively. cDNAs from leaflets and roots of Motelle (*Mi-1*/*Mi-1*) and Moneymaker (*mi-1*/*mi-1*) plants subjected to a time-course exposure to aphids and nematodes, respectively, were used for *SlWRKY70* transcript evaluation using quantitative PCR. Values represent means of three technical replicates, normalized relative to the internal control *Ubiquitin* and calibrated to the expression in the TRV control sample. Bars represent standard error of means.

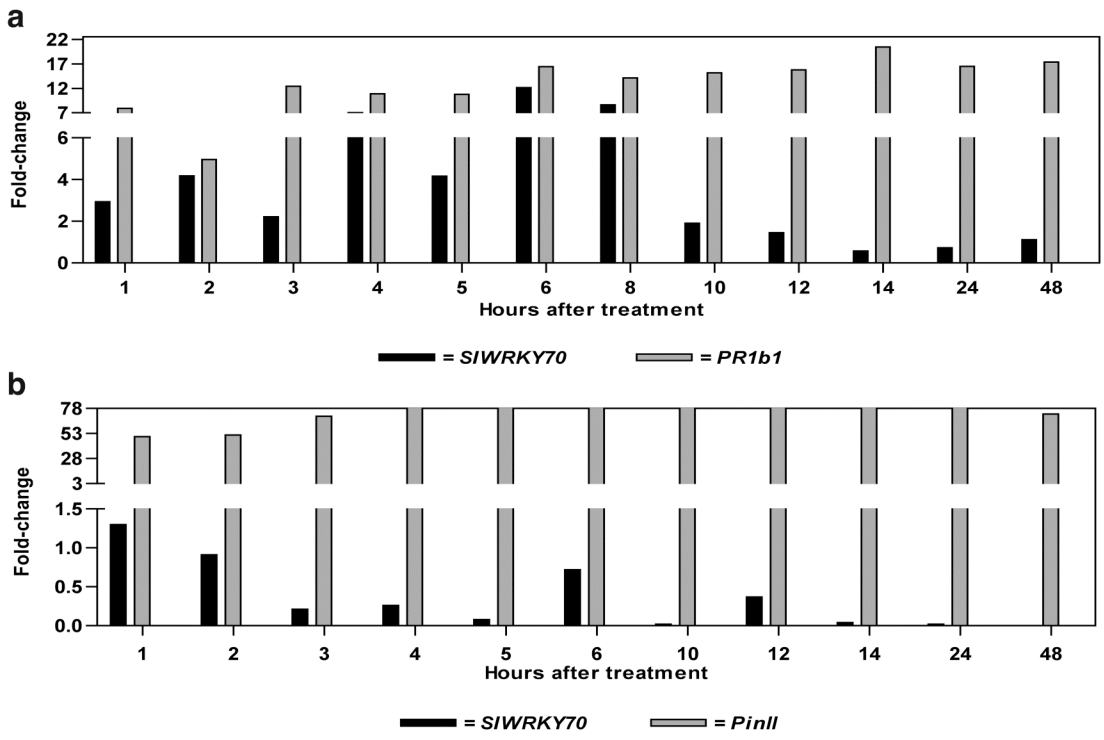


Fig. 1.5A-B Semi-quantitative RT-PCR analysis of *SlWRKY70* transcript levels after salicylic acid (A) or methyl jasmonate (B) hormone treatments. cDNAs synthesized from leaflets of Motelle (*Mi-1/Mi-1*) treated with SA or JA and collected at 0, 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 24, and 48 h after treatment (hat) were used for *SlWRKY70* transcript evaluation. Values were normalized relative to the internal control *Ubiquitin* and calibrated to the expression in the 0 h time-point. *PR1b1* and *PinII* were used as markers for induction of SA and JA signaling pathways, respectively.

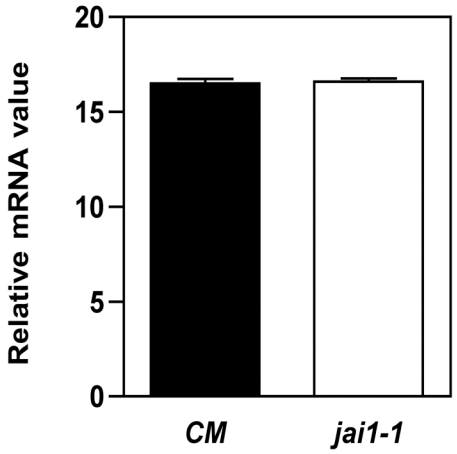


Figure 6

Fig. 1.6 *SlWRKY70* basal expression analysis in wild type cv. Castelmart (CM) and *jai1-1* mutant tomato plants. cDNA synthesized from leaflets of CM and *jai1-1* mutant plants were used for *SlWRKY70* transcript evaluation using quantitative PCR. Values represent means of three biological replicates normalized relative to the internal control *Ubiquitin* and calibrated to the expression in the wild type CM. Bars represent standard error of means.

CHAPTER TWO

A systemic phylogenetics-based nomenclature for WRKY transcription factors.

Abstract

WRKY transcription factors are encoded by large gene families in higher plant species. Based on conserved sequence motifs and phylogenetic relationships the 74 members of the *Arabidopsis thaliana* WRKY family have previously been classified into seven groups/subgroups. Since completion of the *A. thaliana* genome sequence, numerous additional plant genomes have been sequenced and WRKY genes identified. However, often the naming of new *WRKY* genes occurred in a random fashion. Here we report on the annotation of the WRKY families in 15 completed plant genome sequences. In depth analysis of the WRKY transcription factor sequences identified multiple conserved sequence motifs. The incorporation of this information in the multiple sequence alignment process resulted in considerable improvement of the phylogenetic trees manifested by highly resolved branches with significant posterior probability values. Using a combination of phylogenetic relatedness and presence or absence of conserved motifs, the group I and group III WRKYs were divided into five and four subgroups, respectively. Moreover, based on additional structural relatedness among predicted WRKY transcription factors, in each subgroup the members were assigned into types. Collectively this information allowed for the desig of a systematic nomenclature for the WRKY transcription factor family that allowed inferred orthology relationships to be determined. The proposed WRKY nomenclature will enable systematic naming of WRKY transcription factors in additional genomes. In addition, the proposed nomenclature

can be refined and expanded and may also serve as a model for the naming of TFs from other families.

Introduction

Plants and other organisms have the ability to respond to a multitude of external and internal stimuli by comprehensive transcriptome changes. Transcription factors (TFs), which specifically bind to the promoters of target genes and affect their rate of transcription, are of central importance for the coordination of such transcriptional reprogramming (Ramirez and Basu 2009). The genome of the model eudicot plant *Arabidopsis thaliana* (Arabidopsis) encodes for more than 1500 transcription factors belonging to 64 different families with some families consisting of more than 100 members (Guo et al. 2005). One of the largest families of plant TFs are WRKY proteins (Riechmann et al. 2000; Wu et al. 2005).

The WRKY transcription factor family is defined by the ~60 amino acid WRKY DNA-binding domain containing the nearly invariant motif "WRKY" and a conserved cysteine-histidine array of zinc ligands. The gene family is present throughout the plant kingdom including multicelluar and unicellular lineages such as green algae, but is absent in prokaryotes, fungi and metazoans. The 74 Arabidopsis WRKY members have been categorized into three major structural classes: Groups I, II, and III (Eulgem et al. 2000). Generally members of group I have two WRKY DNA-binding domains, while those of groups II and III harbor a single copy. The WRKY DNA-binding domains of group I and II members feature a conserved $C_{x4-5}C_{x22-23}H_xH$ pattern of zinc ligands, while those of group III members contain the zinc-finger motif $C_{x7}C_{x23}H_xC$. The single WRKY DNA-binding domains found in group II

and III members are more closely related to the C-terminal WRKY DNA-binding domain in the (two domain-containing) group I members (Eulgem et al. 2000; Zhang and Wang 2005). Several studies have shown that the C-terminal, but not the N-terminal WRKY DNA-binding domain of group I members serve as specific DNA-binding domains (de Pater et al. 1996; Eulgem et al. 1999) and structural analyses revealed specific physicochemical interactions between the WRKYGQK residues of C-terminal WRKY DNA-binding domain and base pairs of their cognate DNA target site (Duan et al. 2007; Yamasaki et al. 2005). Based on further phylogenetic analyses of WRKY DNA-binding domain sequences and the presence of additional conserved primary-structural features, group II was divided into the five subgroups A, B, C, D and E (Eulgem et al. 2000). This general categorization has also been adopted to classify WRKYs of other plant species (Rushton et al. 2010; Wu et al. 2005).

The WRKY DNA-binding domain is a representative of the WRKY-GCM1 fold superfamily of DNA binding domains that are present in various classes of eukaryotic TFs and transposases (Babu et al. 2006). The domain is characterized by a core structure of four anti-parallel beta sheets stabilized by a zinc-finger and multiple amino acid side chain interactions. Several lines of evidence suggest that TF families containing this domain evolved from ancestral transposases. In plants WRKY-GCM1-fold TFs are represented by the WRKY family and the NAM family. TFs of the latter one, however, have secondarily lost the zinc-finger structure. The WRKYGQK motif seems to be exclusively present in the DNA-interfacing region of

WRKY TFs and is replaced by unrelated sequences in the corresponding regions of other WRKY-GCM1-fold proteins (Babu et al. 2006).

While most non-plant eukaryotes appear to lack *WRKY* genes, the existence of a single *WRKY* member was reported in the slime mold *Dictyostelium discoideum* and the intestinal parasite *Giardia lamblia* (Ulker and Somssich 2004; Zhang and Wang 2005). *G. lamblia* is an ancient protist that diverged about 1,500 million years ago (Mya) from the lineages containing the metazoans, fungi (Opisthokont) and plants (Zhang and Wang 2005). The lineage of slime molds split after these crown eukaryotes emerged and share an ancestor with the plant ancestor between 948 and 1126 Mya (Berney and Pawlowski 2006). Thus, most likely an ancestral *WRKY* gene emerged prior to the establishment of the plant lineage and was secondarily lost in the Opisthokont clade (Zhang and Wang 2005).

While the unicellular green alga *Chlamydomonas reinhardtii* harbors only a single *WRKY* gene, the family substantially expanded in plant species (Eulgem et al. 2000; Wu et al. 2005). The single *WRKY* genes of *D. discoideum, G. lamblia and C. reinhardtii* encode group I members with two WRKY DNA-binding domains (Zhang and Wang 2005). The moss *Physcomytrella patens* contains the whole spectrum of WRKY groups, including group III (Rensing et al. 2008) indicating the family diversified early in land plant evolution.

Similarly the gymnosperm *Pinus monticola* features a wide variety of WRKYs representing group I and all of the subgroups of group II (Liu and

Ekramoddoullah 2009), but lacks members of group III, which may have gotten secondarily lost in at least this gymnosperm lineage after the separation of the monocot and dicot lineages about 140-150 Mya (Chaw et al. 2004). Previous phylogenetic analyses revealed that group II is not a monophyletic group but is comprised of sister clades/subgroups IIA and IIB as well as IID and IIE (Zhang and Wang 2005). Members of subgroup IIC cluster separately from other group II members and appear very diverse. Of all group II clades, subgroup IIC is most closely related to group I, while IID and IIE are most closely related to group III (Eulgem et al. 2000; Zhang and Wang 2005).

Taken together these observations suggested the following evolutionary scenario proposed by Zhang and Wang (Zhang and Wang 2005). An ancestral WRKY TF with one WRKY DNA-binding domain emerged early in unicellular eukaryotes. Prior to the separation of the *G. lamblia* lineage from those leading to plants, animals and the slime mold ancestor, this domain was duplicated giving rise to a group I prototype. Likely a group I derivative that had secondarily lost its N-terminal WRKY DNA-binding domain emerged early in the land plants after the divergence from the green algae ancestor. This hypothetical WRKY must have given rise to the diversification of the WRKY family manifested in the presence of subgroups IIA, IIB, IIC, IID and IIE in all land plants and group III in most land plant lineages. Subgroup IIC is more ancestral than the other subcategories of group II as indicated by the breadth of representatives and diversity of the group.

Originally identified by their ability to bind to pathogen-responsive W box promoter elements [(T)TGACC/T] (Rushton et al. 1996), genetic evidence has primarily implicated members of the WRKY family in the regulation of plant immunity (Eulgem and Somssich 2007). Numerous recent reports showed that some of these TFs are also involved in responses to abiotic stresses, such as heat, drought, high salinity, as well as developmental programs including seed development, trichome formation and senescence (Rushton et al. 2010). Often individual WRKY TFs seem to contribute to multiple biological functions (Rushton et al. 2010).

Besides the WRKY DNA-binding domain, several other conserved amino acid sequence motifs hav been noted in the WRKY TFs and are associated with defined molecular functions (Eulgem and Somssich 2007). For example, coiled-coil (CC) structures present in IIA and IIB WRKYs mediate homo- and hetero-dimerization (Robatzek and Somssich 2002; Xu et al. 2006). "Motif C" is conserved among IID WRKYs and serves as a calmodulin-binding domain (Park et al. 2005), and "motif D", which is found in group I members, can be phosphorylated by MAP kinases (Andreasson et al. 2005). In addition, short basic motifs reminiscent of nuclear localization signals as well as acidic-, glutamine- or proline-rich regions similar to known transcriptional activation domains are frequently present in WRKY primary structures (Eulgem et al. 2000).

During the past 10 years, over 300 descriptive and functional studies have been published on WRKY TFs. Several of these reports identified orthologous

WRKY TFs in different plant species and showed that structural relatedness correlates with conservation of their biological roles (Berri et al. 2009; Bhattarai et al. 2010). However, no universal nomenclature exists for members of this family and there are no systematic guidelines for designating orthologous *WRKY* genes in different plant species. While often the naming of new *WRKY* genes appears to be based on their order of identification (Encinas-Villarejo et al. 2009; Park et al. 2006; van Verk et al. 2008), some authors resorted to using the Arabidopsis nomenclature of this family as a reference (Bhattarai et al. 2010; Levee et al. 2009; Liu et al. 2010; Yao et al. 2011). As a result, the names of hundreds of WRKYs are inconsistent, complicating communication among researchers working on these TFs in different plant species.

Here the annotation of WRKY TFs from 15 plant species for which full genome sequences are available is reported. This analysis included various monocots and dicots, as well as the moss *Physcomytrella patens,* and the green algae *C. reinhardtii*. Based on structural relatedness between predicted WRKY TFs, a systematic nomenclature for this family that allowed recognition of inferred orthology relationships was designed. The WRKY nomenclature proposed can be refined and expanded and may also serve as a model for the naming of TFs from other families.

Material and Methods

Gene retrieving and prediction

The WRKY TFs of Arabidopsis and *O. sativa* were retrieved from TAIR and NCBI websites, respectively (URL3; URL4) based on their gene names. The complements of WRKY TFs of *Vitis vinifera*, *G. max*, *Medicago truncatula*, *Populus trichocarpa*, *Brachypodium distachyon*, *Sorghum bicolor, Zea mays*, *C. reinhardtii*, *Chlorella* sp. NC64A, and *Ostreococcus tauri* were predicted by performing TBLASTN searches (URL5) with the *At*WRKY1 protein sequence against all hypothetical or confirmed transcript sequences of each of the fully sequenced genomes. Genomic sequences of *Carica papaya*, *P. patens*, and *Coccomyxa* sp.C-169 encoding WRKY core motifs (WRKY) were retrieved by performing TBLASTN searches (URL5) with the *At*WRKY1 protein sequence, as predictions of transcripts were not available for these genomes. In each case, a genomic region of 20 kb upstream and 20 kb downstream of the TBLASTN identified locus was used to predict and annotate the respective full-length *WRKY* gene using the online programs Genscan and FGenesh (URL6; URL7). In some cases, the retrieved and predicted *WRKY* genes that appeared incomplete were re-annotated. All WRKY sequences used in this study were mapped against genomic sequences (URL5) to avoid duplicate naming of alleles and alternative splice forms.

Phylogenetic analysis

Multiple alignments of each gene family were generated with T-Coffee (Notredame et al. 2000) and automatically trimmed with Trimal (Capella-Gutierrez et al. 2009) using the –automated1 setting. This approach removed poorly aligning regions. Phylogenetic trees were constructed with MrBayes 3.2-cvs (Ronquist and Huelsenbeck 2003) with parameters of aamodel=fixed and 5 million generations over 4 runs of 4 chains each. The tree search was allowed to stop if runs converged at posterior value difference of < 0.01.

Results and discussion

Identification of conserved motifs and classification

The WRKY TF sequences used in this study were retrieved or predicted by performing TBLASTN searches. The *At*WRKY1 protein sequence [at2g04880] as the query against transcript or genomic sequences of each of the fully sequenced genomes (URL1; URL2). The WRKY TFs of Arabidopsis and *Oryza sativa* were retrieved using existing gene annotation and gene names from TAIR and NCBI websites, respectively (URL3; URL4). Using BLASTP against Arabidopsis WRKY sequences, we initially assigned all the WRKY sequences from 15 plant species to groups/subgroups based on the *Arabidopsis* WRKY family classification (Figure 1) (Eulgem et al. 2000). These assignments were further confirmed by constructing a phylogenetic tree with all the WRKY sequences from the 15 plant species and Arabidopsis using the conserved WRKY DNA-binding domain. Based on phylogenetic analyses separately performed for group I and group III members and the presence of conserved motifs (CMs), identified by the MEME motif discovery tool (Bailey and Elkan 1994), group I and III WRKY TFs were further designated to subgroups. For each of the subgroups within groups I, II and III, further types of closely related family members were defined using phylogenetic relatedness and type-specific CMs as criteria.

As outlined below, simpler approaches that directly aligned group specific sequences to construct phylogenetic trees were unsuccessful in assessing orthology relationships among the WRKY sequences from multiple plant species. Our current approach used initial detailed CM analysis, for consensus CM identification within group specific sequences, prior to sequence alignment. Sequences that did not have the respective consensus CMs were not included in the phylogenetic analysis. The remaining sequences were trimmed before performing the alignment for phylogenetic tree construction so that the alignments have the respective consensus CMs at their N- and C-termini whenever possible. This approach allowed for better-resolved trees and proper inference of orthology among each of the group phylogenies using sequences from the 15 different plant species. Members not included in the phylogenetic analysis were secondarily assigned to a subgroup or type based on the presence of additional subgroup and type-defining CMs.

A universal systemic nomenclature for WRKY TFs

The categorization of WRKYs in distinct groups, subgroups and types enabled the design of a universal nomenclature for this TF family that allowed easy recognition of possible orthology relationships between these TFs throughout the plant kingdom. Each WRKY member was designated by a Roman number indicating its group, a capital letter indicating its subgroup within this group and an Arabic number indicating its type within each subgroup. Paralogs within each type are further differentiated by lower case letters. For example, the three structurally close-related Arabidopsis members of group I,

originally termed *At*WRKY25 [at2g30250], *At*WRKY26 [at5g07100] and *At*WRKY33 [at2g38470] will be named *At*WRKY-IC2a, *At*WRKY-IC2b and *At*WRKY-IC2c, respectively (Table S1 columns B and C rows 45, 46 and 57). Their orthologs from other species will share the acronym "IC2" to indicate their structural and evolutionary relationship, while a two-letter prefix preceding "WRKY" identifies the respective species (e.g. *Os*WRKY for rice WRKY members) (Table S1 column C rows 56, 65 and 70).

A small number of WRKY sequences considered in this study could not be unequivocally classified at the subgroup or type level by our system. We provisionally named them using the capital letter "X", These acronyms are followed by a lower case letter to discriminate family members from the same species that are unclassified (Table S1). For example, a group I member from *Glycine max* that is unclassified was provisionally named *Gm*WRKY-IXa.

Topology of group I

A critical prerequisite for phylogenetic protein analyses is the proper alignment of the respective amino acid sequences (Pace 2009). First we produced an alignment with the full-length sequences of all putative group I members. A phylogenetic analysis of the alignment did not result in a tree with sufficient resolution and well-supported branches (Figure 2). This was due to the fact that the automatic trimming of the full length sequences with Trimal (Capella-Gutierrez et al. 2009), to remove the poorly aligning regions, left only the majority of the WRKY DNA-binding domains (total of

104 amino acids) for phylogenetic analysis. The WRKY DNA-binding domains showed high level of sequence identity and similarity among the different group I members.

In general group I members contained two WRKY DNA-binding domains representing a defining feature of this group. The single WRKY DNA-binding domain found in group II and III members are more closely related to the C-terminal WRKY DNA-binding domain in the (two domain containing) group I members. However, a small number of group I WRKY sequences have only the N-terminal or C-terminal WRKY DNA-binding domain (Appendix 1). The N-terminal WRKY DNA-binding domain of group I members features a perfectly conserved glycine amino acid differentiating it from the group I C-terminal and groups II and III WRKY DNA-binding domains (Appendix 1 indicated by ‡). Few group I WRKY sequences have only the C-terminal WRKY DNA-binding domain, which is more closely related to the WRKY DNA-binding domain of groups II and III. These group I sequences have a universally conserved arginine just preceding the “WRKY motif”, within their DNA-binding domain that is also present in subgroup IIC members (Appendix 1, indicated by ‡). However, the absence of a conserved glycine after the “WRKY motif” in subgroup IIC members distinguishes between group I and group IIC C-terminal WRKY DNA-binding domains (Appendix 1, indicated by **¥).**

We constructed a highly resolved phylogenetic tree with the sequences that contained MEME-identified consensus CMs (Figure 3). This tree had most nodes with significant support indicated by posterior probabilities > 0.95 dividing group I into 5 different subgroups (IA, IB, IC, ID and IE) and further defining four different types within subgroup IC (IC1, IC2, IC3 and IC4). Using this sample tree each of the remaining group I members were assigned to one of the five subgroups and types based on the presence of CMs unique to each subgroup and type. To confirm these secondary assignments, another phylogenetic tree including a subset of the secondarily assigned sequences with the alignment of the sample tree sequences was constructed (Figure 4). This overall strategy enabled generation of a tree of sufficient quality and accurately categorized as many family members as possible. A similar strategy was pursued for the analyses of the other WRKY groups and subgroups.

CMs defining sub-clades of group I

Table 1 lists CMs defining all WRKY subgroups and types proposed in this study. Motif D and a motif enriched with charged amino acids, previously referred to as "basic-motif-2" (Eulgem et al., 2000) are the most prominent CMs within group I. Motif D is present in the N-terminal region of members of the large subgroup IC. As basic-motif-2 contains several conserved acidic residues besides basic ones, we renamed this motif to charged-1. Charged-1 is located between the two WRKY DNA-binding domains of most group I members. Motifs D and charged-1 can be

conveniently used as major criteria defining the five subgroups of group I. Motif D is generally present in most members of subgroup IC, while this motif is strictly absent in subgroups IA, IB, ID, and IE. Subgroups IA and ID both lack motif D and contain charged-1, but members of the latter clade do not feature the YKPxAK and VASR motifs, which are unique for subgroup IA. Thus, subgroup IA is unequivocally defined by the presence of charged-1, YKPxAK and VASR, as well as the absence of motif D. Subgroup IB is defined by the absence of motif D and the presence of either [Q-rich-1, FxxLLxG and P-rich-1] or VEEV motifs. Subgroup ID members generally harbor the CM PSIIR and share the conservation of hydrophobic residues at a single position at the C-terminal end of their first WRKY DNA-binding domain (Appendix 1 indicated by an arrow). The majority of the members of the other subgroups have positively charged amino acids at that position. Besides being defined by the absence of motif D, subgroup IE members harbor the SvxxS CM and feature a highly conserved negatively charged amino acid (aspartic acid[D] or glutamic acid[E]) followed by two proline (P) residues at the C-terminal of their first WRKY DNA-binding domain (Appendix 1 indicated by arrows).

Additional conserved motifs allowed for the definition of distinct types within the subgroup IB and the complex subgroup IC. Besides motif D conserved among all IC members, the Q-rich-1, P-rich-1, and FxxLLxG motifs define IC1-type WRKYs. Moreover, members of this type do not contain the SPTTG CM present in types IC2, IC3, and IC4 members. Motifs unique for IC2-type WRKYs are RTGSG and ENSS.

An additional motif defining IC2-type WRKYs is a derivative of charged-1, termed charged-2 with an invariant tryptophan (W) inserted after lysine-arginine (K-R) (Appendix 1 indicated by an arrow). Type IC3 WRKYs are defined by CMs DxSS and LxER. Type IC4 WRKYs are defined by the VxxxE and ARYK motifs.

Group I appears to be the most ancestral clade of extant WRKYs, as it features members from a phylogenetically diverse collection of species including single-celled algae and the single-cell diplomonad *G. lamblia*. Motif D, which defines subgroup IC, is missing in the single-celled basal lineages of algae and *G. intestinalis* suggesting that this CM was gained later in evolution. Subgroups IB and ID appear to be of more recent origin and are derived from subgroup IC (Figure 3).

Refined classification of group II

Group II has been subdivided into five subgroups based on variations of their WRKY DNA-binding domain primary structure and CMs (Eulgem et al. 2000). While previous analyses showed that this group is not monophyletic (Rushton et al. 2010; Zhang and Wang 2005), group II members share the occurrence of a single $C_{X4-5}C_{X22-23}H_XH$-type zinc finger-containing WRKY DNA-binding domain that is most closely related to the C-terminal WRKY DNA-binding domain of group I members. They are distinguished from the group I members with only a C-terminal WRKY DNA-binding domain and by the absence of a conserved glycine (G) within the WRKY DNA-binding domain.

Subgroup IIA

The majority of subgroup IIA sequences featured a putative coiled-coil domain (CC-A) at their N-termini and an alanine-rich motif (alanine-rich-1) at their C-termini (Appendix 2, Table S2). The coiled-coil motif is characterized by invariant heptad repeats of bulky hydrophobic amino acids. An alignment of the sequences stretching from the beginning of CC-A to the end of alanine-rich-1 (Appendix 2) was used to infer a phylogenetic tree for subgroup IIA (Figure 5).

The tree is subdivided into four clades supported by high posterior support (> 0.90) and specifying four WRKY types IIA1, IIA2, IIA3 and IIA4. The IIA1 and IIA4 are exclusively monocot sequences, while those representing types IIA2 and IIA3 are exclusively dicot sequences. The fact that each of the monocot-specific clades is a sister clade of a dicot specific one suggests the existence of two ancestral group IIA WRKYs that separately underwent further diversification after the division between monocots and dicots. All three *Arabidopsis* IIA members are classified as type IIA2 [*At*WRKY18 (*At*WRKY-IIA2a), *At*WRKY60 (*At*WRKY-IIA2b) and *At*WRKY40 (*At*WRKY-IIA2c)] and there are no Arabidopsis members of IIA3, the other dicot-specific group.

IIA1-type sequences are defined by the SINS motif. The IIA1-type sequences further share a perfectly conserved alanine in the C-terminal half of the WRKY DNA-binding domain (Appendix 2 indicated by an arrow). Type IIA2 sequences are defined by the CM ESSSTDE, while IIA3-type sequences typically contain a highly conserved

glutamine preceding the WRKY DNA-binding domain (Appendix 2 indicated by an arrow). This type also harbors a perfectly conserved tyrosine within the alanine-rich-1 motif (Appendix 2 indicated by an arrow) and a conserved cysteine in the zinc-finger region of the WRKY DNA-binding domain. Members of type IIA4 feature the CMs EGGS and ECTS, which are preceding their WRKY DNA-binding domain (Appendix 2). A conserved motif LDL-1 is absent in IIA4-type members while present in the remaining three types (IIA1, IIA2, and IIA3).

Subgroup IIB

Previous phylogenetic analyses of WRKY sequences from multiple plant species revealed a sister relationship of the subgroups IIA and IIB, representing two sub-clades of a larger monophyletic group II clade (Rushton et al. 2010; Wu et al. 2005; Zhang and Wang 2005). Consistent with this, there are some primary structural features specifically conserved between these two subgroups. Subgroup IIB sequences contain derivatives of the subgroup IIA CC-A, LDL-1 and alanine-rich-1 motifs, which were termed CC-B, LDL-2 and alanine-rich-2, respectively (Appendix 3, Table S3). We generated a high quality sample tree using the sequences stretching between CC-B and alanine-rich-2 motifs. (Figure 6).

Based on high bootstrap values supporting key branchpoints and the presence of CMs, we subdivided this subgroup into four separate clades representing types IIB1, IIB2, IIB3 and IIB4. IIB1-type WRKY sequences share the highly conserved VPRQF, RxxLPC

and QSKF motifs (Appendix 3). Members of the IIB2-type WRKYs share a perfectly conserved proline directly preceding the WRKY DNA-binding domain (Appendix 3 indicated by an arrow) and the conserved PYAS motif. The acidic-2 motif and a highly conserved cysteine preceding the WRKY DNA-binding domain are specific for type IIB3. Also in the WRKY DNA-binding domain of the IIB-3-type members, the highly conserved methionine is mostly replaced by valine or other non-polar amino acids. IIB4-type WRKYs are defined by the highly conserved MMPLP and FLAR motifs. WRKYs of this type appear only to be present in monocots (Figure 6).

Subgroup IIC

No single CM defines subgroup IIC (Appendix 4, Table S4). We used an alignment with sequences of the identified consensus CMs to construct a high quality sample tree for subgroup IIC (Figure 7). The resulting sample tree suggested the existence of seven distinct WRKY types within this subgroup. Type IIC1 members were defined by the CMs YSGPTI, KYTxK and LLxD (Appendix 4). The later CM is also found in IIC6 type members. Type IIC2 members are defined by the invariant stretch RMVI within the WRKY DNA-binding domain and a serine residue immediately following the WRKY DNA-binding domain (Appendix 4 indicated by an arrow). Type IIC3 is defined by the conserved motif FEHIL (Appendix 4). Members of this type also harbor a perfectly conserved glutamic acid in the zinc-finger region of the WRKY DNA-binding domain. The sequences of the other members of this subgroup do not

have a negatively charged amino acid at this position. Most type IIC4 WRKYs contain the LxPxL and DWxxL motifs (Appendix 4).

The members of IIC5-type are defined by a highly conserved serine (Appendix 4) in the center of the WRKY DNA-binding domain. The majority of the remaining subgroup IIC members have threonine at this position and none has serine (Appendix 4 indicated by an arrow). Moreover, members of this type have two highly conserved asparagine residues in the N-terminal half of their WRKY DNA-binding domain (Appendix 4 indicated by arrows). No other IIC sequences show this combination of conserved residues within their WRKY DNA-binding domain. Type IIC6 members are defined by the highly conserved TPNSS and LLxD motifs and the absence of YSGPTI and KYTxK motifs.

All type IIC7 members are sequences from *P. paten*s. Although they clustered together in a well-defined subclade in our sample tree (Figure 7), we identified no obvious CMs defining this type.

Subgroup IID

The basic-1 motif described previously (Eulgem et al. 2000), is conserved among all IID members. The great majority of subgroup IID sequences are characterized by the presence of the CM acidic-3 at their N-termini, and motif C, which serves as a calmodulin-binding domain (Appendix 5, Table S5) (Park et al. 2005). We generated a

high quality sample tree using sequences stretching from the N-terminal of acidic-3 to the C-terminal of their WRKY DNA-binding domain (Figure 8).

This sample tree features four separate clades representing the WRKY types IID1, IID2, IID3 and IID4. The HARF motif, which was previously found to define subgroup IID in Arabidopsis (Eulgem et al. 2000), is broadly present in this subgroup, although IID4-type WRKYs feature a distinct derivative of this motif, we termed HARV (Appendix 5).

IID1-type WRKYs are defined by the motif PxxxP. This type has a perfectly conserved serine immediately following the three positively charged amino acid triplet (Appendix 5 indicated by an arrow) conserved among most IID-types. Moreover, members of this type feature a highly conserved methionine directly preceding the WRKY DNA-binding domain (Appendix 5 indicated by an arrow). Methionine is strictly absent at this position in WRKY sequences of other IID types. IID2-type WRKYs generally share the unique VTLDF and GSVSxG motifs. IID3-type WRKYs are defined by the MSDAA motif and a perfectly conserved asparagine in the N-terminal region of the WRKY DNA-binding domain (Appendix 5 indicated by an arrow). All the members of the other IID-types have negatively charged amino acids at this position. IID4-type members are defined by the motifs HARV, SxxGS and FHLxG. IID5-type WRKYs feature the unique AGVFLE and KCAI motifs as well as a perfectly conserved serine in the zinc-finger region of the WRKY DNA-binding domain (Appendix 5 indicated by an arrow). No WRKYs of other IID-types have a serine at this position.

Subgroup IIE

As in the case of subgroups IIA and IIB, the subgroups IID and IIE represent two subclades of a larger monophyletic clade (Rushton et al. 2010; Zhang and Wang 2005). Consistent with this, a derivative of the basic-1, which we termed basic-2, is conserved among all IIE members (Appendix 6, Table S6).

The majority of subgroup IIE sequences contain the motif E at their N-termini and/or the acidic-4 motif at their C-termini. Sequences of subgroup IIE WRKYs harboring both or either of these two CMs were used for the sample tree construction (Figures 9). The sample tree showed clear subdivision of this subgroup into three clades representing the WRKY types IIE1, IIE2 and IIE3.

IIE1-type WRKYs are defined by the presence of the GExxxP, WFxD, LESP and acidic-4 motifs but unlike the majority of IIE2- and IIE3-type WRKYs IIE1 members lack the N-terminal motif E (Appendix 6). Morevover, the CM basic-2 of the IIE1-type members contains a perfectly conserved proline (Appendix 6 indicated by an arrow). The IIE2-type members have a highly conserved tryptophan, within the motif E, which is strictly missing in the IIE3-type members. In addition they lack the acidic motifs conserved among the IIE1- and IIE3-type members (acidic-4 and acidic-5, respectively).

Instead of GExxxP, which precedes the WRKY DNA-binding domain of IIE1-type sequences, the IIE3-type members have the CM RxxxTG. IIE3-type members also have the CMs DPFxxxxDP and acidic-5, in addition to a perfectly conserved proline-alanine-proline trimer and a perfectly conserved glycine in their WRKY DNA-binding domain (Appendix 6 indicated by arrows).

Topology of group III

Group III is defined by the zinc-finger motif $C_{x7}C_{x23}H_xC$ within the WRKY DNA-binding domain. The majority of group III members have the hydrophobic-1 CM at their N termini (Appendix 7, Table S7). We generated a high quality sample tree using the sequences stretching from hydrophobic-1 to the C-terminal of their WRKY DNA-binding domain (Figure 10). Consistent with previous studies focused on Arabidopsis and rice (Zhang and Wang 2005) and similar to subgroup IIA, this group features two major clades, each of which contains a monocot and a dicot specific subclade. This suggests the existence of two ancestral group III WRKYs that further diverged after the separation between monocots and dicots (Zhang and Wang 2005). Based on this clear subdivision, which is supported by nodes with high boostrap values in our tree (Figures 10), we propose to split group III into four subgroups (IIIA, IIIB, IIIC and IIID). Subgroups IIIA and IIID are monocot-specific, while subgroups IIIB and IIIC are dicot-specific.

CMs defining sub-clades of group III

The monocot-specific subgroups IIIA and IIID can be discriminated by the DILGAK CM and a highly conserved serine within the WRKY DNA-binding domain, both of which are present only in IIIA members (Appendix 7 indicated by double arrows). Similarly the two dicot-specific subgroups IIIB and IIIC can be discriminated by the DILGAK CM, in addition to a perfectly conserved serine, a highly conserved basic residue (arginine or lysine) and a negatively charged amino acid within the WRKY DNA-binding domain as well as a conserved glycine before the WRKY DNA-binding domain (Appendix 7 indicated by arrows). These amino acids are absent in the respective positions of subgroup IIIC members.

Additional conserved motifs allowed for the definition of distinct types within subgroups IIIA, IIIC, and IIID. Subgroup IIIA was further divided into three types. Type IIIA1 is defined by the CMs VRVQ, LxAP, TARWT and PHAQ. IIIA2-type members are defined by the CMs LVVT, ENWG and have a perfectly conserved leucine-glycine-glycine-serine tetramer at their N-termini (Appendix 7 indicated by arrows). IIIA3-type members have the CM RVSAVQD. In addition they harbor DGDP tetramer within the WRKY DNA-binding domain and two methionine residues preceding the hydrophobic-1 motif (Appendix 7 indicated by arrows) instead of the two highly conserved leucine residues present in the members of the other types.

The dicot subgroup IIIC was further divided into two types. IIIC1-type members are defined by the absence of two highly conserved glutamic acid residues and the

presence of a perfectly conserved cysteine preceding the hydrophobic-1 CM (Appendix 7 indicated by arrows), while IIIC2-type members are defined by the CMs RRKA and RGCY. The monocot subgroup IIID was further divided into eight types. The CMs QLRA and LETP define the IIID1-type members. IIID2-type members are defined by the CMs KILES and HxGS. IIID3-type members are defined by the CMs RQQVEL and PIKEIL. IIID4-type members are defined by the CM EHEAV. Although IIID5-type members clustered together in a well-defined sub-clade in our sample tree (Figure 10), we could not identify a CM defining this type. Members of this type lack the CMs defining the other types of subgroup IIID. IIID6-type members are defined by the CMs YELIK and SWAT. IIID7-type members are defined by the CM GxxSAA (Appendix 7). IIID8-type members clustered together in a well-defined sub-clade in our sample tree (Figure 10). However, we could not identify a CM defining this type.

References

Andreasson, E., Jenkins, T., Brodersen, P., Thorgrimsen, S., Petersen, N. H., Zhu, S., Qiu, J. L., Micheelsen, P., Rocher, A., Petersen, M., Newman, M. A., Bjorn Nielsen, H., Hirt, H., Somssich, I., Mattsson, O., and Mundy, J. 2005. The MAP kinase substrate MKS1 is a regulator of plant defense responses. EMBO J. 24:2579-2589.

Babu, M. M., Iyer, L. M., Balaji, S., and Aravind, L. 2006. The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. Nucleic Acids Res. 34:6505-6520.

Bailey, T. L., and Elkan, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. 2:28-36.

Berney, C., and Pawlowski, J. 2006. A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record. Proc. R. Soc. Lond. B Biol. Sci. 273:1867-1872.

Berri, S., Abbruscato, P., Faivre-Rampant, O., Brasileiro, A. C., Fumasoni, I., Satoh, K., Kikuchi, S., Mizzi, L., Morandini, P., Pe, M. E., and Piffanelli, P. 2009. Characterization of WRKY co-regulatory networks in rice and Arabidopsis. BMC Plant Biol. 9:120.

Bhattarai, K. K., Atamian, H. S., Kaloshian, I., and Eulgem, T. 2010. WRKY72-type transcription factors contribute to basal immunity in tomato and Arabidopsis as well as gene-for-gene resistance mediated by the tomato *R*-gene *Mi-1*. Plant J. 63:229-240.

Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972-1973.

Chaw, S. M., Chang, C. C., Chen, H. L., and Li, W. H. 2004. Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. J Mol. Evol. 58:424-441.

de Pater, S., Greco, V., Pham, K., Memelink, J., and J, K. 1996. Characterization of a zinc-dependent transcriptional activator from Arabidopsis. Nucleic Acids Res. 24:4624-4631.

Duan, M. R., Nan, J., Liang, Y. H., Mao, P., Lu, L., Li, L., Wei, C., Lai, L., Li, Y., and Su, X. D. 2007. DNA binding mechanism revealed by high resolution crystal structure of *Arabidopsis thaliana* WRKY1 protein. Nucleic Acids Res. 35:1145-1154.

Encinas-Villarejo, S., Maldonado, A. M., Amil-Ruiz, F., de los Santos, B., Romero, F., Pliego-Alfaro, F., Munoz-Blanco, J., and Caballero, J. L. 2009. Evidence for a positive regulatory role of strawberry (*Fragaria x ananassa*) *Fa*WRKY1 and Arabidopsis *At*WRKY75 proteins in resistance. J. Exp. Bot. 60:3043-3065.

Eulgem, T., and Somssich, I. E. 2007. Networks of WRKY transcription factors in defense signaling. Curr. Opin. Plant Biol. 10:366-371.

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. 2000. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 5:199-206.

Eulgem, T., Rushton, P. J., Schmelzer, E., Hahlbrock, K., and Somssich, I. E. 1999. Early nuclear events in plant defence signalling: Rapid activation by WRKY transcription factors. EMBO J. 18.

Guo, A., He, K., Liu, D., Bai, S., Gu, X., Wei, L., and Luo, J. 2005. DATF: a database of Arabidopsis transcription factors. Bioinformatics 21:2568-2569.

Levee, V., Major, I., Levasseur, C., Tremblay, L., MacKay, J., and Seguin, A. 2009. Expression profiling and functional analysis of *Populus* WRKY23 reveals a regulatory role in defense. New Phytol. 184:48-70.

Liu, H., Yang, W., Liu, D., Han, Y., Zhang, A., and Li, S. 2010. Ectopic expression of a grapevine transcription factor *Vv*WRKY11 contributes to osmotic stress tolerance in Arabidopsis. Mol. Biol. Rep. 38:417-427.

Liu, J. J., and Ekramoddoullah, A. K. 2009. Identification and characterization of the WRKY transcription factor family in *Pinus monticola*. Genome 52:77-88.

Notredame, C., Higgins, D. G., and Heringa, J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302:205-217.

Pace, N. R. 2009. Mapping the tree of life: progress and prospects. Microbiol. Mol. Biol. Rev. 73:565-576.

Park, C. J., Shin, Y. C., Lee, B. J., Kim, K. J., Kim, J. K., and Paek, K. H. 2006. A hot pepper gene encoding WRKY transcription factor is induced during hypersensitive response to Tobacco mosaic virus and *Xanthomonas campestris*. Planta 223:168-179.

Park, C. Y., Lee, J. H., Yoo, J. H., Moon, B. C., Choi, M. S., Kang, Y. H., Lee, S. M., Kim, H. S., Kang, K. Y., Chung, W. S., Lim, C. O., and Cho, M. J. 2005. WRKY group IId transcription factors interact with calmodulin. FEBS Lett. 579:1545-1550.

Ramirez, S. R., and Basu, C. 2009. Comparative analyses of plant transcription factor databases. Curr. Genomics 10:10-17.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin, I. T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W. B., Barker, E., Bennetzen, J. L., Blankenship, R., Cho, S. H., Dutcher, S. K., Estelle, M., Fawcett, J. A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K. A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D. R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P. J., Sanderfoot, A., Schween, G., Shiu, S. H., Stueber, K., Theodoulou, F. L., Tu, H., Van de Peer, Y., Verrier, P. J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A. C., Hasebe, M., Lucas, S., Mishler, B. D., Reski, R., Grigoriev, I. V., Quatrano, R. S., and Boore, J. L. 2008. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science 319:64-69.

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C.-J., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J. Z., Ghandahari, D., Sherman, B. K., and Yu, G.-L. 2000. *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. Science 290:2105-2110.

Robatzek, S., and Somssich, I. E. 2002. Targets of *At*WRKY6 regulation during plant senescence and pathogen defense. Genes Dev. 16:1139-1149.

Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. 2010. WRKY transcription factors. Trends Plant Sci. 15:247-258.

Rushton, P. J., Tovar Torres, J., Parniske, M., Wernert, P., Hahlbrock, K., and Somssich, I. E. 1996. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. EMBO J. 15:5690-5700.

Ulker, B., and Somssich, I. E. 2004. WRKY transcription factors: from DNA binding towards biological function. Curr. Opin. Plant Biol. 7:491-498.

URL1. The plant genome database.

URL2. Joint Genome Institute.

URL3. The Arabidopsis Information Resource.

URL4. The National Center for Biotechnology Information.

URL5. Phytozome: a comparative platform for green plant genomics.

URL6. The GENSCAN Web Server at MIT.

URL7. FGENESH: HMM-based gene structure prediction.

van Verk, M. C., Pappaioannou, D., Neeleman, L., Bol, J. F., and Linthorst, H. J. 2008. A novel WRKY transcription factor is required for induction of *PR-1a* gene expression by salicylic acid and bacterial elicitors. Plant Physiol. 146:1983-1995.

Wu, K. L., Guo, Z. J., Wang, H. H., and Li, J. 2005. The WRKY family of transcription factors in rice and Arabidopsis and their origins. DNA Res. 12:9-26.

Xu, X., Chen, C., Fan, B., and Chen, Z. 2006. Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. Plant Cell 18:1310-1326.

Yamasaki, K., Kigawa, T., Inoue, M., Tateno, M., Yamasaki, T., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Tomo, Y., Hayami, N., Terada, T., Shirouzu, M., Tanaka, A., Seki, M., Shinozaki, K., and Yokoyama, S. 2005. Solution structure of an Arabidopsis WRKY DNA binding domain. Plant Cell 17:944-956.

Yao, D., Zhang, X., Zhao, X., Liu, C., Wang, C., Zhang, Z., Zhang, C., Wei, Q., Wang, Q., Yan, H., Li, F., and Su, Z. 2011. Transcriptome analysis reveals salt-stress-regulated biological processes and key pathways in roots of cotton (*Gossypium hirsutum* L.). Genomics 98:47-55.

Zhang, Y., and Wang, L. 2005. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol. Biol. 5:1.

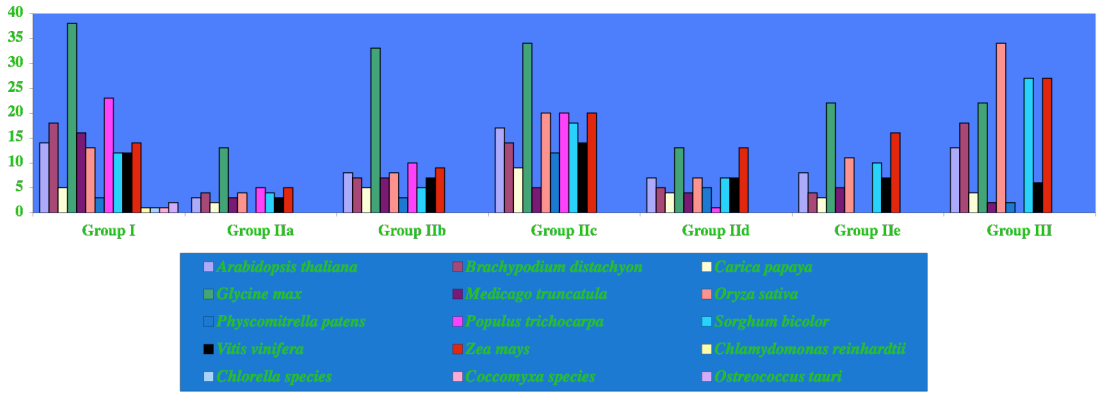


Fig. 2.1 Assignment of WRKY sequences from 15 plant species to groups/subgroups based on the *Arabidopsis* WRKY family classification.

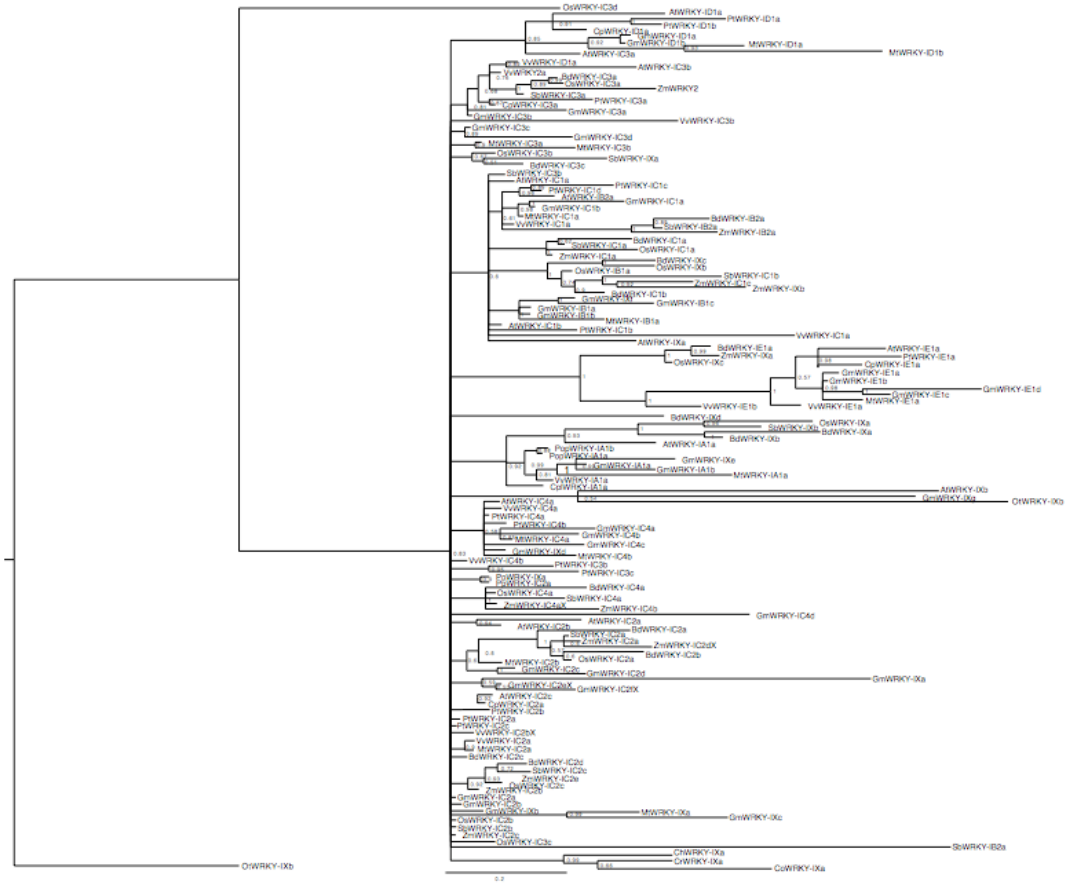


Fig. 2.2 A phylogenetic tree constructed for group I from direct alignment of full length sequences. Phylogenetic tree was constructed with MrBayes 3.2-cvs. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. The resulting phylogenetic tree has poorly supported branches with the phylogenetic position of most of the members not resolved.

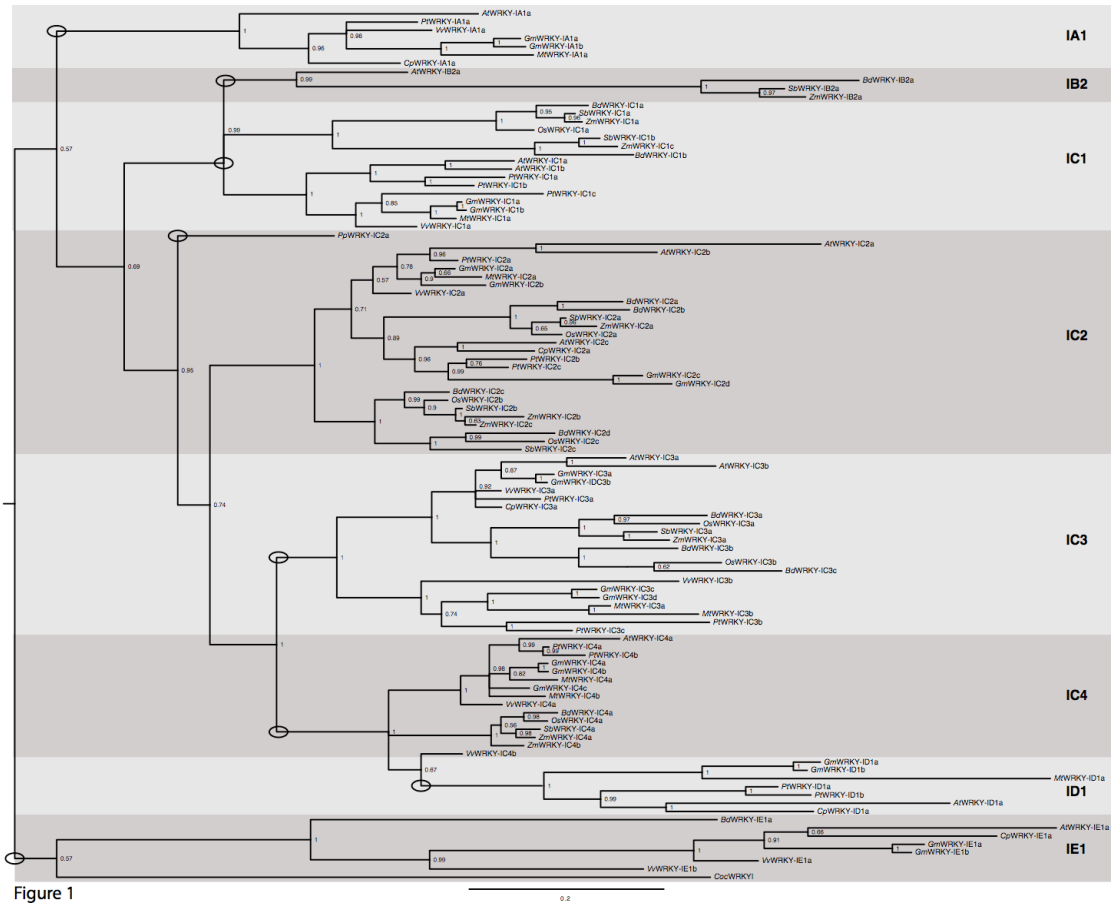


Fig. 2.3 Evolutionary relationships of putative group I WRKY transcription factors from 15 sequenced plant genomes inferred from Bayesian phylogenetic analysis. The tree is unrooted. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. Based on phylogeny and conserved motifs (CMs), five subgroups and four types are identified. These include subgroups IA, IB, IC, ID and IE. The subgroup IC is further divided into four types, IC1, IC2, IC3 and IC4, based on additional unique CMs.

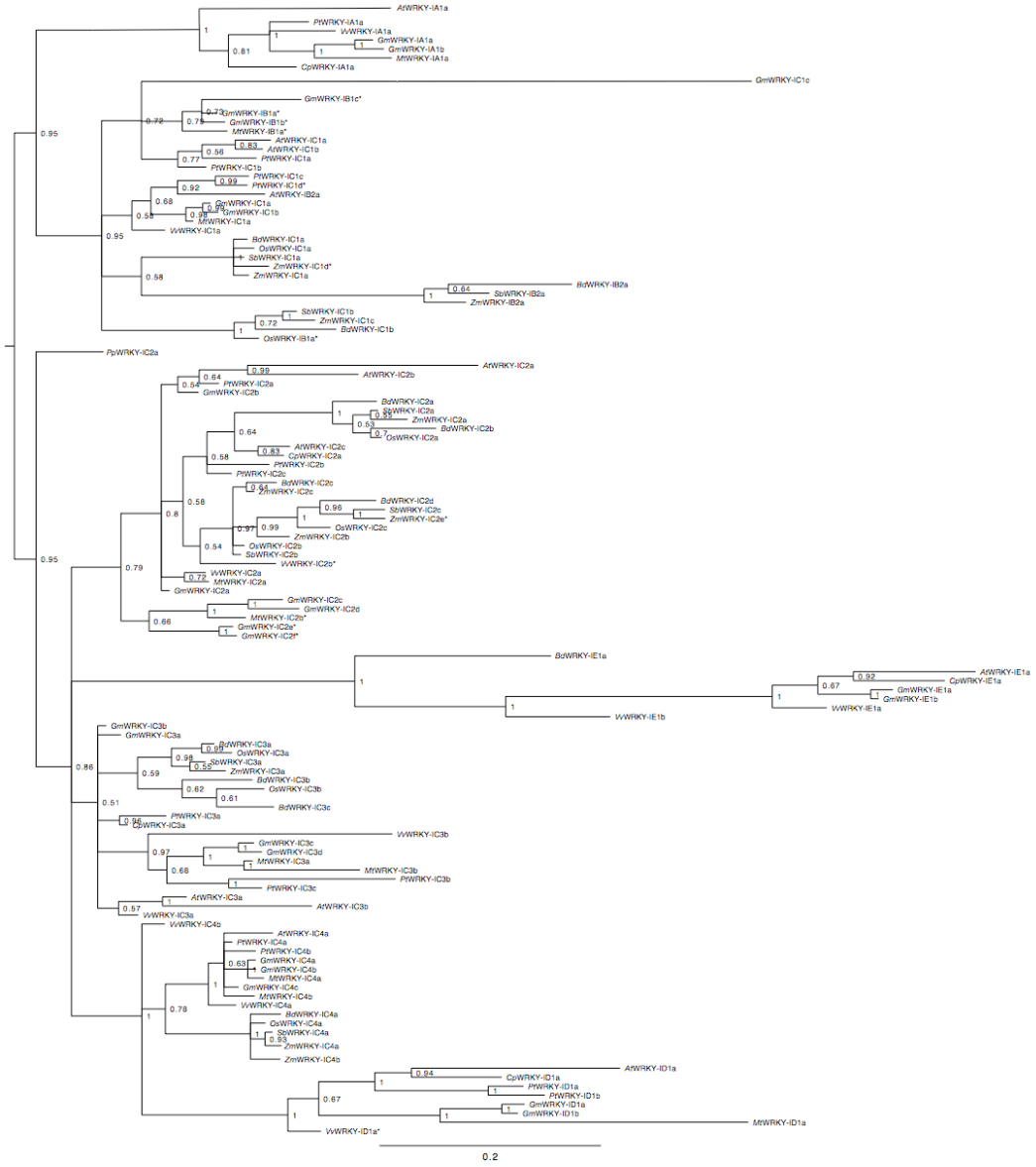


Fig. 2.4 A phylogenetic tree of group I including secondarily assigned members. Phylogenetic tree was constructed with MrBayes 3.2-cvs. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. All the 13 secondarily assigned sequences clustered with their respective assigned subgroups and types.

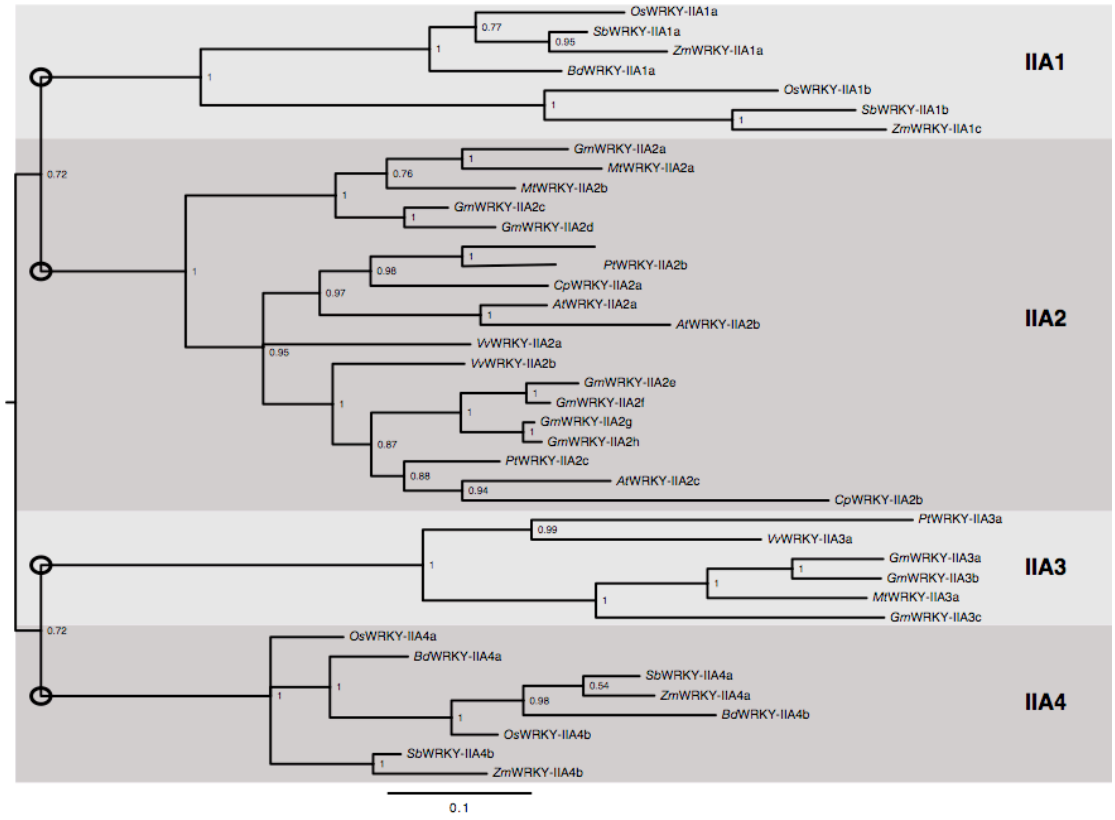


Fig. 2.5 Evolutionary relationships of putative subgroup IIA WRKY transcription factors from 11 sequenced plant genomes inferred from Bayesian phylogenetic analysis. The tree is unrooted. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. Based on the phylogeny and conserved motifs, this subgroup is divided into four types, IIA1, IIA2, IIA3 and IIA4. The IIA1- and IIA4-types consisting of monocot sequences, while those representing types IIA2 and IIA3 are dicot sequences.

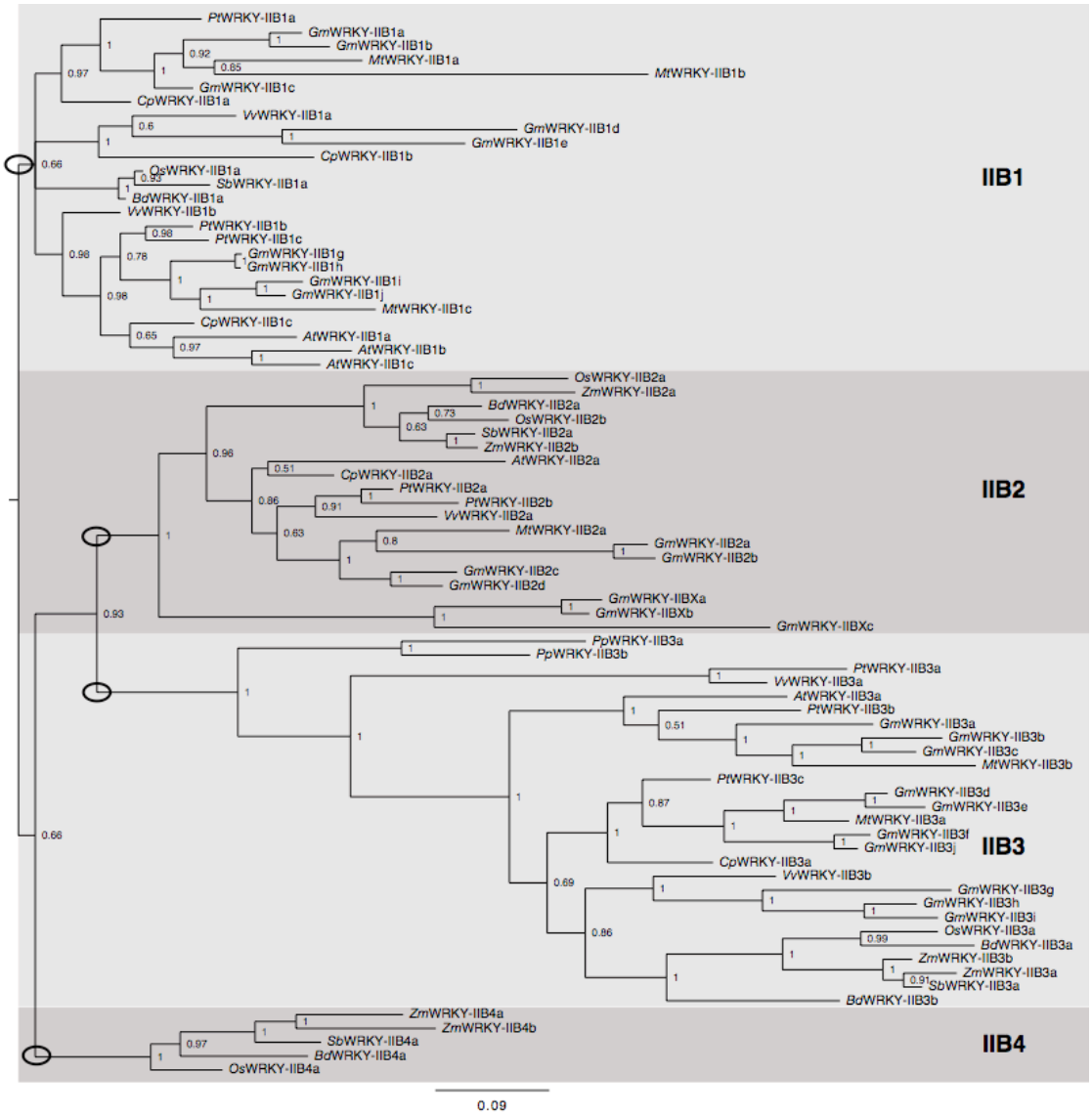


Fig 2.6 Evolutionary relationships of putative subgroup IIB WRKY transcription factors from 11 sequenced plant genomes inferred from Bayesian phylogenetic analysis. The tree is unrooted. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. Based on the phylogeny and conserved motifs, this subgroup is divided into four separate clades representing types IIB1, IIB2, IIB3 and IIB4.

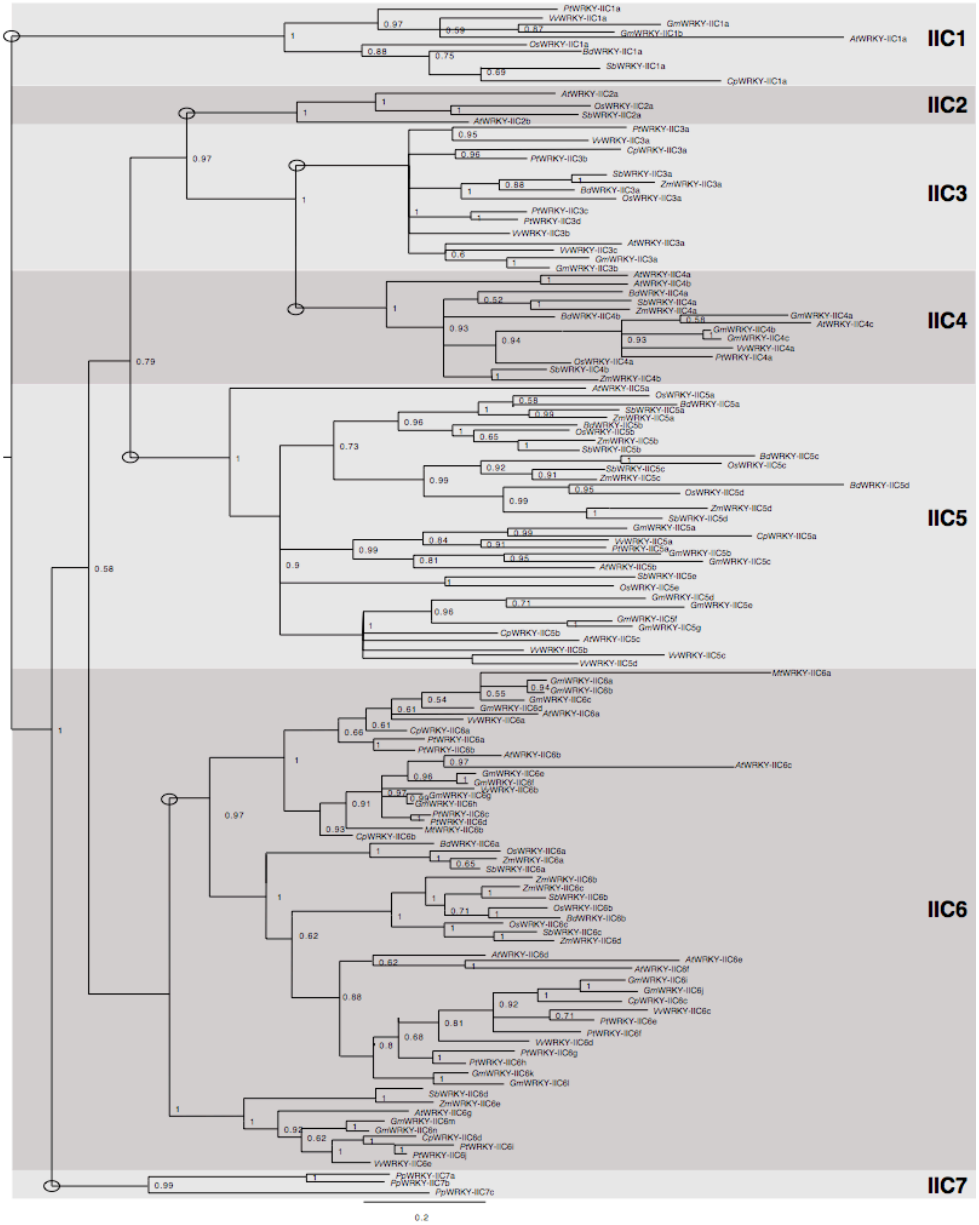


Fig. 2.7 Evolutionary relationships of putative subgroup IIC WRKY transcription factors from 11 sequenced plant genomes inferred from Bayesian phylogenetic analysis. The tree is unrooted. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. Based on the phylogeny and conserved motifs, this subgroup is divided into seven types, IIC1, IIC2, IIC3, IIC4, IIC5, IIC6, and IIC7.

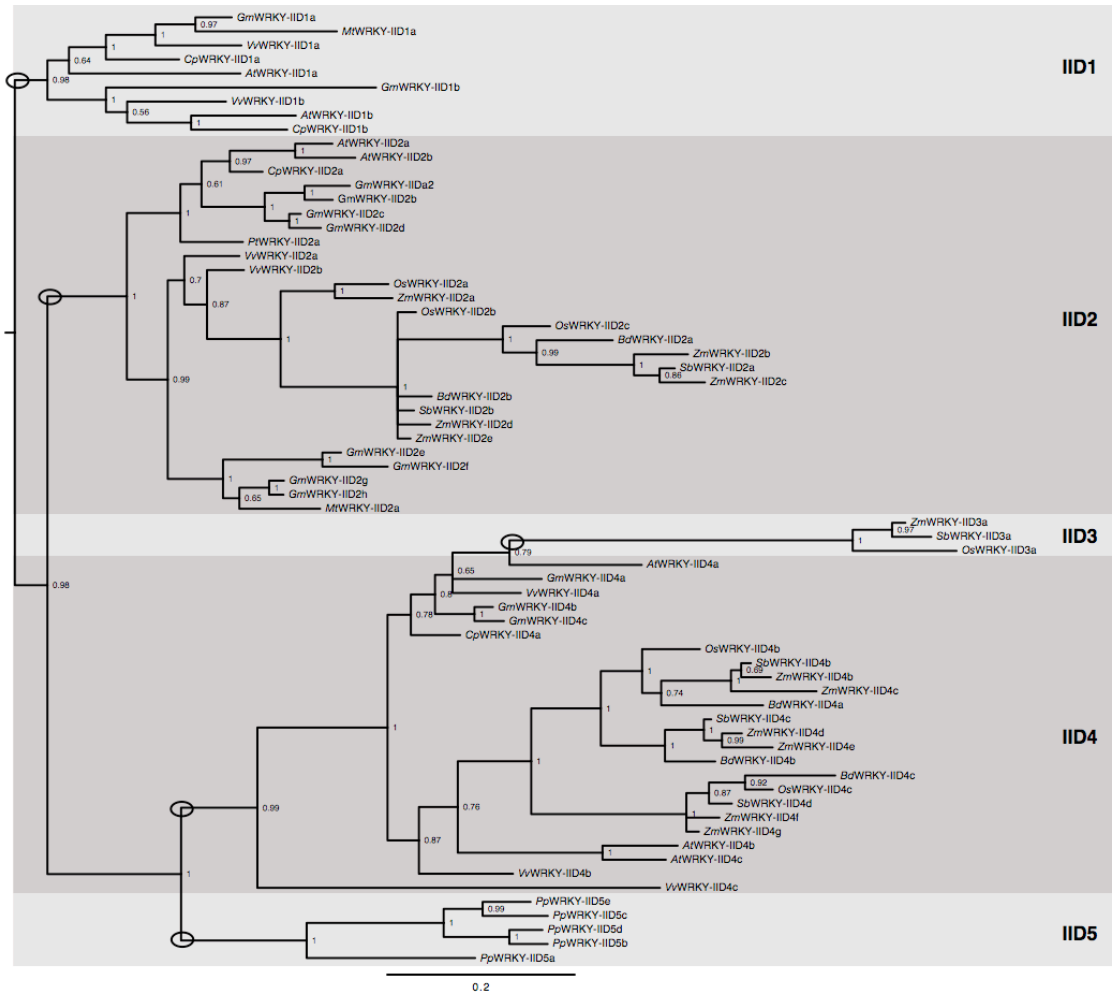


Fig. 2.8 Evolutionary relationships of putative subgroup IID WRKY transcription factors from 11 sequenced plant genomes inferred from phylogenetic Bayesian analysis.The tree is unrooted. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. Based on the phylogeny and conserved motifs, this subgroup is divided into five types, IID1, IID2, IID3, IID4, and IID5

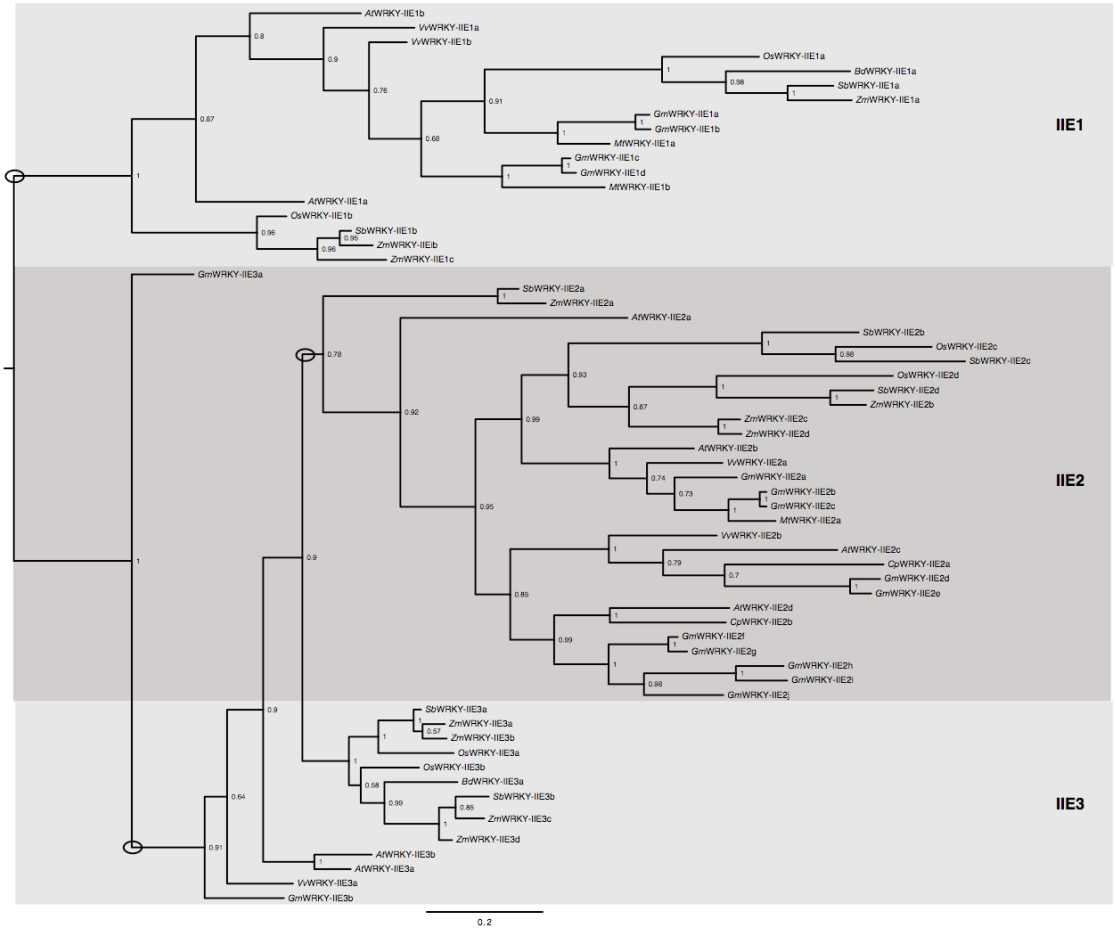


Fig. 2.9 Evolutionary relationships of putative subgroup IIE WRKY transcription factors from 11 sequenced plant genomes inferred from Bayesian phylogenetic analysis. The tree is unrooted. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. Based on the phylogeny and conserved motifs, this subgroup is divided into three types, IIE1, IIE2, and IIE3.

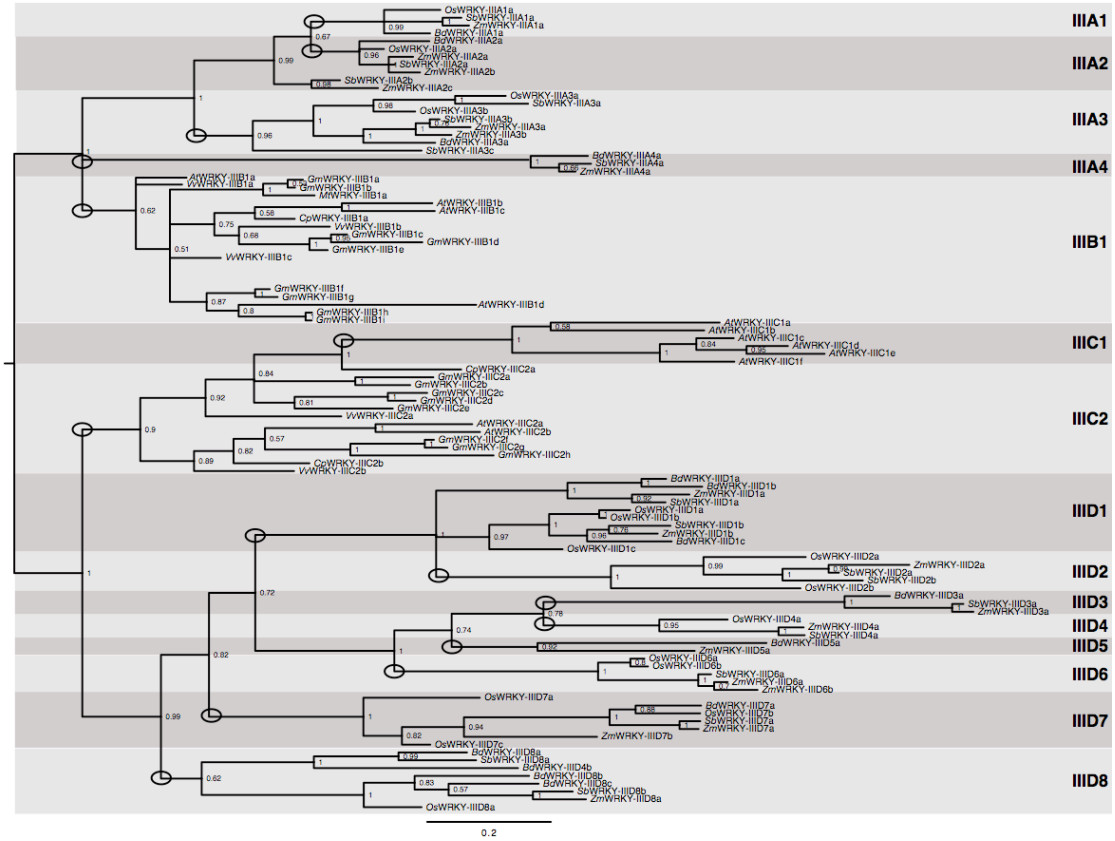


Fig. 2.10 Evolutionary relationships of putative group III WRKY transcription factors from 11 sequenced plant genomes inferred from Bayesian phylogenetic analysis. The tree is unrooted. The posterior probabilities are indicated at each node. The highest support possible is 1.0. Scale bar represents number of substitution per site. Based on the phylogeny and conserved motifs, four subgroups and 14 types are identified. These include subgroups IIIA, IIIB, IIIC, and ID. The subgroup IIIA is further divided into three types, IIIA1, IIIA2, IIIA3 and IIIA4. The subgroup IIIC is further divided into two types, IIIC1 and IIIC2. The subgroup IIID is further divided into eight types, IIID1. IIID2, IIID3. IIID4, IIID5, IIID6, IIID7, and IIID8.

Table 2.1: Conserved motifs defining all WRKY subgroups and types proposed in this study.

Motif name	Consensus sequence[1]	subgroups /types
Motif D	[PS]Y[FL]T[IV]PPGLSP[TAS]x[LF]L[DE]SPV[LF][LF]	IC
charged-1	GD[ED]x[DE][DE]D[ED]P[ED]SKR[RW]KK[ED]xxNE	IA, IB1, IC1, IC3, IC4, ID, & IE
YKPxAK	[IV]YKP[QT]AKLVSK[TA]TVSLLANM[GL]N[CF][NS]T[NS][QS]QQ[PT][QL]Q[PS][AV]E[AT][RN]PQHS[NS][QH]DK[HF][RN]	IA
VASR	VASRP[TS]CS[ST]F[KR]SFSELLAGAIN[ATV][ST]P[APS][IT]P[SC]	IA
Q-rich-1	[QM][GS][GPS]FGMSHQQALAQVTAQA[VA][HLQ][SA][QPN][SL][RHN]M[FH]D[QH]	IB1 & IC1
P-rich-1	[PT]PPPPRP[TR][LI][ATS]LPPR[SP][ASF]AE[SA][FL]F[TS][GAS][AG][GA][DGS]ASP	IB1 & IC1
FxxLLxG	[VF][AG]S[RS][PS][ST][AC][AS]D[FC][KR]SF[ST][QE]LLAGA[IML][AN]SP[PA][AP]x[PS]S	IB1 & IC1
VEEV	[GA]VEEVT[AS]A[PT]AINL[VT]PQ[SF]	IB2
SPTTG	[NS]S[NK][IAV][LQE][PA]SPTTG[ASK][FL]Px	IC2, IC3, & IC4
RTGSG	RGG[GS]GV[PS]K[FY]K[SA][AM][QTP]PPSLP[LI]S	IC2
ENSS	[DT][SP][VA]ATPE[ND]SS[AIV][ST]FG[DE]D[ED][FA][ED]	IC2
charged-2	G[DG][ED][DF]D[EDG]DEP[DE][AS]KRW[KR][GK][DE][GN][EDG]NE	IC2
DxSS	F[ED]SP[DE][AG][VP][DE][AV][ST]S[TA][FLV]S[ND][EDH][EDV]D[DG][DE]D	IC3
LxER	QK[SP]S[SP][RG][GS][GS]L[AVS]ER[MI][AQ]ARA	IC3

	GFN[AV]P[KR][LI][ND][TM][EPS][SFR][IV]	
VxxxE	[LS][SA][LP][SN][SQ]PV[QHS][MI]V[CST][SP][SG][AD][SN][AM]P[AV][ED]V[DG][TLS][DS]E[LM][NH]Q[IM][GN][SN]S[EA][NT][GA]LQ[AE]SQ[SV][ED]	IC4
ARYK	[AGS][GL][AGQ][GN][ASG]GARYK[LA]MSPA[KR]LPISR	IC4
PSIIR	[SP]I[IM]REKVS	ID
SVxxS	VCST[SP]LS[ED]LSPTSV[TS][QH]S	IE
CC-A	L[EV][EA][EK]LRRVSEEN[KR][RK]LTEML[AT][VA][VML]C[EA][NK]YNAL[QR][KS][HQ]Lx[ED][LM][VM]x[KA][NA]	IIA (majority)
alanin-rich-1	SPEI[RQ][QK][NV]L[VA][EQ][QE]MA[ST][ST]LT[RK]DPNF[TK]AAL[AV][AT]A[IVL]SGR[IL]L	IIA (majority)
SINS	S[ST]LP[CR]SIISI[DN]S	IIA1
LDL-1	SGPT[VI]TLDLT[KQ]S[GK]GS	IIA1, IIA2 & IIA3
ESSSTDE	NG[IG][NSV][NG][GN]N[ST]ESSS[TS]DE[DE]S[SC]KKP	IIA2
ECTS	[AF][AV]A[PAH]DQ[TAM]ECTS[AG][AE][PA][CA]	IIA4
EGGS	N[HN][PQ]SSTSEGGS	IIA4
CC-B	ELAA[LA][QK]AE[LM][GE][RE][VM][KNR]EENQ[RK]L[KR]xML[SD][QR]	IIB (majority)
alanin-rich-2	DTV[ED][TA][AV]T[AK]AI[TA][AS]DP[NS]FT[AS]ALAAAI[TS]SIIG	IIB (majority)
LDL-2	[MN]A[TS][IL]S[AST]SA[PS][FH]PT[IV]TLDLT[SQ]	IIB (majority)
QSKF	LYNQSKFSGLQ	IIB1
RxxLPC	MN[SP]N[FL]L[AT]R[AT][IVL]LPCSS	IIB1

VPRQF	Kxx[KG]xGGx[ML]VPRQF[ML]DLG[PL][AS][AS]	IIB1
PYAS	S[GY][FV][SFH]SS[FIL]PYAS	IIB2
acidic-2	H[DE]E[ED]xEE[SD][ED]LVSL[SC]LG[RT]S[PS][SN]	IIB3
FLAR	LMAGSNFLARAV	IIB4
MMPLP	MMPLP[AH]F[ED][HL][GH][HN][HQ][QH][HP][LPQ][AIQ]H	IIB4
YSGPTI	[FILV]N[KR][LF][IT]S[TN][VI]YSGPTI[SQ]DIE[NS]ALS[FLV][TS][AGN][AQ][RG]D[HAQ]	IIC1
KYTxK	[SK]K[IM]ENKYTLKI	IIC1
LLxD	xQxxLLRD[YH]GLLQD[IM]VP[SP]xMR	IIC1 & IIC6
RMVI	RMVI	IIC2
FEHIL	DNFEHIL[ST]QMQIY	IIC3
LxPxL	CEKLME[AT]L[ST]P[IL]L[KR]Q[LI]QFLS[QRS]	IIC4
DWxxL	[GIL][PL]AD[IV]DW[AS]SL[LF][SLQ][AGP]QS	IIC4
TPNSS	P[AS]TPNSS[SV][SI]SSSS[SE][ED]	IIC6
basic-1		IID (all)
acidic-3	[AE][VI][EQ]EA[AN][RSA]A[GA][LV][EKR]S[CM][EH][RH][LV][LI][RAS][LS]LSQ[QP]QDQxx	IID (majority)
motif C	[IE][TA][DG]E[AT]V[SA][KR]F[KR]KV[IV]SLL[NS]R[TG]G[HG]	IID (majority)
HARF	HARFRR[AG]PVVS[PS][PS][PS]	IID1, IID2, IID3, & IID5
PxxxxP	[GT]S[AE][FTS][RK]VY[CH][PA]TP[LI][HQ][QR][ILV]PPL[PS]H[NH][NHQ][HIQ][HP][NQ]	IID1
VTLDF	TLDFTKP	IID2
GSVSxG	G[DE]GSVS[NKDG	IID2
MSDAA	[EK][VE][IV][SI][FS][SF][FS][DF][ND]NSVCTSSAATSFFTSI[SG]SQLISMSDAAT[SN]	IID3
SxxGS	[MA]SS[TAS]RSF[LI]SSLS[MI]DGS	IID4

HARV	H[AP]R[VG]R[KMFL][KRSI]	IID4
FHLxG	[DGK][GKS][SN][SAP]FHLIG[AG]P[AHV][AMS]SD[PQ][NAV][NS][AQ][QH][QH]	IID4
AGVFLE	[YC]AGVFLE[SN]SNF[FC][TR][DE]N[AST]Q	IID5
KCAI	[GS][AG][TK]CA[IT][LA]G	IID5
basic-2	[PS][GP][KS][KR][RS][RK][KR][SN]Q[QA]K[KR]VVC[VIH][PV][AT][AP][ADE][ANV]	IIE (all)
motif E	D[WG]DLT[AD][VI]VR[SGA][CG]	IIE2 & IIE3
acidic-4	[EG][EH][DE][DE][DSA][LF]FA[DG]L[GA]EL[PE][ES][CD][PSA][MV][VS][FL][RI]	IIE1
GExxxP	[SP][KR][GS][KV]GEG[NY][PT]P[ST]	IIE1
WFxD	F[GR]W[FL][SYFGD]	IIE1
LESP	[VIM]LESPI[CFM][AG][EGT][VG][DY]	IIE1
RxxxTG	AA[GN][GS]R[PT][TS]G	IIE3
DPFxxxxDPF	[FG][GA]D[PA]F[SA]G[LM][VPR]DP[FL][ALS][SHT][DE]	IIE3
acidic-5	[GH][HQ][PSA][DE]DFF[AS]DLAELESDP[ML]SL[IL]	IIE3
hydrophobic-1	SS[ED]LAExL[VA]xK[IV]L[RS][SC]FE[KR][AS][LI]	III
DILGAK	[DE]ILGAK	IIIA & IIIB
VRVQ	V[SA][ED]L[GCV]RVQ	IIIA1
LxAP	L[QRS][AS]P	IIIA1
TARWT	RK[AT]TARWTS	IIIA1
PHAQ	PHAQ[AST][AL]LQ[GS]L[AS]A[RGS]	IIIA1
LVVT	LVVTEL[SG]HIK	IIIA2
ENWG	PST[PS][EN]NWGVSPA[TS]SDSNH[VA][AV]	IIIA2
RVSAVQD	[KR]RK[TGA][LMT]P[KC][WV][SR][RT]Q[VL]R[VA][SA][SA]VQD	IIIA3
RGCY	RG[CS]YKR[RK][KRS][NTS][AES][QP]TW[TE][IK]	IIIC2

	[EV][SA][SQ]	
RRKA	[DK][KR][KR][KR][VA]IEEL[VL][KR]G[RH]	IIIC2
QLRA	EA[ML][ER]E[IMV][ARG]R[EGQ]Q[ES]LV[TA]QLRA[LI][VL][LF]	IIID1
LETP	[KQH][RKH][RK]R[RK][RLN][DF][KDG][RDE]S[RV]SL[EV]T[PHN]VP	IIID1
KILES	M[NK]ILES[SF][TGN][HLR][SG][GD][CY][QK][VE]	IIID2
HxGS	[NS]KRRKNA[QEN]H[TI][GS]S[VIT][VM][TA][QA][AT]P	IIID2
RQQVEL	[RK]G[TA]QLAE[LF]LRQQVELIPE[PH]	IIID3
PIKEIL	G[ADQ]ELPIKEILTE	IIID3
EHEAV	[DE]EHEAV[IV]RELTRGHELTA[QR]	IIID4
YELIK	AVREVAQVYELIK[LT][QH]QPLLL	IIID6
SWAT	SSW[ASV][TQY][LFHV]T[APV]V	IIID6
GxxSAA	[VA][VA][LTS]ELM[TA][MK]G[RQ][EQ]	IIID7

[1]: "[ ]" indicate alternative amino acids at the same position

CHAPTER THREE

Sequencing and comparative analysis of the potato aphid *Macrosiphum euphorbiae* transcriptome.

Abstract

The potato aphid, *Macrosiphum euphorbiae*, is an important agricultural pest that causes economic losses to potato and tomato production. Resistance of tomato to this aphid is well characterized making it a model species for the study of plant-aphid interaction. However, no genomic resources exist for *M. euphorbiae*. To establish transcriptome data for this aphid, we used Illumina sequencing and generated 52.6 million 75-105 bp paired-end reads that were assembled into 24,137 contigs. About 70% of the assembled contigs were of annotatable length (>300 bp), and represented more than 7,000 transcripts previously predicted for the related pea aphid, *Acyrthosiphon pisum,* genome. We found 55% of the contigs to represent the Gene Ontology (GO) molecular function categories protein binding and catalytic activity, while 36% of them were classified in the cellular component ontology as nuclear localized. Through comparative analysis of transcriptomes from 11 insect species, including four aphids, *A. pisum, Myzus persicae*, *M. euphorbiae*, *Aphis gossypii* and one planktonic crustacean, *Daphnia pulex*, we identified a set of conserved genes including those involved in insect immunity that are missing in aphids as well as a set of sequences that are specific to aphids. Moreover, a 15,198 bp *M. euphorbiae* transcript predominantly expressed in the gut tissue was identified.

Introduction

Aphids (Hemiptera: Aphididae) are among the most destructive agricultural insect pests worldwide (Dixon et al. 1998). They have a short generation time often resulting in vast population expansion during a single growing season. Aphids damage their host plants by both directly feeding on them and indirectly by transmitting viruses or supporting the growth of saprophytic fungi. Being phloem feeders, they are nourished by sucking the phloem sap, which is needed for plant growth and reproduction. Moreover, during the feeding process, they inject saliva that can be phytotoxic (Evert et al. 1968) or contain effectors that suppress plant defenses (Bos et al. 2010). Indirect damage is caused by plant viruses transmitted by aphids or by black sooty mold growth on plant tissues covered with honeydew which is excreted by the aphid. The latter compromises the plant's photosynthetic activity and lowers the marketability of its fruits. The economic losses as a result of virus transmission often far exceed aphid's direct impact on crops (Katis et al. 2007; Nault 1997).

Aphids have a somewhat complex life cycle, comprising of both sexual and asexual (parthenogenetic) modes of reproduction, wing dimorphism (Blackman and Eastop 2000; Braendle et al. 2006) as well as a high diversity in terms of host range and host plant specialization. Moreover, they possess a diverse symbiont community that includes the obligate bacterial symbiont *Buchnera aphidicola* (Buchner 1965) as well as several facultative symbionts, a subset of which are believed to contribute to the aphid host range (Leonardo and Muiru 2003). In addition, aphids have extremely complex relationship with their plant hosts. Under low infestation levels, their highly

specialized mode of phloem feeding causes little apparent plant damage enabling them to evade the plant immune system. The unusual biology of aphids makes them ideal models for the study of several biological processes that are not readily studied in other genetic model systems. Some of these aphid-associated characteristics are expected to be the result of unique sets of genes found in this genus. Analysis of the pea aphid (*Acyrthosiphum pisum*) genome, the only publicly available aphid genome, identified orthologous gene relationships with other arthropods (Huerta-Cepas et al. 2010) and revealed a wave of gene duplications in the aphid lineage that is larger than that of any other sequenced insect clade (consortium 2010). This genome analysis also identified loss of evolutionarily conserved genes central to the IMD immune pathway, selenoprotein utilization, purine salvage, and the entire urea cycle (consortium 2010).

Like *A. pisum*, the potato aphid *Macrosiphum euphorbiae* belongs to the tribe Macrosiphini (Von Dohlen and Teulon 2003). *M. euphorbiae* infests many plant species including those from the solanaceae such as potato and tomato and transmits a number of plant viruses (Chan et al. 1991; Moeller 1973; Radcliffe and Ragsdale 2002). In tomato, resistance to this aphid is mediated by the *Mi-1* gene that encodes a nucleotide-binding leucine-rich repeat protein (Kaloshian et al. 1995; Rossi et al. 1998). Both *Mi-1*-virulent and avirulent *M. euphorbiae* are available (Hebert et al. 2007; Kaloshian et al. 1997). In spite of this interesting host-aphid interaction, only a few transcript sequences for this aphid are currently available in public databases.

In this study, we assembled a *M. euphorbiae* reference transcriptome by generating 128 giga bases (Gb) of high-quality sequence information using Illumina

technology. Sequences generated by paired-end 75 or 105 bp reads were *de novo* assembled into 24,137 contigs and annotated. Based on Gene Onthology (GO) analysis the contigs were assigned to diverse molecular function and biological process categories suggesting a comprehensive representation of the *M. euphorbiae* transcriptome. Through comparative analysis we identified several evolutionary highly conserved gene sets that have either been lost or diverged in aphids. The inference of these observations regarding interactions of aphids with their primary or secondary symbionts is discussed. Interestingly, we identified potentially unique sets of genes operating in the aphid lineage. Functional characterization of these sequences will add to our understanding of unconventional aspects of aphid biology as some of these sequences may be involved in the phenotypic plasticity of aphids and their interaction with hosts. Furthermore, we identified a *M. euphorbiae* transcript over 15,198 bp in size predominantly expressed in the aphid gut.

Materials and Methods

Plant and aphid colonies

Tomato cultivar (cv.) UC82B (*mi-1*/*mi-1*) and near isogenic cv. Motelle (*Mi-1*/*Mi-1*) and cv. Moneymaker (*mi-1*/*mi-1*) as well as mustard india (Burpee & Co., Warminster, PA) were grown in UC mix II (agops.ucr.edu/pdfs/soil_mix_recipes.pdf) in a growth room at 24ºC with 16h light/8h dark. Colonies of parthenogenetic potato aphid (*M. euphorbiae*) and green peach aphid (*M. persicae*) were reared on the susceptible tomato cv. UC82B and mustard, respectively. The colonies were maintained inside insect cages in a pesticide-free greenhouse at 22-26ºC. The pea aphids (*A. pisum*) maintained on fava bean (*Vicia faba*) variety Windsor were kindly provided by Dr. Greg Walker (Department of Entomology, UCR).

Aphid material

Age-synchronized, one-day-old, adult aphids were generated as described previously (Bhattarai et al. 2007). About 200 one-day-old adult aphids were exposed to resistant Motelle for 12 h and 24 h or to susceptible Moneymaker tomato plants for 24 h. Additional 200 one-day old adult aphids were subjected to starvation for 24 h in a Petri dish. Mixed stage aphids were collected from the colony reared on the susceptible tomato cv. UC82B.

RNA and DNA extraction

For Illumina library preparation, RNA was extracted using the RNeasy Midi kit according to manufacturer's recommendation (Qiagen, Valencia, CA). Twenty μg of RNA was treated with DNase I enzyme (New England BioLabs, Ipswich, MA) followed by phenol-chlorophorm extraction and isopropanol precipitation. The RNA quality and integrity were evaluated by Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA). For RT-PCR analysis, total RNA was extracted from 40 dissected aphid heads or guts using Trizol reagent (Invitrogen, Grand Island, NY). Tissues were homogenized in Trizol followed by chlorophorm extraction and total RNA was precipitated using isopropanol. DNase treatment was done as mentioned above. For RNA blot analysis RNA was extracted from mixed developmental stages of *M. euphorbiae, A. pisum* and *M. persicae* using Trizol reagent (Invitrogen). DNA was isolated using a standard phenol/chloroform extraction protocol.

Library construction and sequencing

mRNA-seq libraries were prepared for high-throughput sequencing on the Illumina Cluster Station and Genome analyzer as described by (Atamian and Kaloshian 2012). In brief, mRNA was isolated from 4 μg of the DNase-treated total RNA using Sera-mag Magnetic oligo(dT) beads and fragmented with divalent cations under elevated temperatures. The cleaved mRNA fragments were copied into first- and second-strand cDNA using random primers. The overhangs were converted into blunt ends using T4 DNA polymerase and Klenow DNA polymerase, followed by the addition of an "A"

base to the 3' end of the blunt phosphorylated cDNA fragments. Adapters were ligated to the ends of the cDNA fragments, purified on gel and 300 bp templates selected for downstream enrichment with PCR using primers complementary to the adapter sequences. The size, purity and concentration of the prepared library were evaluated by running 1μl on 2% agaorse gel. To assess the diversity of the library, 1 μl of the library was cloned into Zero Blunt TOPO vector (Invitrogen) and 10 clones were sequenced using Sanger technology.

Paired-end 75 or 105 nucleotide long sequencing was performed with the Illumina Cluster Station and Genome Analyzer II at the Institute for Integrative Genome Biology, University of California Riverside.

***De novo* assembly of reads and annotation**

Data from Illumina Genome Analyzer II sequencing runs were processed using the Illumina pipeline version 1.4 to generate sequencing reads, base-call quality scores, and remove low quality reads. The sequence data generated in this study will be deposited in NCBI's Sequence Read Archive (SRA), and accession number will be provided. The reads were preprocessed with SEED and assembled with Velvet/Oases as described in Bao et al. 2011 (Bao et al. 2011) and the assembly is available at http://biocluster.ucr.edu/~hatamian/Macrosiphum_euphorbiae_Whole_body_transcriptome.txt. The resulting contigs were annotated by BLASTX searches against the UniProt database and the predicted transcripts from pea aphid provided by AphidBase.

Cluster analysis and annotation

The predicted transcript sequences of *A. mellifera* (Amel_pre_release2_OGS_cds), *N. vitripennis* (nasonia_automated_gene_model_v1), *A. gambiae* (agambiae.TRANSCRIPTS-AgamP3.6), *D. melanogaster* (dmel-all-CDS-r5.41), *D. pulex* (daphnia_genes2010_beta3), *T. castaneum* (Tribolium_mRNA3.0), *B. mori (*silkcds), *A. pisum* (ACYPI mRNA v2.1) and available EST sequences of *B. tabaci*, *M. persicae*, and *A. gossypii* were downloaded from genome databases, NCBI and AphidBase 2.1 EST collection. The EST sequences of *B. tabaci, M. persicae*, and *A. gossypii* were assembled into contigs using the CAP3 assembly program (Huang and Madan 1999). Similarity matrices were computed with TBLASTX and subjected to the MCL software (van Dongen 2000). The main parameter of MCL, called inflation or granularity value, was set to 2. For group I clusters, TBLASTX searches were performed against the aphid EST sequences downloaded from NCBI (taxids: 7029, 13164 and 80765) and the *A. pisum* genomic sequences from AphidBase 2.1 to confirm their assignments by the MCL software. The clusters containing sequences with matches (E value < $1e^{-5}$) were discarded. Moreover, manual conserved domain analysis (Marchler-Bauer et al. 2011) was performed for the sequences within individual cluster and clusters containing sequences with multiple non-overlapping or no conserved domains were discarded. Group II individual cluster sequences were analyzed using TBLASTX against NCBI's nucleotide and EST databases. Those clusters with sequences having matches (E value < $1e^{-5}$) against sequences other than aphids were discarded. Group III cluster sequences were searched against the NCBI

nucleotide and EST databases and *A. pisum* genomic sequences from AphidBase 2.1. Clusters with sequences having matches (E value < $1e^{-5}$) were discarded.

Tissue-specific expression analysis

Semi-quantitative RT-PCR (sqPCR) analysis was performed using cDNA prepared from whole aphids or dissected aphid heads and guts using primer pairs *Me*_WB10316F1 (5′-ATGCTAGGCAATTCTCCTCT-3′) and *Me*_WB10316R1 (5′-CTTGTTCTCAGTGGTTTGAG-3′), *Me*_WB26189F1 (5′-CTCACTGGGGCA CCTAATGT-3′) and *Me*_WB26189R1 (5′-GCAACAAGGGAACCAAGAG-3′), *Me*_WB28140F1 (5'-CAACATGCACATGTACACACCTC-3') and *Me*_WB28140R1 (5'-TGAAGAAGCCGGGAAGGTAG-3'). The ribosomal gene L27 (NM_001126221) was used as internal control and amplified using primer pair L27F (5′-CCGAAAAGCTGT CATAATGAAGAC-3′) and L27R (5′-GGTGAAACCTTGTCTACTGTTACATCTT-3′). As control for the tissue specificity, the expression of two marker genes, *C002* (Mutti et al. 2008) and *Sucrase (APS1)* (Price et al. 2007) was assessed using *M. euphorbiae* gene-specific primer pairs *Me*C002F2 (5′-GAGCAGGAAGAAGCGTCTGT-3′), *Me*C002R2 (5′-CTTG GTGGGAGC ATTGGTTA-3′) and *Me*SucF (5′- GAGATCGATCCTATTTATGGC-3′), *Me*192R2 (5′-CATTCCATTCCCACGGAGATC-3′), respectively. PCR was performed in 25 µl with 40 ng of cDNA template, 1X PCR buffer, 2.5 mM $MgCl_2$, 0.4 mM dNTPs, 2 units of *Taq* DNA polymerase, and 10 µM of each forward and

reverse primers. The PCR program was initialized at 94°C for 5 min, followed by 24 cycles of 94°C for 1 min, 60°C for 45 s, and 72°C for 1 min, with a final extension at 72°C for 10 min.

RNA and DNA blot analysis

Total RNA (8 μg) from each aphid species was fractionated on 1.2% denaturing formaldehyde agarose gels and blotted to nylon membranes according to standard protocols. For DNA blot, 7 μg of genomic DNA was digested with *Hind*III, *EcoR*V or *Bgl*II restriction enzymes, fractionated on a 0.8% agarose gel and blotted to a nylon membrane according to standard protocols. The blots were hybridized with 32P labeled probes in 50% (v/v) formamide at 42°C overnight. After washes, blots were exposed to X-Ray film (Research Products International Corp.) with an intensifying screen at -80°C.

Radioactive ^{32}P dCTP labeled probes were prepared using random priming (Pharmacia) and purified on a sephadex G-50 column. For probes, 450 bp, 472 bp and 561 bp subclones were amplified from Me_WB10316, Me_WB26189 and Me_WB28140 using the primer pairs Me_WB10316F2 (5′-CACACTTTTGCTATCAACGAT-3′) and Me_WB10316R2 (5′-GACTTTAGCGTTC CAGTCAT-3’), Me_WB26189F2 (5’-CTCACTGGGGCACCTAATGT-3’) and Me_WB26189R2 (5’-AGCGCGTGAGCTCAGGCA-3’), Me_WB28140F2 (5’-GAGTCGGACACGTGGTTACCT-3’) and Me_WB28140R2 (5’-CATCATCTTGATCCACATAAGC-3’).

Amplification of the gaps between contigs and sequencing

The gap between Me_WB28140 and Me_WB26189 was amplified using the primer pair MeCF1 (5'-GGAGGAGAGGTGTCAAGACTG-3') and MeCR1 (5'-ATGTTGGTGTGTGAGCCCTTC-3'). The gap between Me_WB10316 and Me_WB26189 was amplified using the primer pair MeCF2 (5'-GCTATAGGTCGT TTCGTGTTTAATG-3') and MeCR2 (5'-CTAAGGTTATACTTGTGCCAAAGAC-3'). The amplified products were cloned using TOPO TA cloning kit (Invitrogen) and subjected to Sanger sequencing.

Results

Transcriptome sequencing and assembly

To maximize the genome coverage of the *M. euphorbiae* expressed genes in our experimental material, we prepared RNA-Seq libraries representing transcripts from mixed aphid developmental stages as well as aphids exposed to various biotic or abiotic stresses. For the biotic stress conditions, three libraries were prepared from 200 age-synchronized, one day-old, adult aphids either exposed to *Mi-1* resistant cv Motelle plants for 12 or 24 h, or to the susceptible cv. Moneymaker for 24 h. For abiotic stress, a single library was prepared from 200 age-synchronized, one day-old, adult aphids starved for 24 h. An aliquot of each library was cloned and 10 clones from each library were sequenced. BLAST searches of the sequences against NCBI's nucleotide database identified different sequences for each clone within each library suggesting that they were not biased by containing only a limited number of distinct transcripts. The RNA-Seq library from mixed aphid developmental stages was run on two flowcell lanes, while the remaining four libraries were combined together and run on a single flowcell lane.

A total of 52.6 million paired-end reads were generated. Two main approaches were employed to generate a draft assembly of the *M. euphorbiae* transcriptome. First, *de novo* assemblies were optimized using four different assembly tools, including Velvet/Oases (Schulz et al. 2012), Trinity (Grabherr et al. 2011), ABySS (Robertson et al. 2010), and SOAP_denovo (Li et al. 2010). Since Velvet/Oases gave consistently

the best results, we used the contigs from this assembly for all downstream analyses. To further improve the assemblies, we developed a novel clustering algorithm for next generation sequencing data that greatly facilitated and improved the quality of the downstream assemblies (Bao et al. 2011). This method, called SEED, was able to reduce the time and memory requirements for assembling our data with Velvet/Oases by 60-85% and 21-41%, respectively. In addition, the assemblies contained longer contigs than non-preprocessed data as indicated by 12-27% larger N50 values. The final assembly contained 24,137 contigs with an N50 value of 2,130bp. Over 70% of the assembled contigs were greater than 300 bp in length (Figure 1A). The identified transcripts were aligned to the predicted *A. pisum* transcriptome (ACYPI mRNA v2.1) in AphidBase 2.1. More than 7000 of the *A. pisum* transcripts had at least 50% coverage by their corresponding *M. euphorbiae* contigs (Figure 1B).

Gene onthology assignments and annotation

The consensus contigs obtained from the optimized *de novo* assembly were annotated by BLASTX searches against the UniProt database and TBLASTN searches against the predicted transcripts from *A. pisum* (ACYPI mRNA v2.1) in AphidBase 2.1. In both cases we used an E-value cutoff of $<1e^{-5}$. Gene Onthology (GO) terms were assigned to the *M. euphorbiae* contigs based on the best ranking BLAST hits against the UniProt and AphidBase databases (Figure 2). Within the molecular function ontology, a high percentage of genes were assigned to the catalytic activity (35%), protein binding (20%) and nucleotide binding (15%) categories. The cellular

component terms showed a high representation of genes with nuclear (36%) localization, while the biological process terms were associated predominantly with multicellular organismal development (19%), nucleobase-containing compound metabolic process (18%), cellular component organization (16%), and transport (14%).

Comparative analysis

To identify sequences that are either absent or conserved among various aphid species or specific for *M. euphorbiae*, we performed sequence similarity clustering combining the Markov clustering (MCL) algorithm with subsequent BLAST searches and stringent criteria. We applied this approach to 11 insect species (*Apis mellifera, Nasonia vitripennis, Anopheles gambiae, Drosophila melanogaster, Tribolium castaneum, Bombyx mori, Bemisia tabaci, A. pisum, Myzus persicae*, *Aphis gossypii* and *M. euphorbiae*) and one planktonic crustacean (*Daphnia pulex*) used as an outgroup. The resulting clusters were divided into three groups. Group I clusters included sequences restricted to all seven non-aphid insect species and *D. pulex*, group II clusters included sequences restricted to the all four aphid species, while group III clusters contained sequences restricted to only *M. euphorbiae* (Figure 3).

Our analysis identified 17 group I clusters. Additional TBLASTX searches with all sequences in group I, against the *A. pisum*, *M. persicae* and *A. gossypii* sequences in the NCBI nucleotide and EST databases, *M. euphorbiae* contigs generated in this study and the *A. pisum* genomic sequences on AphidBase 2.1,

identified no similarity ($E < 1e^{-5}$) confirming their absence in the four aphid species (Table S8). The group I clusters were annotated with GO terms from the biological process and molecular function ontologies (Table 3.1) using the most recent Drosophila annotations (FlyMine v32.0). Among conserved genes not identified in aphids, are those with roles in immunity (cylindromatosis), peroxisome division (similar to peroxisomal biogenesis factor 11), phagocytosis (Haemolysin-III related), carbohydrate metabolism (alpha-N-acetylglucosaminidase) and amino acid metabolism (homogentisate 1,2-dioxygenase).

To identify aphid-specific clusters, the sequences from the group II clusters were subjected to TBLASTX searches against NCBI's nucleotide and EST databases. These analyses identified 119 putative aphid-specific clusters, representing 729 sequences, with no matches (E value $< 1e^{-5}$) in other organisms (Table S9). BLASTN analysis of the *A. pisum* sequences from the 119 clusters against the recently generated salivary gland transcriptome for this species (Carolan et al. 2011) showed that sequences of 22 clusters are present in the salivary gland. All *A. pisum* sequences in this group were annotated as hypothetical or uncharacterized in the latest *A. pisum* genome annotation in AphidBase 2.1 (ACYPI mRNA v2.1).

Putative *M. euphorbiae*-specific sequences and their spatial expression

The sequence clustering results also contained 24 clusters, representing 29 sequences that contained exclusively *M. euphorbiae* sequences or sequences of group III (Table S10) . None of them was found in the recently generated *M. euphorbiae* salivary gland

transcriptome (Atamian et al. 2012). The sequences in these clusters ranged in length between 117-4640 bp with only 4 sequences longer than 400 bp suggesting that most transcripts in this set are incomplete. Three clusters, each with a single *M. euphorbiae* contig, which represent the three largest sequences in group III, were selected for further analyses. These three contigs, *Me*_WB10316, *Me*_WB26189, and *Me*_WB28140 have large yet incomplete ORFs encoding 510, 1546, and 1037 amino acids, respectively. To exclude the possibility of an improper assembly, we confirmed the identity of these contigs by amplifying and re-sequencing them from independently prepared cDNAs. Using clone-specific PCR primers (Table 3.2), fragments similar in size to each of the three contigs *Me*_WB10316, *Me*_WB26189, and *Me*_WB28140 were amplified. Sequencing the PCR-amplified fragments confirmed the identity of the three amplified products indicating accurate assembly of the Illumina reads.

To confirm that these contigs were specific to *M. euphorbiae* and to identify the size of their full-length transcripts, RNA blot analysis was performed. Blots were prepared from mixed developmental stage RNA of *M. euphorbiae*, *A. pisum* or *M. persicae*. Interestingly, the RNA blot analysis of the three contigs detected only a single transcript greater than 10 kb in *M. euphorbiae* (Figure 4). No signal was detected in the *A. pisum* or *M. persicae* lanes suggesting that the contigs were *M. euphorbiae*-specific sequences. To determine if each contig encoded a different > 10 kb transcript or if the three contigs constituted non-overlapping parts of a single transcript, the tissue-specific expression of these contigs was determined by RT-PCR.

The respective cDNA templates were from whole adult aphids, dissected heads or gut tissues. As controls, *M. euphorbiae* tissue specific markers were needed. To this end, *M. euphorbiae* homologs of the *A. pisum*, sucrase (*APS1*) and *ApC002* genes, which are expressed primarily in gut and salivary gland tissues, respectively, were identified using TBLASTN analysis (Mutti et al. 2008; Price et al. 2007). *MeS1* and *MeC002* gene-specific primers were designed and used in RT-PCR. Consistent with the reported expression in *A. pisum*, the candidate *M. euphorbiae* sucrase *MeS1* and *MeC002* were primarily expressed in the gut and head tissues, respectively (Figure 5). Moreover, the tissue-specific expression analysis of the three putative *M. euphorbiae*-specific contigs showed similar expression pattern for all three contigs being expressed mainly in the gut (Figure 5). This tissue expression result further suggested that the three contigs constituted parts of a single transcript.

Using a set of primer combinations from the three identified contigs, we were able to amplify the gaps to join the three contigs (Figure 6). The assembled transcript was 15,198 bp in size, hereafter is referred to as *Me*_WB29764. It has an ORF of 5065 amino acids, but is incomplete transcript as both start and stop codons are missing.

Using the *Me*_WB29764 sequence, TBLASTN analysis against the *A. pisum* EST and genomic sequences was performed. Surprisingly, matches ($E < 1e^{-5}$) to four EST sequences (CN750073, CN750957, CN750862 and GD185970) were detected. The *A. pisum* ESTs were localized to a gap between *Me*_WB26189 and *Me*_WB28140 not represented in the initial *M. euphorbiae* contigs (Figure 6). Interestingly, no match was detected in the *A. pisum* genomic sequences. TBLASTN analysis of

*Me*_WB29764 sequences against sequences in the NCBI nucleotide (nr/nt) database identified additional matches to sea urchin, sea squirt, marine bacteria, and viral sequences (E value < $1e^{-8}$) as well as insect sequences annotated as ATP-dependent RNA helicase (E value < $1e^{-6}$).

To address the discrepancy that ESTs matching *Me*_WB29764 could be identified, while sequences in the current assembly of the *A. pisum* genome matching this contig appear to be absent, we performed DNA blot analysis with *A. pisum*, *M. euphorbiae* as well as *M. persicae* using a subclone of *Me*_WB10316 as a probe. As expected, hybridization signals were detected in *M. euphorbiae* (Figure 7). In addition, signals were also detected in both *A. pisum* and *M. persicae* genomes indicating the presence of similar sequences in both aphid species (Figure 7).

Discussion

For organisms for which full genome sequences are not available, transcriptome sequencing and *de novo* assembly provides an alternative to built genomics resources to guide future studies. High-throughput sequencing technologies, with deep coverage at base level resolution, ease of library preparation and requirement for low quantity of total RNA as starting material, made possible the inclusion of sequencing in studies aimed at finding answers to numerous biological questions. Moreover, transcriptome sequencing addresses the expressed part of the genome, which cannot be unequivocally predicted from the genome sequence alone. Upon genome sequence availability, the transcriptome sequences represent valuable resources for accurate gene prediction and identification of splice patterns. The discovery of a comprehensive set of expressed genes in an organism requires the construction of libraries from different tissues and biological conditions.

In this study, we sequenced libraries derived from mixed developmental stages of *M. euphorbiae* and *M. euphorbiae* exposed to different biotic or abiotic stresses using the Illumina sequencing platform and *de novo* assembled the reads into 24,137 contigs (N50= 2130 bp) using the SEED/Velvet/Oases approach (Bao et al. 2011). Various *de novo* transcriptome assembly algorithms are freely available, such as Velvet/Oases (45), Trinity (46), ABySS (47) and SOAP_denovo (48). Each of them has advantages and disadvantages with respect to sensitivity, precision, run time and memory usage, and one has to choose among different assemblers the most suitable for the specific application (Zhao et al. 2011). For this study we developed a method

to improve transcriptome assemblies by pre-processing the reads with a novel clustering approach (20). Being closely related to *M. euphorbiae* (Von Dohlen and Teulon 2003), we utilized the predicted gene set of the *A. pisum,* the only aphid genome available publicly, as reference to assess the quality of the contigs assembled in this study. More than 7000 *A. pisum*-predicted transcripts have at least 50% coverage by the *M. euphorbiae* transcriptome generated in this study providing a valuable resource for future gene expression analysis and identifying genes regulated by host-aphid interactions as well as other aphid related processes.

Gene onthology assignment predicted 38% of the contigs assigned to cellular component category to encode proteins with nuclear localization. Numerous nuclear proteins are known to regulate of a wide variety of biological processes including development and stress responses. Moreover the contigs assembled in this study were assigned to diverse molecular functions suggesting that our libraries provided a comprehensive representation of the *M. euphorbiae* transcriptome.

Using stringent criteria and a combination of MCL algorithm with subsequent BLAST searches, we identified sequences that are absent or unique in the aphid species used in this study, as well as sequences unique to *M. euphorbiae*. From the sequences absent in aphids but conserved in other insects as well as in *D. pulex*, inferences can be drawn related to the aphid biology. We identified 17 such clusters with sequences from all eight non-aphid species, used in this study, and no similarity to aphid EST and genomic sequences ($E < 1e^{-5}$) indicating that these clusters have been lost in aphids. Although unlikely, we cannot rule out the possibility that

sequencing the genomes of additional aphid species may identify sequences belonging to these 17 clusters.

Aphids possess a diverse symbiont community that includes the obligate bacterial symbiont *Buchnera aphidicola* (Buchner 1965; Munson et al. 1991) as well as several facultative symbionts. The establishment of obligate symbiotic relationship between *Buchnera* and aphid is estimated to be 150 to 250 million years ago (Baumann et al. 1997). This strict obligate symbiotic relationship has lead to loss of genes involved in certain amino acid biosynthetic pathways in both organisms. Consequently, the production of some amino acids in aphids has become dependent on host-symbiont cooperation (Hansen and Moran 2011). Similarly, it is suggested that the absence of some immunity-related genes in aphid species could be related to its relationship with symbionts (Gerardo et al. 2010). As was demonstrated by analyzing the *A. pisum* genome, this aphid lacks genes involved in recognition, signaling and killing of microbes mostly present in other insect species (Gerardo et al. 2010). These include many crucial components of the immune deficiency (IMD)-signaling pathway as well as the peptidoglycan-receptor proteins (PGRPs) that recognize the peptidoglycans present in cell walls of bacteria and lead to the activation of both the Toll and IMD/c-Jun NH(2)-terminal protein Kinase (JNK) pathways. Consistent with the *A. pisum* genome, *M. euphorbiae* also lacks clusters of immunity-related sequences. One such cluster contained the Drosophila gene cylindromatosis (*CYLD*), an ortholog of the human cylindromatosis tumor suppressor gene. In Drosophila, *CYLD* has been shown to be involved in regulation of *JNK*-induced cell death (Xue et

al. 2007), triglyceride content and antibacterial defense (Tsichritzis et al. 2007). Other clusters contain the unnamed Drosophila gene products CG2765 and CG4615. Using RNAi screens in Drosophila S2 cells, CG2765 was shown to play an essential role in the internalization of *Escherichia coli* and *Staphylococcus aureus*, Gram-negative and Gram-positive bacteria, respectively (Ulvila et al. 2011), while CG4615 was shown to be involved in phagocytosis of the fungal pathogen *Candida albicans* (Stroschein-Stevenson et al. 2006).

Experimental evidence suggests that the aphid immune system, similar to its amino acid metabolism, is shaped by inputs from both the host and the symbionts. Although no role has been identified for symbionts in aphid protection against bacterial pathogens, in *A. pisum* the facultative symbiont *Regiella insecticola* has a major effect on aphid resistance to the fungal pathogen *Pandora neoaphidis* (Scarborough et al. 2005). Moreover, the facultative symbionts *R. insecticola* and *Hamiltonella defensa* provide protection against the aphid parasitoid *Aphidius ervi* (Hansen et al. 2012). Therefore, aphids can be considered as "extended organisms" comprised of aphid and symbiont genomes and featuring an immune system resulting from inputs of two genomes (Poirié and Coustau 2011).

The remaining clusters with sequences missing in aphids are not functionally characterized in arthropods, except for one, CG4779, which was reported to be up-regulated by starvation in Drosophila (Gronke et al. 2005). It is worth mentioning that the other clusters with sequences absent in aphids, but not represented by each of the

eight non-aphid species, were not included in group I as we only focused on the highly conserved sequences.

More direct insight regarding the possible genes responsible for the aphid's unconventional biology can be obtained through the identification of sequences unique to these species. Initially, individual clusters represented by sequences from all four aphid species, but none of the eight non-aphid species used in this study, were selected. Extending our search to all organisms, by performing TBLASTX analysis against the NCBI nucleotide and EST databases ($E < 1e^{-5}$), we identified 119 clusters as putatively aphid-specific. However none of these genes have been characterized in aphids and therefore their functions are unknown. We hypothesize that some of these sequences are involved in signaling pathways responsible for the aphid's phenotypic plasticity or interaction with its primary and secondary symbionts. Moreover, the 22 clusters with sequences from salivary gland may play roles in interactions of aphids with their plant host.

We expect to have missed several additional potential aphid-specific clusters for the following reasons. The genomes of three of the aphids used in this study, *M. persicae, A. gossypii* and *M. euphorbiae*, have not been sequenced and transcriptome sequences are limited by expression. Furthermore, our criteria for cluster selection were highly stringent, requiring clusters to have sequences from each of the four aphid species. Future *wet-lab* experiments designed to analyze the set of uncharacterized aphid-specific sequences identified in this study are expected to considerably add to our current understanding of pathways operating in aphids.

The *M. euphorbiae*-specific clusters were mainly comprised of transcripts with relatively short and incomplete transcript sequences suggesting that they could be misrepresented in this category. Sequencing the genomes or developing additional full-length aphid transcriptomes will address this possibility. Although the relationship of these *M. euphorbiae* contigs with sequences from *M. persicae, A. gossypii* is not clear in the absence of full genome sequences, these contigs represent sequences that have diverged significantly ($E < 1e^{-5}$) from that of *A. pisum*.

We identified three *M. euphorbiae*-specific contigs, encoding long ORFs that were hypothesized to originate from the same transcript based on RNA blot analysis and tissue expression pattern. The accuracy of this assumption was demonstrated by amplifying the gaps between the three contigs. The *Me*_WB29764 assembled transcript (15,198 bp) was lacking its 5' and 3' ends and the exact size of the full-length transcript was difficult to estimate from RNA blots. Nevertheless, sequence information indicates that the full-length transcript is longer than 15.2 kb. Among the 36,961 *A. pisum*-predicted transcripts, only 31 transcripts are longer than 15 kb highlighting the rare nature of such large aphid transcripts. Although BLAST analysis of *Me*_WB29764 identified revealed weak similarity to sequences annotated as ATP-dependent RNA helicase (E value $< 1e^{-6}$), it is difficult to predict whether this protein functions as a helicase. RNA helicases function within cellular processes that involve RNA including transcription, splicing, translation, RNAi and RNA editing (Owttrim 2006; Rocak and Linder 2004; Tanner and Linder 2001). RNA helicase like genes are expected to have orthologs with high sequence similarity. The Drosophila ATP-

dependent RNA helicase sequence (GK11214) matching to *Me*_WB29764 (E value < $1e^{-6}$) encodes only a 1432 amino acid long protein and has high amino acid sequence identity (E value < $1e^{-176}$) to a predicted protein in *A. pisum* (ACYPI005239) with a similar size annotated as ATP-dependent RNA helicase. Thus, it is likely that *Me*_WB29764, although containing helicase related sequences, has a different function.

Although the *Me*_WB29764 sequence was not identified in the *A. pisum* genome by TBLASTX analyses, DNA blot analysis detected *Me*_WB29764-related sequences in both the *A. pisum* and *M. persicae* genomes. Therefore, it is likely that this gene is located in a part of the *A. pisum* genome, which has not been sequenced yet. The absence of detectable *Me*_WB29764-related transcripts in *A. pisum* and *M. persicae* in our RNA blot analyses may reflect that this gene is not expressed or expressed at low levels under the conditions in which these two aphids were reared. Since this gene is expressed mainly in the gut, it is possible that its expression is influenced by the plant host species on which the aphids are reared. Future experiments will address this possibility.

References

Atamian, H. S., and Kaloshian, I. 2012. Construction of RNA-Seq libraries from large and microscopic tissues for the Illumina sequencing platform. Pages 47-57 in: RNA Abundance Analysis, H. Jin and W. Gassmann, eds. Humana Press.

Atamian, H. S., Chaudhary, R., Dal Cin, V., Girke, T., and Kaloshian, I. *In planta* expression or delivery of potato aphid *Macrosiphum euphorbiae* effectors *Me*10 and *Me*23 enhances aphid fecundity. Mol. Plant-Microbe Interact. in press.

Bao, E., Jiang, T., Kaloshian, I., and Girke, T. 2011. SEED: efficient clustering of next-generation sequences. Bioinformatics 27:2502-2509.

Baumann, P., Moran, N. A., and Baumann, L. 1997. The evolution and genetics of aphid endosymbionts. BioScience 47:12-20.

Bhattarai, K. K., Xie, Q. G., Pourshalimi, D., Younglove, T., and Kaloshian, I. 2007. *Coi1*-dependent signaling pathway is not required for *Mi-1*-mediated potato aphid resistance. Mol. Plant-Microbe Interact. 20:276-282.

Blackman, R. L., and Eastop, V. F. 2000. Aphids on the World's Crops. John Wiley & Sons, Ltd, New York.

Bos, J. I., Prince, D., Pitino, M., Maffei, M. E., Win, J., and Hogenhout, S. A. 2010. A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). PLoS Genet. 6:e1001216.

Braendle, C., Davis, G. K., Brisson, J. A., and Stern, D. L. 2006. Wing dimorphism in aphids. Heredity 97:192-199.

Buchner, P. 1965. Endosymbiosis of animals with plant microorganisms. John Wiley, New York.

Carolan, J. C., Caragea, D., Reardon, K. T., Mutti, N. S., Dittmer, N., Pappan, K., Cui, F., Castaneto, M., Poulain, J., Dossat, C., Tagu, D., Reese, J. C., Reeck, G. R., Wilkinson, T. L., and Edwards, O. R. 2011. Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrthosiphon pisum*): a dual transcriptomic/proteomic approach. J. Proteome Res. 10:1505-1518.

Chan, C. K., Forbes, A. R., and Raworth, D. A. 1991. Aphid-transmitted viruses and their vectors of the world. Agric. Can. Tech. Bull. 3E:1–216.

Dixon, A. F. G., Kindlmann, P., Leps, J., and Holman, J. 1998. Why there are so few species of aphids especially in the tropics? Am. Soc. Nat. 129:580-592.

Evert, R. F., Eschrich, W., Medler, J. J., and Alfieri, F. J. 1968. Observations on penetration of linden branches by stylets of the aphid *Longistigma caryae*. Am. J. Bot. 55:860-874.

Gerardo, N. M., Altincicek, B., Anselme, C., Atamian, H., Barribeau, S. M., de Vos, M., Duncan, E. J., Evans, J. D., Gabaldon, T., Ghanim, M., Heddi, A., Kaloshian, I., Latorre, A., Moya, A., Nakabachi, A., Parker, B. J., Perez-Brocal, V., Pignatelli, M., Rahbe, Y., Ramsey, J. S., Spragg, C. J., Tamames, J., Tamarit, D., Tamborindeguy, C., Vincent-Monegat, C., and Vilcinskas, A. 2010. Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome Biol. 11:R21.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644-652.

Gronke, S., Mildner, A., Fellert, S., Tennagels, N., Petry, S., Muller, G., Jackle, H., and Kuhnlein, R. P. 2005. Brummer lipase is an evolutionary conserved fat storage regulator in Drosophila. Cell Metab. 1:323-330.

Hansen, A. K., and Moran, N. A. 2011. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. Proc. Natl. Acad. Sci. U.S.A. 108:2849-2854.

Hansen, A. K., Vorburger, C., and Moran, N. A. 2012. Genomic basis of endosymbiont-conferred protection against an insect parasitoid. Genome Res. 22:106-114.

Hebert, S. L., Jia, L., and Goggin, F. L. 2007. Quantitative differences in aphid virulence and foliar symptom development on tomato plants carrying the *Mi* resistance gene. Environ. Entomol. 36:458-467.

Huang, X., and Madan, A. 1999. CAP3: A DNA sequence assembly program. Genome Res. 9:868-877.

Huerta-Cepas, J., Marcet-Houben, M., Pignatelli, M., Moya, A., and Gabaldon, T. 2010. The pea aphid phylome: a complete catalogue of evolutionary histories and arthropod orthology and paralogy relationships for *Acyrthosiphon pisum* genes. Insect Mol. Biol. 19:13-21.

International Aphid Genomic Consortium 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biol. 8:e1000313.

Kaloshian, I., Lange, W. H., and Williamson, V. M. 1995. An aphid-resistance locus is tightly linked to the nematode-resistance gene, *Mi*, in tomato. Proc. Natl. Acad. Sci. U.S.A. 92:622-625.

Kaloshian, I., Kinsey, M. G., Ullman, D. E., and Wiliamson, V. M. 1997. The impact of *Meu1*-mediated resistance in tomato on longevity, fecundity and behavior of the potato aphid, *Macrosiphum euphorbiae*. Entomol. Exp. Appl. 83:181–187.

Katis, N. I., Tsitsipis, J. A., Stevens, M., and Powell, G. 2007. Transmission of Plant Viruses. Pages 353-377 in: Aphids as Crop Pests, H.F. van Emden and R. Harrington, eds. CABI, London.

Leonardo, T. E., and Muiru, G. T. 2003. Facultative symbionts are associated with host plant specialization in pea aphid populations. Proc. R. Soc. Lond. B Biol. Sci. 270:209-212.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J., and Wang, J. 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. Genome Res. 20:265-272.

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39:D225-229.

Moeller, F. W. 1973. The host plants of the potato aphid *Macrosiphum euphorbiae* and of closely related species. Wiss. Z. Univ. Rostock. Math. Naturwiss. Reihe. 22:1179-1184.

Munson, M. A., Baumann, P., and Kinsey, M. G. 1991. Buchnera gen. nov. and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. Int. J. Syst. Evol. Microbiol. 41:566–568.

Mutti, N. S., Louis, J., Pappan, L. K., Pappan, K., Begum, K., Chen, M. S., Park, Y., Dittmer, N., Marshall, J., Reese, J. C., and Reeck, G. R. 2008. A protein from the salivary glands of the pea aphid, *Acyrthosiphon pisum*, is essential in feeding on a host plant. Proc. Natl. Acad. Sci. U.S.A. 105:9965-9969.

Nault, L. R. 1997. Arthropod transmission of plant viruses: a new synthesis. Ann. Entomol. Soc. Am. 90:521-541.

Owttrim, G. W. 2006. RNA helicases and abiotic stress. Nucleic Acids Res. 34:3220-3230.

Poirié, M., and Coustau, C. 2011. The evolutionary ecology of aphids' immunity. Invertebrate Surviv. J. 8:247-255.

Price, D. R., Karley, A. J., Ashford, D. A., Isaacs, H. V., Pownall, M. E., Wilkinson, H. S., Gatehouse, J. A., and Douglas, A. E. 2007. Molecular characterisation of a candidate gut sucrase in the pea aphid, *Acyrthosiphon pisum*. Insect Biochem. Mol. Biol. 37:307-317.

Radcliffe, E. B., and Ragsdale, D. W. 2002. Aphid-transmitted potato viruses: The importance of understanding vector biology. Am. J. Potato Res. 79:353-386.

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S. D., Mungall, K., Lee, S., Okada, H. M., Qian, J. Q., Griffith, M., Raymond, A., Thiessen, N., Cezard, T., Butterfield, Y. S., Newsome, R., Chan, S. K., She, R., Varhol, R., Kamoh, B., Prabhu, A. L., Tam, A., Zhao, Y., Moore, R. A., Hirst, M., Marra, M. A., Jones, S. J., Hoodless, P. A., and Birol, I. 2010. *De novo* assembly and analysis of RNA-seq data. Nat. Methods 7:909-912.

Rocak, S., and Linder, P. 2004. DEAD-box proteins: the driving forces behind RNA metabolism. Nat. Rev. Mol. Cell Biol. 5:232-241.

Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E., and Williamson, V. M. 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc. Natl. Acad. Sci. U.S.A. 95:9750-9754.

Scarborough, C. L., Ferrari, J., and Godfray, H. C. 2005. Aphid protected from pathogen by endosymbiont. Science 310:1781.

Schulz, M. H., Zerbino, D. R., Vingron, M., and Birney, E. 2012. *Oases*: robust *de novo* RNA-seq assembly across the dynamic range of expression levels. Bioinformatics 28:1086-1092.

Stroschein-Stevenson, S. L., Foley, E., O'Farrell, P. H., and Johnson, A. D. 2006. Identification of Drosophila gene products required for phagocytosis of *Candida albicans*. PLoS Biol. 4:e4.

Tanner, N. K., and Linder, P. 2001. DExD/H box RNA helicases: from generic motors to specific dissociation functions. Mol. Cell 8:251-262.

Tsichritzis, T., Gaentzsch, P. C., Kosmidis, S., Brown, A. E., Skoulakis, E. M., Ligoxygakis, P., and Mosialos, G. 2007. A Drosophila ortholog of the human cylindromatosis tumor suppressor gene regulates triglyceride content and antibacterial defense. Development 134:2605-2614.

Ulvila, J., Vanha-aho, L. M., Kleino, A., Vaha-Makila, M., Vuoksio, M., Eskelinen, S., Hultmark, D., Kocks, C., Hallman, M., Parikka, M., and Ramet, M. 2011. Cofilin regulator 14-3-3zeta is an evolutionarily conserved protein required for phagocytosis and microbial resistance. J. Leukoc. Biol. 89:649-659.

van Dongen, S. (2000). Graph Clustering by Flow Simulation (University of Utrecht).

Von Dohlen, C. D., and Teulon, D. A. J. 2003. Phylogeny and historical biogeography of New Zealand indigenous Aphidini aphids (Hemiptera, Aphididae): An hypothesis. Ann. Entomol. Soc. Am. 96:107–116.

Xue, L., Igaki, T., Kuranaga, E., Kanda, H., Miura, M., and Xu, T. 2007. Tumor suppressor CYLD regulates JNK-induced cell death in Drosophila. Dev. Cell 13:446-454.

Zhao, Q. Y., Wang, Y., Kong, Y. M., Luo, D., Li, X., and Hao, P. 2011. Optimizing *de novo* transcriptome assembly from short-read RNA-Seq data: a comparative study. BMC Bioinformatics 12 Suppl 14:S2.

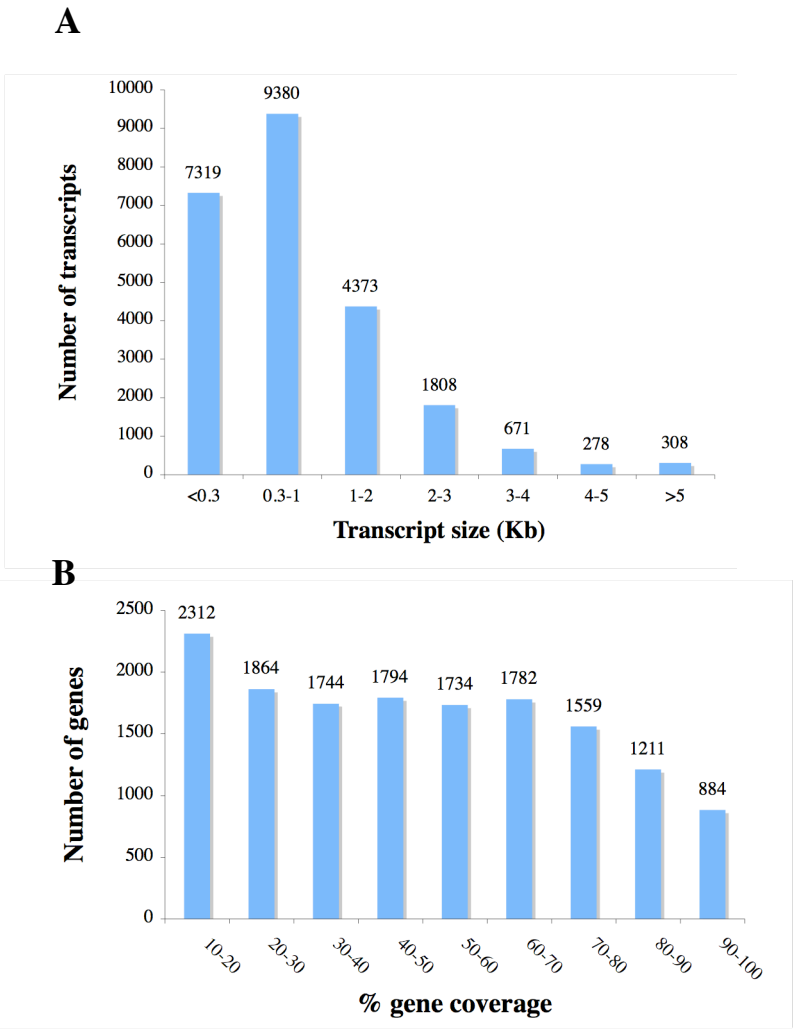
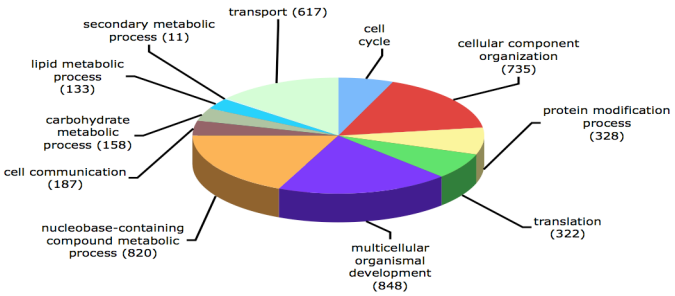
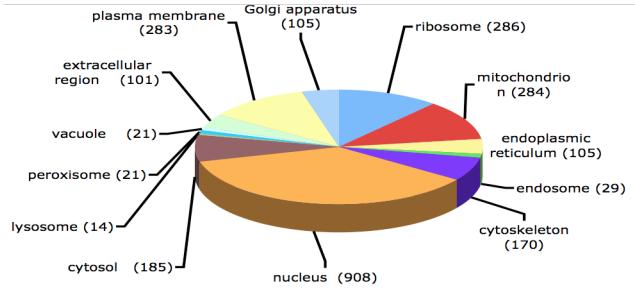


Fig. 3.1A-B Overview of *Macrosiphum euphorbiae* transcriptome assembly. (A) Size distribution of contigs obtained from *de novo* assembly of the high quality reads. The longest contig is 12,437 bp. (B) Histogram showing the coverage of *Acyrthosiphon pisum* predicted genes by the assembled *M. euphorbiae* contigs.





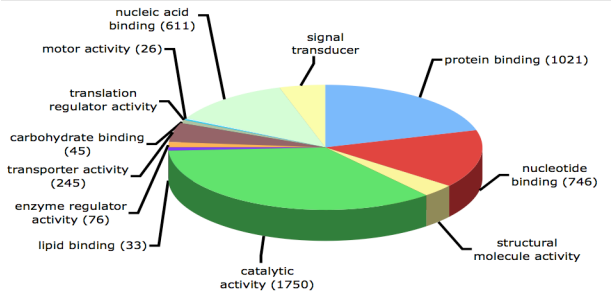


Fig. 3.2A-C Insect GO-slim terms associated with *Macrosiphum euphorbiae*. Pie charts giving the distribution of insect GO-slim terms associated with *M. euphorbiae* contigs represented in (A) biological process, (B) cellular component, and (C) molecular function categories.

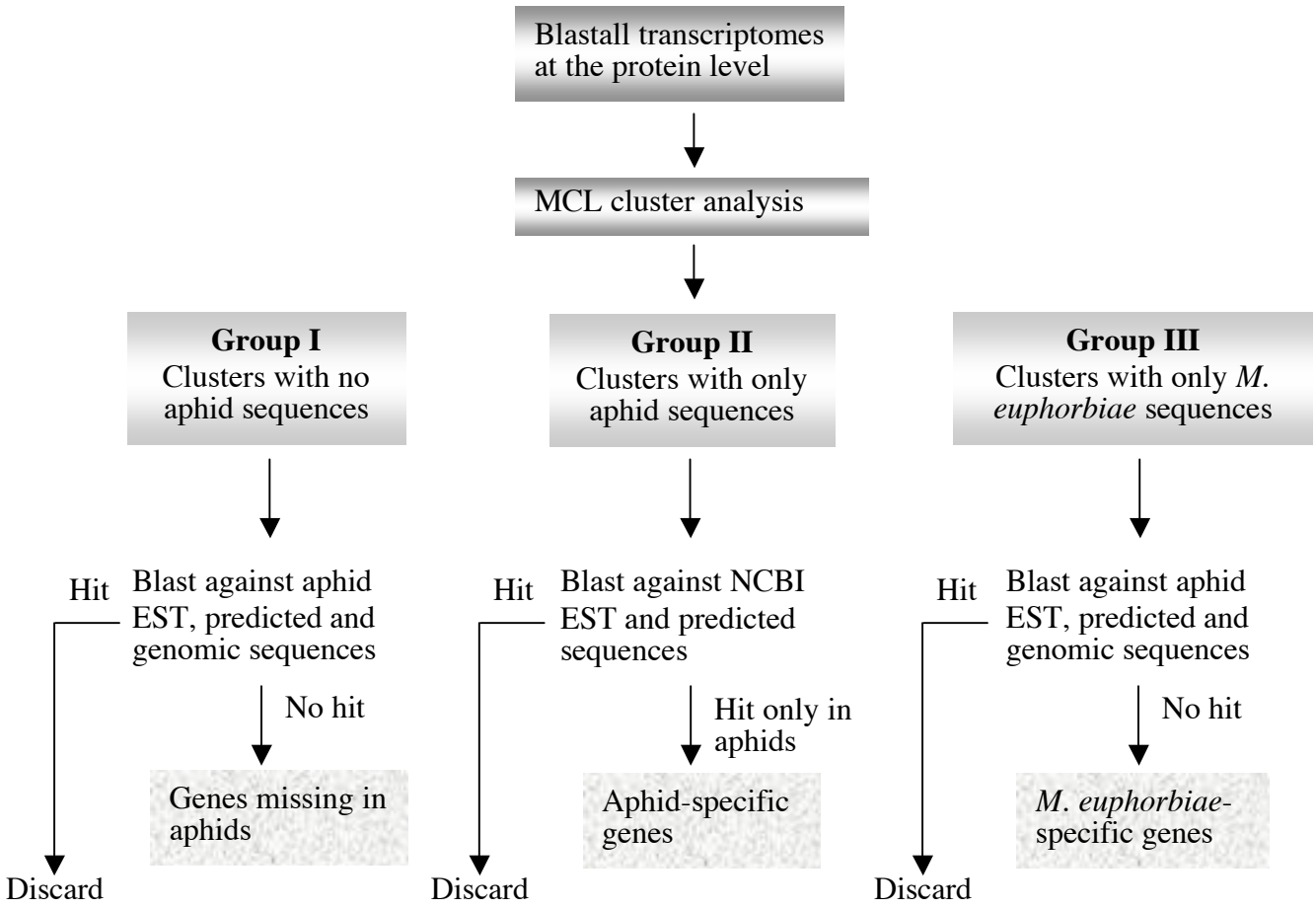


Fig. 3.3 Schematics of the comparative transcriptome analysis. Overall workflow of the bioinformatics analysis.

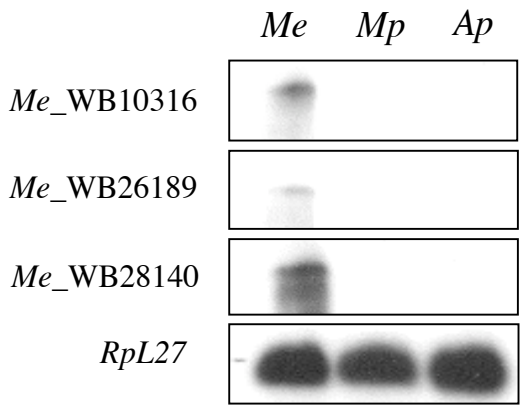


Fig. 3.4 RNA blot analysis of the putative *Macrosiphum euphorbiae*-specific transcripts. RNA blots were used to detect full-length Me_WB10316, Me_WB26189 and Me_WB28140 mRNA. Total RNA (7 μg) from mixed developmental stages of aphids was fractionated onto 1.2% agarose denaturing gels and blotted onto nylon membranes. Blots were hybridized with the potato aphid indicated probes. Ribosomal RNA *RpL27* probe was used as control for equal loading of RNA. *Me*: *M. euphorbiae*; *Mp: M. persicae*; *Ap*: *A. pisum*. The final wash of the blots was in 0.5x SSC and 0.1% SDS at 55°C and exposed to X-ray film with an intensifying screen for 1-3 days at -80°C except for RpL27 exposed for 16. Two blots were hybridized with each probe and a representative blot per probe is presented.

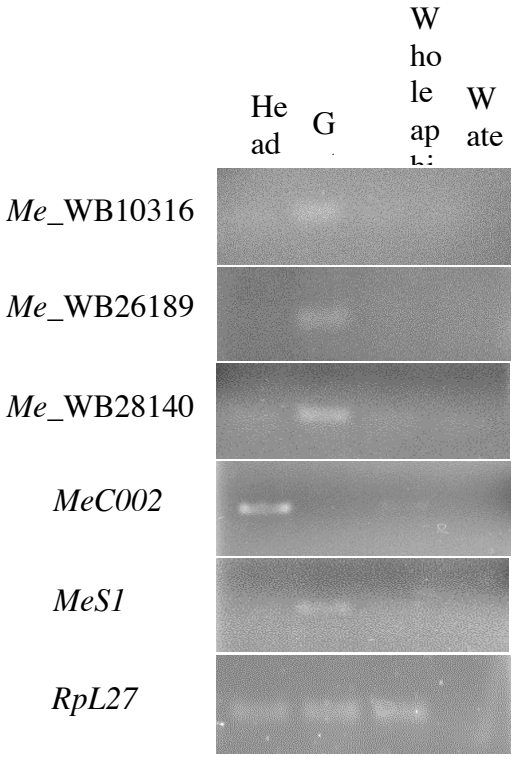


Fig. 3.5 Tissue-specific expression of putative *Macrosiphum euphorbiae*-specific transcripts. RNA isolated from whole mature aphids or dissected aphid body parts were used for RT-PCR with gene-specific primers for *Me*_WB10316, *Me*_WB26189, *Me*_WB28140, *MeC002*, or *MeS1*. Amplification of the ribosomal RpL27 gene was used as control.

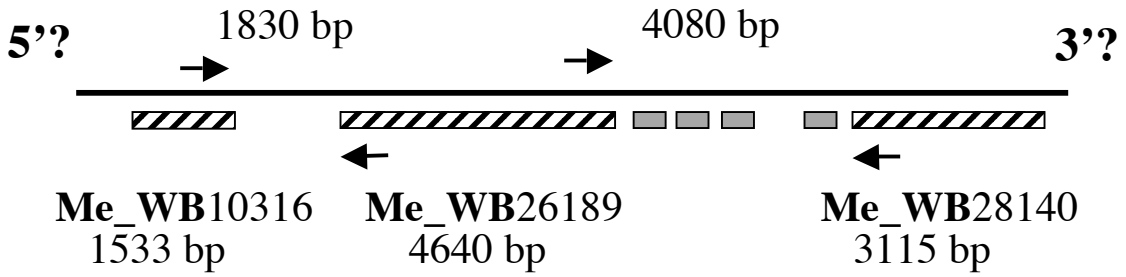


Fig. 3.6 Schematic of *Me*_WB29764 transcript. Schematic of the *Me*_WB29764 transcript showing the locations of the three potato-aphid contigs (shaded bars) and the primers used to amplify the gaps. Grey boxes represent locations of the matching *A. pisum* ESTs (GD185970, CN750073, CN750862 and CN750957).

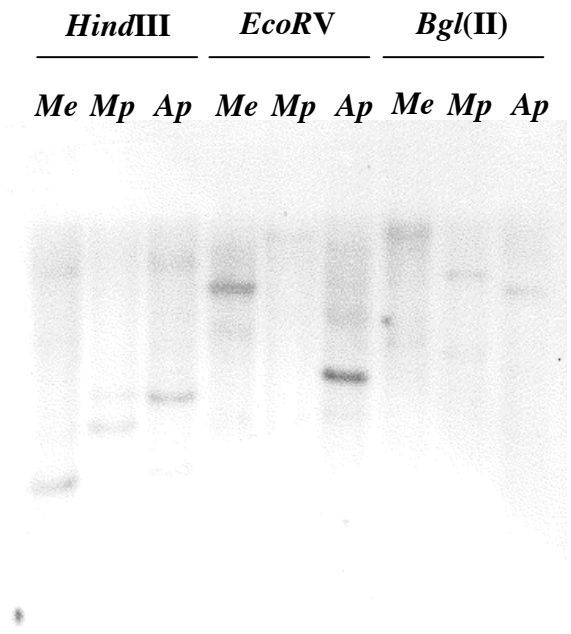


Fig. 3.7 DNA blot analysis of *Me*_WB29764. Genomic DNA from *M. euphorbiae (Me)*; *M. persicae (Mp)*; and *A. pisum (Ap)* were digested with the indicated restriction enzymes, separated on 0.9% agarose gel, blotted onto a nylon membrane, and hybridized with radiolabeled probe. The final wash of the blot was with 0.5x SSC and 0.1% SDS at 50°C and exposed to an X-ray film with an intensifying screen for two days at -80°C.

Table 3.1. Gene Onthology (GO) assignment for the clusters absent in aphids.

Onthology term	**Number of Clusters**
Biological process	
Immunity	1
Peroxisome fission	1
Phagocytosis	1
Carbohydrate metabolism	1
Amino acid metabolism	1
Unknown	12
Molecular function	
Sulfotransferase activity	1
Ubiquitin-specific protease activity	1
Nucleic acid binding	1
Binding	1
Protein binding	1
Rab GTPase binding	1
Oxidoreductase activity	1
Cation binding	1
Homogentisate 1,2-dioxygenase activity	1
Microtubles associated complex	1
Unknown	7

Table 3.2. Primers designed for sequencing Me_WB29764

Primer name	Sequence 5'-3'	Description
Me_WB10316F3	TGTTTAGCTTGATGGGGGCA	Primers to amplify and sequence Me_WB10316
Me_WB10316R3	AGTAGCCGTTAGCTTCATTG	
Me_WB26189F3	GGTAATATCTGGTATTGTGGCTTG	Primers to amplify and sequence Me_WB26189
Me_WB26189R3	ACTGTCGTTGTGTTTCACAGTAAC	
Me_WB26189F4	AGGGAGTTGACCCCTCATTTG	
Me_WB26189R4	CGGCAAGAACCATCAATGAC	
Me_WB26189F5	GCATATATACGTCTTGTTACCCGATG	
Me_WB26189R5	ACTCCAGGCCTCTATGAAGCAG	
Me_WB26189F6	CGTTAATCTCACGGTGATAGGT	
Me_WB26189R6	CCGGCTGACAATTAGACCAC	
Me_WB28140F3	AGTCTATCTTTGAAGGGCTCACAC	Primers to amplify and sequence Me_WB28140
Me_WB28140R3	CAACCTGCTGACATTCCTAACG	
Me_WB28140F4	CTATACTTGCCGCACCTACTTG	
Me_WB28140R4	GCTGGTCGCCATCGTTAAG	
Me_WB28140F5	GAGTCGGACACGTGGTTACCT	
Me_WB28140R5	TGAAGAAGCCGGGAAGGTAG	
MeCpwF1	AACAAAGACTGCAAGAACAAATG	Oligos used to sequence the gaps between the contigs.
MeCpwF2	GATAACACGTCCGAAAGTGAAG	
MeCpwF3	GAAGGCTGCTAAGCTCAACCT	
MeCpwF4	AAAGATCAACTCTGGTGGTACTTG	
MeCpwF5	AAGTACTGCAACCACATTGCTG	
MeCpwF6	ATCAAAACCGTTGATGTGTGGT	
MeCpwF7	TCTTATTTTCTGCTGGTATCATTG	

CHAPTER FOUR

In planta expression or delivery of potato aphid *Macrosiphum euphorbiae* effectors *Me*10 and *Me*23 enhances aphid fecundity.

Abstract

The interactions between aphids and their host plants seem to be analogous to those of plant-microbial pathogens. Unlike microbial pathogen effectors, little is known about aphid effectors and their ability to interfere with host immunity. To date, only three functional aphid effectors have been reported. To identify potato aphid (*Macrosiphum euphorbiae*) effectors, we developed a salivary gland transcriptome using Illumina technology. We generated 85 million Illumina reads from salivary glands and assembled them into 646 contigs. *Ab initio* sequence analysis predicted secretion signal peptides in 24% of these sequences suggesting that they might be secreted into the plant during aphid feeding. Eight of these candidate effectors with secretion signal peptides were functionally characterized using *Agrobacterium tumefaciens*-mediated transient overexpression in *Nicotiana benthamiana*. Two candidate effectors, *Me*10 and *Me*23, increased aphid fecundity suggesting their ability to suppress *N. benthamiana* defenses. Five of these candidate effectors, including *Me*10 and *Me*23, were also analyzed in tomato by delivering them through the *Pseudomonas syringae* type three secretion system. In tomato, only *Me*10 increased aphid fecundity. This work identified two additional aphid effectors with ability to manipulate the host for their advantage.

Introduction

Aphids (Hemiptera: Aphididae) are soft-bodied insects with piercing-sucking mouthparts that cause serious economic losses to cultivated crops. They damage plants directly by depleting nutrients and altering plant development, and indirectly by vectoring plant viruses and support the growth of the sooty mold fungus (Blackman and Eastop 2000). Some aphid species are globally distributed due to their polyphagous nature and ability to adapt to different environmental conditions (Margaritopoulos et al. 2009). The life cycle of aphids is somewhat complex, comprising of both sexual and asexual (parthenogenetic) modes of reproduction, the latter giving rise to live progeny (Blackman and Eastop 2000). Sexual reproduction occurs only in the fall season where eggs are laid on perennial plants for overwintering. Being hemimetabolous insects, aphids have no morphologically distinct larval or pupal stages. During asexual reproduction, females lay nymphs which, after three successive molts become adults (Moran 1992; Van Emden and Harrington 2007).

The potato aphid (*Macrosiphum euphorbiae*) has a broad host range including plants in Solanaceae, transmits a number of plant viruses and represents an aphid species of worldwide significance (Chan et al. 1991; Moeller 1973; Radcliffe and Ragsdale 2002). Its abundance and propensity to develop alatae makes this aphid species very important in viral epidemiology (Cerato et al. 1994; Singh and Boiteau 1986). In tomato (*Solanum lycopersicum*), resistance to potato aphids is mediated by the *Mi-1* gene that encodes a coiled-coil nucleotide-binding leucine-rich protein and

requires the Somatic Embryogenesis Receptor Kinase 1 (*SlSERK1*) (Mantelin et al. 2011; Rossi et al. 1998). The aphid effector recognized directly or indirectly by Mi-1 remains elusive.

Aphids are phloem feeders that use a pair of slender stylets to mostly move between cells until they reach the sieve element, where they feed for a prolonged period of time. Although the stylets path is mainly intercellular, stylets do also puncture cells and cause cell wall disturbance and damage to the plasma membranes of mesophyll and parenchyma cells (Moran et al. 2002; Pollard 1973; Tjallingii and Hogen Esch 1993). Unlike chewing insects that cause extensive tissue damage, this specialized aphid feeding behavior and interaction with its host, avoiding extensive mechanical tissue damage, is analogous to plant-biotrophic pathogen interactions where the pathogen is sustained in a localized area and is dependent on living host plant cells. Consistent with these observations, induction of plant genes associated with pathogen-induced response pathways have been reported as a result of aphid feeding (De Vos et al. 2005; Kaloshian and Walling 2005; Martinez de Ilarduya et al. 2003; Thompson and Goggin 2006).

Aphids release two types of saliva during feeding, soluble saliva and gelling saliva. The soluble saliva is in liquid form and is delivered along the penetration path and in the sieve element whereas the gelling saliva forms a proteinaceous sheath around the stylets as soon as it exits the stylet tip (Miles 1999; Tjallingii 2006). It has been hypothesized that constituents of aphid salivary secretions play crucial roles in modulating plant responses. Aphid saliva consists of a suite of bio-reactive

compounds, some of which may serve as cues to elicit plant defenses while others are expected to function in suppressing or circumventing plant defenses (Harris et al. 2003; Hogenhout and Bos 2011; Miles 1999).

It is speculated that aphid-host interactions shadow the commonly accepted zig-zag model of plant-microbial pathogen defense evolution described by Jones and Dangl (2006). In pathosystems, conserved sets of molecular signatures called pathogen- or microbe-associated molecular patterns (PAMPs/MAMPs) or general elicitors are recognized by pattern recognition receptors (PRRs) present at the host cell surface. PRR-mediated non-self recognition activates pattern-triggered immunity (PTI) in the host plant. PTI is associated with activation of downstream signaling pathways including mitogen activated protein kinases (MAPK) and WRKY transcription factors, induction of defense responses including production of reactive oxygen species, accumulation of pathogen-related (PR) proteins and callose deposition which collectively restrict microbial growth (Segonzac and Zipfel 2011). Pathogens can secrete small molecules or proteins known as effectors (Abramovitch et al. 2006; Birch et al. 2006; Davis et al. 2008; Kamoun 2006; van der Hoorn 2008) to counteract PTI and effectively parasitize and colonize the host plant. Extensive research using various approaches has led to the identification of hundreds of effectors secreted by bacterial, fungal and oomycete plant pathogens that were essential for understanding pathogenesis.

Unlike PAMPs and pathogen effectors, which have been extensively studied during the past decade, not much is known about herbivore associated molecular

patterns (HAMPs) or herbivore effectors (Hogenhout and Bos 2011; Miles 1999; Tjallingii 2006). Proteinaceous elicitor(s) with a size between 3 and 10 kD in the saliva of green peach aphid, *Myzus persicae,* have been shown to induce defense responses in Arabidopsis (De Vos and Jander 2009). Similarly, expressing the *M. persicae* proteins *Mp*10 and *Mp*42 in *Nicotiana benthamiana* resulted in a decrease in aphid fecundity (Bos et al. 2010). Moreover, induction of plant defense genes by feeding of piercing–sucking herbivores including aphids has been extensively demonstrated (Kempema et al. 2007; Martinez de Ilarduya et al. 2003; Rodriguez-Saona et al. 2010; Thompson and Goggin 2006). In contrast, little is known about how the salivary proteins of piercing-sucking herbivores may interfere with the plant immune system or manipulate host metabolites for their advantage. Aphids alter host primary metabolism and improve nutrient composition of the phloem sap to enhance their growth (Geordanengo et al 2010; Wilson et al 2011). However, it is not clear how aphids are able to cause these changes or the salivary secretion(s) responsible for these changes. A few aphid salivary proteins have been implicated in enhancing aphid performance. The salivary protein *Mp*C002 has been shown to enhance *M. persicae* fecundity or nymph production on *N. benthamiana* (Bos et al. 2010). *In vitro* analysis of the effect of the vetch aphid, *Megoura viciae,* saliva on forisomes provided direct evidence that aphid saliva has the ability to counteract plant defenses and prevent sieve tube plugging providing aphids with access to a continuous flow of phloem sap (Will et al. 2007). Moreover, some piercing-sucking insects have been shown to suppress the expression of plant defense genes and manipulate defense signaling

pathways to their advantage (Zarate et al. 2007; Zhang et al. 2011). Taken together, the ultimate outcome of interactions of piercing-sucking herbivores with their host plant is likely to depend on the salivary secretion and effectors produced by the herbivore and the ability of the plant to perceive these effectors and respond appropriately, and the ability of the salivary secretions to alter host metabolism.

A large number of pathogen effectors have been identified using homology-based searches while others were likely missed, as their sequences are unique. In aphids, candidate effector molecules have been identified using three approaches involving (1) sequencing the aphid salivary gland transcriptome or (2) the salivary gland proteome followed by prediction of the protein secretion signal, or (3) by direct sequencing the aphid salivary proteome (Bos et al. 2010; Carolan et al. 2011; Carolan et al. 2009; Harmel et al. 2008). In this study, we sequenced the salivary gland transcriptome of *M. euphorbiae* using the Illumina sequencing platform and assembled the reads into 646 contigs. Data mining of the assembled contigs identified 159 predicted *M. euphorbiae*-secreted proteins. Eight candidate effectors, with secretion signal peptides identified in the salivary transcriptome, were functionally characterized using *Agrobacterium tumefaciens*-mediated transient overexpression in *N. benthamiana* or delivered by *Pseudomonas syringae* type three secretion system (TTSS) into tomato plants. Using these assays, we identified two effectors, *Me*10 and *Me*23, which enhanced aphid fecundity.

Materials and Methods

Plant material and aphid colonies

Tomato (*Solanum lycopersicum*) cultivars UC82B and Moneymaker, *Nicotiana benthamiana* and tobacco (*Nicotiana tabacum)* NC-95 were used. Seedlings were transplanted into California mix II or sand. The plants were maintained in growth rooms at 24°C with 16-h-light and 8-h-dark photoperiod and 200 μmol m^{-2} s^{-1} light intensity and weekly fertilized with MiracleGro (18-18-21; Stern's MiracleGro Products).

Colonies of the parthenogenetic potato aphid (*Macrosiphum euphorbiae*) and green peach aphid (*Myzus persicae*) were reared on tomato cv. UC82B and tobacco NC-95 plants, respectively. The colonies were maintained in insect cages in a pesticide-free greenhouse at 22-26ºC supplemented with light for 16-h daylength. Age-synchronized, one-day-old, adult *M. euphorbiae* aphids were produced as described in Bhattarai et al. (2007).

RNA extraction and cDNA synthesis

RNA was isolated from 200 dissected *M. euphorbiae* salivary glands and used for RNA-Seq library preparation as described previously (Atamian and Kaloshian 2012). Similarly, RNA was isolated from whole aphids and 20 dissected heads and salivary glands or guts. cDNA was synthesized from 100 ng DNase-treated whole body, head

and salivary glands, or gut RNA using Superscript III (Invitrogen) reverse transcriptase enzyme and oligo-dT primers according to the manufacturer's recommendations.

Library construction, sequencing and *de novo* assembly

A detailed procedure of RNA-Seq library preparation from salivary gland tissues has been described previously (Atamian and Kaloshian 2012). One lane of single-end 75-nucleotide long sequencing was performed with an Illumina Genome Analyzer II instrument at the Institute for Integrative Genome Biology, University of California Riverside. Data from the Illumina sequencing run was processed using the Illumina standard pipeline version 1.4. The sequence data generated in this study have been deposited in NCBI's Short Read Archive, accession number SRR547988. Redundancies in the data set were removed with the SEED NGS clustering tool (Bao et al. 2011). The reaming sequences were assembled with the Velvet/Oases assembler as described in (Bao et al. 2011; Schulz et al. 2012) and deposited at DDBJ/EMBL/GenBank under the accession GAAF00000000in NCBI or presented in Table S11.

Annotation and secretion signal prediction

Reciprocal TBLASTX analyses were performed between *M. euphorbiae* and *Acyrthosiphon pisum* sequences to identify the putative orthologs of the *M. euphorbiae* sequences generated in this study. These putative orthologs were annotated based on the latest *A. pisum* annotation (aphidbase_2.1_peptides). The *M. euphorbiae* sequences with no putative orthologs in *A. pisum* were annotated by performing reciprocal TBLASTX analysis against NCBI nucleotide (nt/nr) database. The annotated sequences were assigned to different GO categories based on available GO analysis. The full-length putative *A. pisum* orthologs of the *M. euphorbiae* sequences were subjected to *de novo* signal peptide prediction analysis using SignalP 4.0 and TargetP 1.1 programs.

Cloning *M. euphorbiae* salivary gland EST sequences

Eight of *M. euphorbiae* salivary gland EST sequences with predicted secretion signal in their putative *A. pisum* orthologs were cloned. Gene-specific primers were designed excluding the secretion signal peptide at the 5'-end, based on the *A. pisum* full-length sequences, and a start codon added (Table 4.1). The PCR amplified products, obtained using Phusion High-Fidelity Polymerase (New England BioLabs) were cloned into the pDONR207 (Invitrogen) and recombined into the binary vector pEarleyGate100 (Earley et al. 2006) and sequenced using Sanger sequencing. The clones were transformed into *Agrobacterium tumefaciens* strain GV3101 for transient

overexpression in *N. benthamiana.* A subset of these sequences were also recombined into the pVSP_*Ps*SPdes vector (Rentel et al. 2008), sequenced and transformed into *Pseudomonas syringae* pv. *tomato* strain DC3000/Δ*avrPto*/Δ*avrPtoB (Pst)* for assay in tomato plants (Nguyen et al. 2010).

Aphid bioassay on *N. benthamiana*

Recombinant *A. tumefaciens* containing candidate effector or GFP were grown overnight in LB media supplemented with rifampicin (25 μg/ml), gentamycin (14 μg/ml) and kanamycin (50 μg/ml) at 28°C. The cultures were diluted to a final OD_{600} of 0.3 in an induction buffer (10 mM MES, 10 mM $MgCl_2$, 150 μM acetosyringone, pH = 5.6). Leaves of 4–5 week-old *N. benthamiana* plants were agroinfiltrated using a needle less syringe. Two leaves of four plants were infiltrated per construct. On day two after infiltration, four *M. persicae* adults were caged on each infiltrated leaf. The following day, the adults were removed leaving four first-instar nymphs. On day 6, the four-day-old nymphs were moved to a leaf of a plant infiltrated 2 days earlier. On day 9, the adult aphids were moved again to a leaf of a plant infiltrated 2 days earlier. On day 12, the same adults were moved to a plant infiltrated 2 days earlier. Aphid survival was counted at 6, 9, 12 and 15 days after the start of the experiment and the number of newly produced nymphs was counted on day 9, 12 and 15. The average number of nymphs produced per leaf sample was calculated by dividing the number of nymphs produced by the number of live adult aphids on days 9, 12 and 15.

Aphid bioassay on tomato

The *Pst* culture was grown on King's B Medium (KBM) plates containing rifampicin (25 μg/ml), carbenicillin (100 μg/ml) and kanamycin (50 μg/ml) at 30°C for two days. A single colony was inoculated into 200 μl of liquid KBM, plated onto KBM and incubated for another day at 30°C. A good healthy lawn on the plate was resuspended in 10 mM $MgCl_2$. Whole plants were vacuum infiltrated with *Pst* (1 x 10^3 CFU/ml) containing a candidate effector or GUS in 1 mM $MgCl_2$ and 0.02% Silwet L-77. Plants were infested with nine one-day-old adult *M. euphorbiae* 24 h after infiltration. Five plants were used per construct. Aphid fecundity was assessed by counting the number of nymphs and removing them daily for five days.

Gene expression analysis

Semi-quantitative RT-PCR analysis was performed using cDNA prepared from whole aphids or dissected aphid heads and salivary glands or guts using gene-specific primer pairs (Table 4.2). *C002* (Mutti et al. 2008) and *Sucrase (Suc)* (Price et al. 2007) were used as tissue-specific markers (Table S4) and the ribosomal gene L27 (NM_001126221) was used as internal control. PCR was performed in 25 µl with 40 ng of cDNA template, 1 X PCR buffer, 2.5 mM $MgCl_2$, 0.4 mM dNTPs, 2 units of *Taq* DNA polymerase, and 10 µM of each forward and reverse primers. The PCR

program was initialized at 94°C for 5 min, followed by 23 cycles of 94°C for 1 min, 58°C for 45 s, and 72°C for 1 min, with a final extension at 72°C for 10 min.

Statistical analyses

For the aphid fecundity assays, data were normalized to GFP or GUS. Statistical analysis with aphid data from *N. benthamiana* plants was performed using one-tailed Flinger-Policello test followed by Bonferrioni adjustment (BenMamoun 2006). Student's *t*-test was used for aphid data from tomato plants.

Results and discussion

Transcriptome sequencing, assembly and annotation

To analyze the *M. euphorbiae* salivary gland transcriptome, an RNA-Seq library was prepared from 200 dissected salivary glands of adult aphids. To determine the quality of the library, an aliquot was cloned and 10 clones were subjected to Sanger sequencing. TBLASTX analysis of the clones against the nonredundant database at National Center for Biotechnology Information (NCBI) identified unique sequences in each clone suggesting that the library was not biased for certain transcripts.

A total of 85 million reads were generated from this RNA-Seq library of which 4 million were unique. Initially, the generated reads were assembled using the Velvet/Oases assembler (Schulz et al. 2012). Further improvement of the assembly was achieved by reducing the redundancy in the data set with the SEED program, a novel clustering algorithm for next-generation sequencing data, which resulted in longer contigs (Bao et al. 2011). In total, 646 contigs were generated with 62% of the assembled contigs longer than 500 bp.

To determine how many of the *M. euphorbiae* salivary gland contigs have putative orthologs in the closely related species *Acyrthosiphon pisum*, whose genome has been sequenced, reciprocal TBLASTX analysis was performed. A total of 551 *M. euphorbiae* contigs were identified with sequences orthologous to 460 *A. pisum* transcripts (Table S11). Some of the *M. euphorbiae* contigs matched to non-overlapping regions of the same *A. pisum* transcript explaining the larger number of

M. euphorbiae orthologous contigs. Consequently, these contigs were annotated according to version 2 of the *A. pisum* annotation (International Aphid Genomics Consortium 2010). Out of the 460 *A. pisum* transcripts identified in this study, 155 were also identified in the *A. pisum* salivary glands by Carolan et al. (2011). Putative orthologs for an additional 41 contigs were identified by performing reciprocal BLASTX analysis against the Uniprot database and annotated accordingly (Table S11). Although these *M. euphorbiae* contigs did not have orthologs in *A. psium,* we cannot exclude the possibility that they may be present in other aphid species. Orthologs could not be identified for 54 *M. euphorbiae* contigs (Table S11). They could be the result of inaccurate assembly of the Illumina reads, although it is possible that some of these sequences are correctly assembled and consequently unique to *M. euphorbiae*.

To determine the putative functions of the *M. euphorbiae* contigs, we used the database [Comment: would expect name of database here.] containing gene ontology (GO) assignments of all the publicly available *A. pisum* ESTs. The functional classification of the contigs based on GO terms showed enrichment for the classes "translation", "metabolic process", and "transport" in the GO category "biological process" (Fig. 1A). Aphids have a pair of salivary glands each consisting of a principal gland and two accessory glands (Ponsen 1972; Weidemann 1968). Besides the expected cell maintenance processes, the cells of the salivary glands undergo cycles of secretory activities (Miles 1999). Thus, transcripts grouped under

the “translation” category are likely to serve both functions. To fulfill their respective roles, those proteins destined to be delivered in the saliva and potentially having roles in the interactions with the plant host, are expected to have secretion signals allowing them to cross cell membranes into the salivary canal. Consequently, it is not surprising that transcripts predicted to have a transport function were enriched in this organ. On the other hand, the GO terms in the “molecular function” category showed more distributed and diverse enrichments of various molecular activity categories (Fig. 1B).

Identification of sequences with secretion signal peptides

The amino acid sequences of the putative full-length *A. pisum* orthologs of the identified *M. euphorbiae* contigs were analyzed with the SignalP 4.0 (Petersen et al. 2011) and TargetP (Emanuelsson et al. 2000) prediction softwares trained to identify signal peptides. Of the 460 examined sequences, 125 and 159 were predicted to have putative signal peptides predicted by SignalP 4.0 (hidden Markov model scores of higher than 0.45) and TargetP (predefined set of cutoffs that yields specificity >0.95 on the TargetP test sets), respectively (Table S11). TargetP predicted a signal peptide in 121 sequences predicted by SignalP. Moreover, it predicted an additional 38 sequences that were not predicted by SignalP. In the *A. pisum* salivary gland transcriptome, 30% (1074/324) of the transcripts were predicted to have signal peptides using SignalP 3.0 (Carolan et al. 2011). Using the version 4.0 of the same

program we predicted 27% (460/125) of the salivary gland transcripts to have signal peptide. Around 42% of the sequences with predicted signal peptides in this study were also identified by Carolan et al. (2011). This low overlap of sequences with signal peptides between the two studies suggests that the salivary gland sequences from both species are incomplete.

Selection of clones for functional analysis and their tissue-specific expression

An efficient way to investigate the roles of aphid candidate effectors *in planta,* is to transiently overexpress them in *N. benthamiana* using *Agrobacterium tumefaciens* and assay with a population of aphids that are adapted to plants with nicotine such as *M. persicae*. A similar approach, using *N. benthamiana* leaf discs, was successfully used to evaluate *M. persicae* effectors (Bos et al. 2010). In order to use this assay system, we choose *M. euphorbiae* effectors with secretion signals and putative orthologs in *M. persicae*. To identify the *M. persicae* orthologs of our set of salivary gland expressed *M. euphorbiae* contigs, we reassembled the publically available *M. persicae* ESTs and used them in reciprocal TBLASTX analysis with the *M. euphorbiae* salivary gland transcripts. We chose eight *M. euphorbiae* salivary gland contigs with putative orthologs in *M. persicae,* four with annotations and four encoding yet uncharacterized proteins (Table 1). We cloned the *M. euphorbiae* ORFs encoding the mature proteins corresponding to these sequences, excluding the signal peptide. These eight ORFs (*Me5*, *Me10*, *Me13*, *Me14*, *Me17*, *Me20*, *Me23*, and

Me25) have 72-85% amino acid sequence identity to the corresponding *M. persicae* EST contig (Table 1). The relationship among the *M. persicae* and *A. pisum* orthologs of these eight *M. euphorbiae* clones was also demonstrated by reciprocal TBLASTX comparisons.

Four proteins were chosen for further study based on possible effector functions suggested by their annotation. Based on *A. pisum* annotation, *Me5* encodes a trehalase. Trehalose accumulation is associated with Arabidopsis defense against aphids (Singh et al. 2011). Therefore, trehalase secreted by the aphid may hydrolyze trehalose and counteract host defenses. *Me14* encodes a lipase; members of this superfamily of proteins have diverse roles including defense against oxidative stress (Horne et al. 2009). *Me23* encodes a glutathione peroxidase (GPX) with potentially protective role against oxidative burst (Lamb and Dixon 1997); while *Me25* encodes a carbonic anhydrase 2, which may function in catalysis of aldehydes induced during aphid feeding (Gosset et al. 2009). Of the four uncharacterized proteins, conserved domain search identified known domains for two, *Me*13 and *Me*17 (Table 1).

To confirm that the eight genes were expressed in the salivary glands, we evaluated the accumulation of their transcripts in dissected head and glands or gut tissues of *M. euphorbiae* (Fig. 2). Seven genes were expressed in the head and glands and not in the gut suggesting that their corresponding proteins are produced in the salivary glands, while *Me5* transcripts were detected in all the tissues tested suggesting either tissue unspecific expression of this gene or expression of its paralogs in different

tissues. As seven of the genes tested exhibited tissue-specific expression, it was unlikely that our tissue-specific cDNAs were contaminated. Orthologous transcripts of five of these genes, *Me10*, *Me13*, *Me17*, *Me20* and *Me23*, were also identified in *A. pisum* salivary gland transcriptome while peptides corresponding to *Me*5, *Me*10 and *Me*23 orthologs were identified in the *A. pisum* salivary gland proteome (Carolan et al. 2011). The presence of transcripts to most of our selected salivary gland genes in both *A. pisum* and *M. persicae*, aphid species with narrow and broad host ranges, respectively, suggests a general rather than specialized roles for these genes.

Evaluating the role of candidate *M. euphorbiae* effectors in aphid defense in *N. benthamiana*

To investigate the roles of *M. euphorbiae* candidate effectors *in planta,* we transiently expressed the selected *M. euphorbiae* proteins in *N. benthamiana* and assayed with a population of *M. persicae* adapted to feeding on tobacco (Kim and Jander 2007). For the transient expression assays, a large area of a *N. benthamiana* leaf was agroinfiltrated with the recombinant binary pEarleyGate100 vector expressing a candidate *M. euphorbiae* protein. Expression of green fluorescent protein (GFP) was used as control to monitor expression and aphid fecundity. Twenty-four hour after infiltration, four *M. persicae* adults were caged on each infiltrated leaf exposing the infiltrated area to aphids. The following day, the adults and newly born nymphs were removed leaving four first-instar nymphs. These remaining aphids were moved to a

fresh plant with recently agroinfiltrated leaves on day 6, day 9 and day 12. This schedule was based on efficient GFP expression. These nymphs became adults on day 8 and their fecundity was evaluated on days 9, 12 and 15. The average nymph production per aphid was calculated and normalized to those on the GFP control. Two candidate effectors *Me*10 ($P = 0.004$) and *Me*23 ($P = 0.01$) increased significantly *M. persicae* fecundity compared to the GFP control (Fig. 3). There was no significant difference in aphid fecundity on the plants expressing the remaining six effectors (Fig. 3). Although the increase in aphid fecundity on plants expressing *Me*10 or *Me*23 was modest, this increase in aphid performance was consistent in three independent experiments. This suggests that *Me*10 and *Me*23 altered *N. benthamiana* responses for aphid's advantage. Contributions of a single effector in manipulating host responses could be minor and this effect could be amplified when combined with additional components of the cocktail of effectors secreted by the aphid. Therefore, simultaneous expression of a combination of effectors, with no effect individually, may also result in enhancement of aphid performance.

To date, only a single aphid effector *Mp*C002, has been identified to increase aphid fecundity (Bos et al. 2010). Although it remains unknown how *Mp*C002 alters *N. benthamiana* responses to *M. persicae*, the *A. pisum* ortholog *ApC002* has been implicated in aphid orientation and feeding (Mutti et al. 2008). *Ap*C002 is also secreted inside the plant host (Mutti et al. 2008). To date, nine and 17 salivary proteins have been identified from *A. pisum* and *M. persicae*, respectively (Carolan et

al. 2009; Harmel et al. 2008). Interestingly, one of the *A. pisum* salivary proteins is the ortholog of *Me*10 the effector with an unknown function. The second *M. euphorbiae* effector, *Me23,* we identified to increase aphid fecundity is predicted to encode a GPX and could be involved in reducing H_2O_2 and function as an antioxidant to enhance aphid virulence and reduce the effect of the oxidative burst triggered by aphid feeding (Martinez de Ilarduya et al. 2003).

Evaluating the role of candidate *M. euphorbiae* effectors in aphid defense in tomato

To feed, aphids probe the plant host tissue with a pair of stylets and secret both watery and gelling saliva in this process. During probing and stylet penetration cells are punctured and saliva is delivered inside the host cytoplasm. Therefore, aphid feeding can be compared with gram-negative plant pathogenic bacteria that possess a TTSS to invade and colonize the host cell by injecting virulence effectors (Buttner and He 2009; Cornelis and Van Gijsegem 2000). Therefore, we took the advantage of an existing bacterial system to deliver candidate *M. euphorbiae* effectors into the tomato cells. The TTSS of the bacterial pathogen *Pseudomonas syringae* pv. *tomato* (*Pst*) have been successfully utilized to deliver oomycyte *Hyaloperonospora parasitica* effectors into Arabidopsis cells (Rentel et al. 2008; Sohn et al. 2007).

We characterized the function of five of the eight candidate effectors tested in *N. benthamiana* by delivering them through the *Pst* TTSS into the tomato cell cytoplasm and assaying aphid performance on the plants. To deliver aphid effectors through the TTSS, we used the pVSP_*Ps*SPdes expression vector that has the promoter and secretion-translocation signal of the *Pst* effector AvrRpm1 (Guttman and Greenberg 2001; Rentel et al. 2008). We introduced this vector in a less virulent strain of *Pst DC3000, Pst* DC3000Δ*avrPto*/Δ*avrPtoB* (Nguyen et al. 2010), lacking two strong virulent effectors *avrPto* and *avrPtoB,* to minimize the interference of bacterial effectors to better evaluate subtle differences in aphid performance. The roles of the five aphid effectors were evaluated in tomato by assessing their effect on the *M. euphorbiae* fecundity. To perform the fecundity assays, 4-5-week-old tomato plants were vacuum infiltrated with the recombinant *Pst* expressing the aphid constructs or β-glucuronidase (GUS as control. Each plant was infested with nine one-day-old, age-synchronized, adult aphids 24 h post infiltration. Nymph production was evaluated for the next five days. The average nymph production per aphid was calculated and compared with the GUS control. Of the five candidate effectors tested, only *Me*10 significantly increased ($P = 0.021$) aphid fecundity on tomato (Fig. 4). Thus, indicating that *Me*10 is able to manipulate tomato plants for aphid's advantage as it did in *N. benthamiana.* None of the remaining effectors affected aphid fecundity (Fig. 4). We did not find statistically significant differences in aphid performance in tomato plants infected with *Pst* expressing *Me*23 compared to the GUS control suggesting that *Me*23 is not able to alter tomato responses to detectable levels. It is

possible that the different mode of effector delivery, the altered aphid assay used for tomato compared to *N. benthamiana* or the shorter time exposure of aphids to plant expressing the effector, did not allow detection of subtle differences in plant responses. Alternatively, *N. benthamiana* may be less tolerant than tomato to aphid infestation and, therefore, more suitable to detect the weak effects of *Me*23 (Goodin et al. 2008).

Since *Me*10 is uncharacterized and has no known functionally conserved domains, it is difficult to speculate how it manipulates plant responses. Future experiments should elucidate this role. Nevertheless, this experiment showed that *Pst* TTSS can be used for delivery of aphid effectors *in planta* to evaluate aphid performance. This approach will allow the evaluation of aphid effectors, without the need to express them by developing stable transgenic plants, in hosts in which Agrobacterium-mediated transient expression does not work consistently or is not feasible to perform.

*Me*10 and *Me*23 and their putative *A. pisum* and *M. persicae* orthologs

Two genes of the *A. pisum* genome (ACYPI002439 and ACYPI38240) are annotated as glutathione peroxidases. The amino acid sequence identity between *Me*23 and the *A. pisum* proteins ACYPI002439 and ACYPI38240 is 87% and 42%, respectively, while the amino acid sequence identity between *Me*23 and a putative *M. persicae*

ortholog is 70% (Table 1). It is not clear how many *Me*23 orthologs are present in the *M. persicae* genome, as the genome of this aphid has not been sequenced. The alignment of the amino acid sequences from the three aphid species shows blocks of conserved regions (Fig. 5). Near the N-terminus, within one of these conserved regions, a deletion of eight amino acids is present in the *M. persicae* protein. The presence of such insertion/deletions might alter the protein's function.

Me10 encodes an uncharacterized protein. BLAST analyses indicated that *Me10* is a single-copy gene in *A. pisum* (Table 1). The amino acid sequence of *Me*10 has 76% and 88% sequence identity to *M. persicae* and *A. pisum* orthologs, respectively. Although these three aphids belong to the same tribe, Macrosiphini, within the aphid subfamily aphidinae, it is not surprising that *M. euphorbiae* sequences are more similar to *A. pisum,* since *M. euphorbiae* is phylogenetically more closely related to *A. pisum* than *M. persicae* (Von Dohlen et al. 2006).

References

Avdiushko, S., Croft, K. P., Brown, G. C., Jackson, D. M., Hamilton-Kemp, T. R., and Hildebrand, D. 1995. Effect of volatile methyl jasmonate on the oxylipin pathway in tobacco, cucumber, and Arabidopsis. Plant Physiol. 109:1227-1230.

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J. W., and Elledge, S. J. 1996. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. Cell 86:263-274.

Branch, C., Hwang, C. F., Navarre, D. A., and Williamson, V. M. 2004. Salicylic acid is part of the *Mi-1*-mediated defense response to root-knot nematode in tomato. Mol Plant-Microbe Interact 17:351-356.

Chao, W., Gu, Y., Pautot, V., Bray, E., and Walling, L. 1999. Leucine aminopeptidase RNAs, proteins, and activities increase in response to water deficit, salinity, and the wound signals systemin, methyl jasmonate, and abscisic acid. Plant Physiol. 120:979-992.

Cooper, W. R., and Goggin, F. L. 2005. Effects of jasmonate-induced defenses in tomato on the potato aphid, *Macrosiphum euphorbiae*. Entomol. Exp. Appl. 115:107-115.

Creelman, R. A., and Mullet, J. E. 1997. Biosynthesis and action of jasmonates in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:355-381.

Dropkin, V. H. 1969. Cellular responses of plants to nematode infections. Annu. Rev. Phytopathol. 7:101-122.

Ellis, C., and Turner, J. G. 2001. The *Arabidopsis* mutant *cev1* has constitutively active jasmonate and ethylene signal pathways and enhanced resistance to pathogens. Plant Cell 13:1025-1033.

Ellis, C., Karafyllidis, I., and Turner, J. G. 2002. Constitutive activation of jasmonate signaling in an *Arabidopsis* mutant correlates with enhanced resistance to *Erysiphae cichoracearum*, *Pseudomonas syringae*, and *Myzus persicae*. Mol. Plant-Microbe Interact. 15:1025-1030.

Farmer, E. E., and Ryan, C. A. 1992. Octadecanoid precursors of jasmonic acid activate the synthesis of wound-inducible proteinase inhibitors. Plant Cell 4:129-134.

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

Hammond-Kosack, K. E., and Parker, J. E. 2003. Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. Curr. Opin. Biotechnol. 14:177-193.

Howe, G. A., Lee, G. I., Itoh, A., Li, L., and DeRocher, A. E. 2000. Cytochrome P450-dependent metabolism of oxylipins in tomato. Cloning and expression of allene oxide synthase and fatty acid hydroperoxide lyase. Plant Physiol. 123:711-724.

Kaloshian, I., Lange, W. H., and Williamson, V. M. 1995. An aphid resistance locus is tightly linked to the nematode resistance gene, *Mi,* in tomato. Proc. Natl. Acad. Sci. USA 92:622-625.

Kaloshian, I., Kinsey, M. G., Ullman, D. E., and Williamson, V. M. 1997. The impact of *Meu1*-mediated resistance in tomato on longevity, fecundity and behavior of the potato aphid, *Macrosiphum euphorbiae*. Entomol. Exp. Appl. 83:181-187.

Kaloshian, I., Kinsey, M. G., Williamson, V. M., and Ullman, D. E. 2000. *Mi* -mediated resistance against the potato aphid *Macrosiphum euphorbiae* (Hemiptera: aphididae) limits sieve element ingestion. Environ. Entomol. 29:690-695.

Kaloshian, I. 2004. Gene-for-gene disease resistance: bridging insect pest and pathogen defense. J. Chem. Ecol. 30:2421-2439.

Kaloshian, I., and Walling, L. L. 2005. Hemipterans as plant pathogens. Annu. Rev. Phytopathol. 43:491-521.

Kohlmann, M., Bachmann, A., Weichert, H., Kolbe, A., Balkenhohl, T., Wasternack, C., and Feussner, I. 1999. Formation of lipoxygenase-pathway-derived aldehydes in barley leaves upon methyl jasmonate treatment. Eur. J. Biochem. 260:885-895.

Lee, G. I., and Howe, G. A. 2003. The tomato mutant spr1 is defective in systemin perception and the production of a systemic wound signal for defense gene expression. Plant J. 33:567-576.

Li, L., Li, C. Y., 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, L., Zhao, Y., McCaig, B. C., Wingerd, B. A., Wang, J., Whalon, M. E., Pichersky, E., and Howe, G. A. 2004. The tomato homolog of *CORONATINE-INSENSITIVE1* is required for the maternal control of seed maturation, jasmonate-signaled defense responses, and glandular trichome development. Plant Cell 16:126-143.

Li, Q., Xie, Q.-G., Smith-Becker, J., Navarre, D., and Kaloshian, I. 2006. *Mi-1*-mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling pathways. Mol. Plant-Microbe Interact. 19:655-664.

Martin, G. B., Bogdanove, A. J., and Sessa, G. 2003. Understanding the functions of plant disease resistance proteins. Annu. Rev. Plant Biol. 54:23-61.

Martinez de Ilarduya, O., and Kaloshian, I. 2001. *Mi-1.2* transcripts accumulate ubiquitously in root-knot nematode resistant *Lycopersicon esculentum*. J. Nematol. 33:116-120.

Martinez de Ilarduya, O., Moore , A. E., and Kaloshian, I. 2001. The tomato *Rme1* locus is required for *Mi-1*-mediated resistance to root-knot nematodes and the potato aphid. Plant J. 27:417-425.

Martinez de Ilarduya, O., Xie, Q.-G., and Kaloshian, I. 2003. Aphid-induced defense responses in *Mi-1*-mediated compatible and incompatible tomato interactions. Mol. Plant-Microbe Interact. 16:699-708.

Martinez de Ilarduya, O., Nombela, G., Hwang, C. F., Williamson, V. M., Muniz, M., and Kaloshian, I. 2004. *Rme1* is necessary for *Mi-1*-mediated resistance and acts early in the resistance pathway. Mol. Plant-Microbe Interact. 17:55-61.

Mewis, I., Appel, H. M., Hom, A., Raina, R., and Schultz, J. C. 2005. Major signaling pathways modulate Arabidopsis glucosinolate accumulation and response to both phloem-feeding and chewing insects. Plant Physiol. 138:1149-1162.

Milligan, S. B., Bodeau, J., Yaghoobi, J., Kaloshian, I., Zabel, P., and Williamson, V. M. 1998. The root-knot nematode resistance gene *Mi* from tomato is a member of leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. Plant Cell 10:1307-1319.

Moran, P. J., and Thompson, G. A. 2001. Molecular responses to aphid feeding in *Arabidopsis* in relation to plant defense pathways. Plant Physiol. 125:1074-1085.

Nombela, G., Williamson, V. M., and Muñiz, M. 2003. The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. Mol. Plant-Microbe Interact. 16:645-649.

Pascual, S., Avilés, M., Nombela, G., Muñiz, M., and Beitia, F. 2000. Development of *Bemisia tabaci* (biotype Q) on tomato cultivars with / without the *Mi* gene. Med. Fac. Landbouww. Univ. Gent. 65/2a:291-292.

Ryan, C. A. 2000. The systemin signaling pathway: Differential activation of plant defensive genes. Biochim. Biophys. Acta 1477:112-121.

Schenk, P., Kazan, K., Wilson, I., Anderson, J., Richmond, T., Somerville, S., and Manners, J. 2000. Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. Proc. Natl. Acad. Sci. USA 97:11655-11660.

Sivasankar, S., Sheldrick, B., and Rothstein, S. J. 2000. Expression of allene oxide synthase determines defense gene activation in tomato. Plant Physiol. 122:1335-1342.

Smith, P. G. 1944. Embryo culture of a tomato species hybrid. Proc. Am. Soc. Hortic. Sci. 44:413-416.

Stintzi, A., Weber, H., Reymond, P., Browse, J., and Farmer, E. E. 2001. Plant defense in the absence of jasmonic acid: The role of cyclopentenones. Proc. Natl. Acad. Sci. USA 98:12837-12842.

Taki, N., Sasaki-Sekimoto, Y., Obayashi, T., Kikuta, A., Kobayashi, K., Ainai, T., Yagi, K., Sakurai, N., Suzuki, H., Masuda, T., Takamiya, K., Shibata, D., Kobayashi, Y., and Ohta, H. 2005. 12-oxo-phytodienoic acid triggers expression of a distinct set of genes and plays a role in wound-induced gene expression in Arabidopsis. Plant Physiol. 139:1268-83.

Tao, Y., Xie, Z. Y., Chen, W. Q., Glazebrook, J., Chang, H. S., Han, B., Zhu, T., Zou, G. Z., and Katagiri, F. 2003. Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. Plant Cell 15:317-330.

Thompson, G. A., and Goggin, F. L. 2006. Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. J Exp Bot 57:755-66.

Turner, J. G., Ellis, C., and Devoto, A. 2002. The jasmonate signal pathway. Plant Cell 14 Suppl:S153-64.

van Kan, J., Cozijnsen, T., Danhash, N., and de Wit, P. 1995. Induction of tomato stress protein mRNAs by ethephon, 2,6-dichloroisonicotinic acid and salicylate. Plant Mol. Biol. 27:1205-1213.

Williamson, V. M., Ho, J.-Y., Wu, F. F., Miller, N., and Kaloshian, I. 1994. A PCR-based marker tightly linked to the nematode resistance gene, *Mi*, in tomato. Theor. Appl. Genet. 87:757-763.

Xie, D. X., Feys, B. F., James, S., Nieto-Rostro, M., and Turner, J. G. 1998. *COI1*: An *Arabidopsis* gene required for jasmonate-regulated defense and fertility. Science 280:1091-1094.

Zhu-Salzman, K., Salzman, R. A., Ahn, J. E., and Koiwa, H. 2004. Transcriptional regulation of sorghum defense determinants against a phloem-feeding aphid. Plant Physiol. 134:420-431.

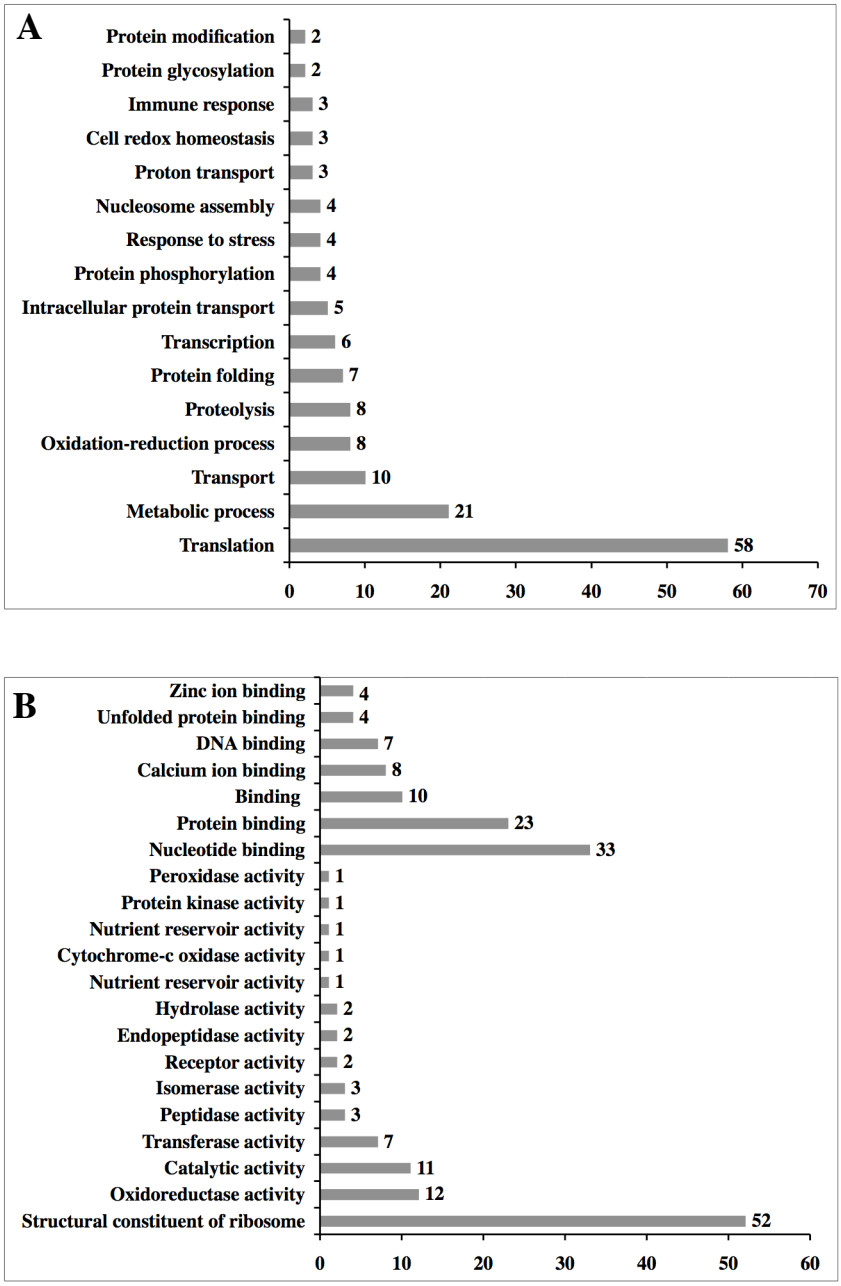


Fig. 4.1A-B Classification of the *Macrosiphum euphorbiae* contigs using the *Acyrthosiuphon pisum* Gene Ontology (GO) terms. The contigs were annotated according to **A,** the biological process that they are predicted or known to be part of or **B,** the known or predicted molecular function.

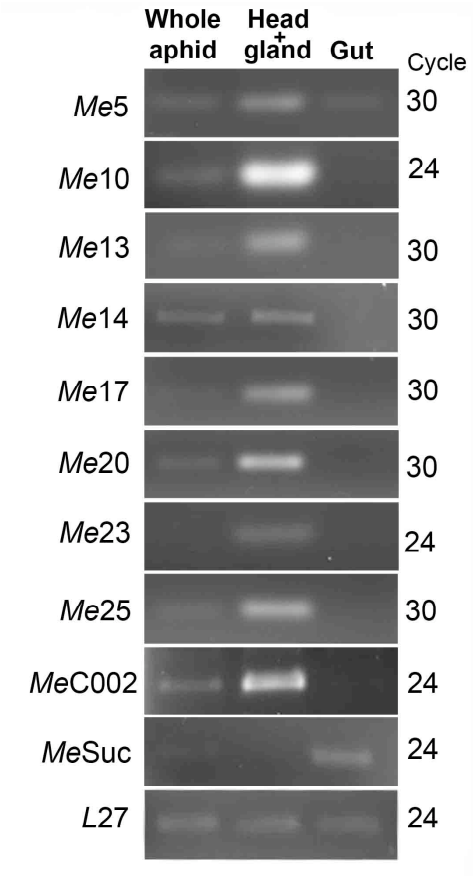


Fig. 4.2 Tissue-specific expression analysis of the eight *Macrosiphum euphorbiae* candidate effectors. RNA isolated from whole aphids or dissected aphid heads and salivary glands or guts was used in reverse transcription and semi-quantitative PCR with gene-specific primers. PCR cycles are indicated to the right of the panel. Expression of *MeC002* and *MeSuc* were used as controls for salivary glands and gut, respectively. Ribosomal gene *L27* was used as an internal control for cDNA.

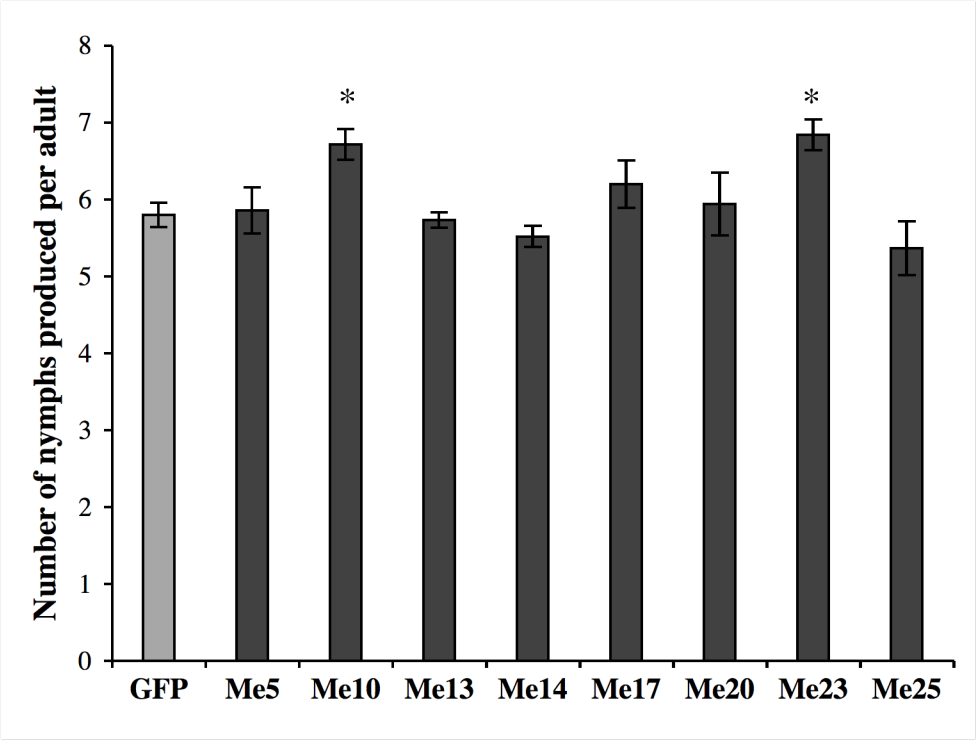


Fig. 4.3 *Myzus persicae* performance on *Nicotiana benthamiana* plants expressing *Macrosiphum euphorbiae* candidate effectors. *M. euphorbiae* candidate effectors were transiently overexpressed in *N. benthamiana* using *Agrobacterium tumefaciens*. One day after agroinfiltration, each leaf sample was caged with four adult *M. persicae*. The following day, adults were removed leaving four first-instar nymphs. The four nymphs were moved to a freshly agroinfiltrated leaf expressing the candidate effector on days 6, 9 and 12. Nymph production was evaluated up to day 15. GFP was used as expression and aphid assay control. Graphs show the average number of nymphs produced per adult from one experiment. Two leaves per plant and 4 plants per construct were used (n=8). Data from one experiment is presented with error bars indicating the standard error. Asterisks indicate statistical significance compared to the GFP control (*Me*10 $P = 0.004$; *Me*23 $P = 0.01$). This experiment was performed three times with essentially identical results.

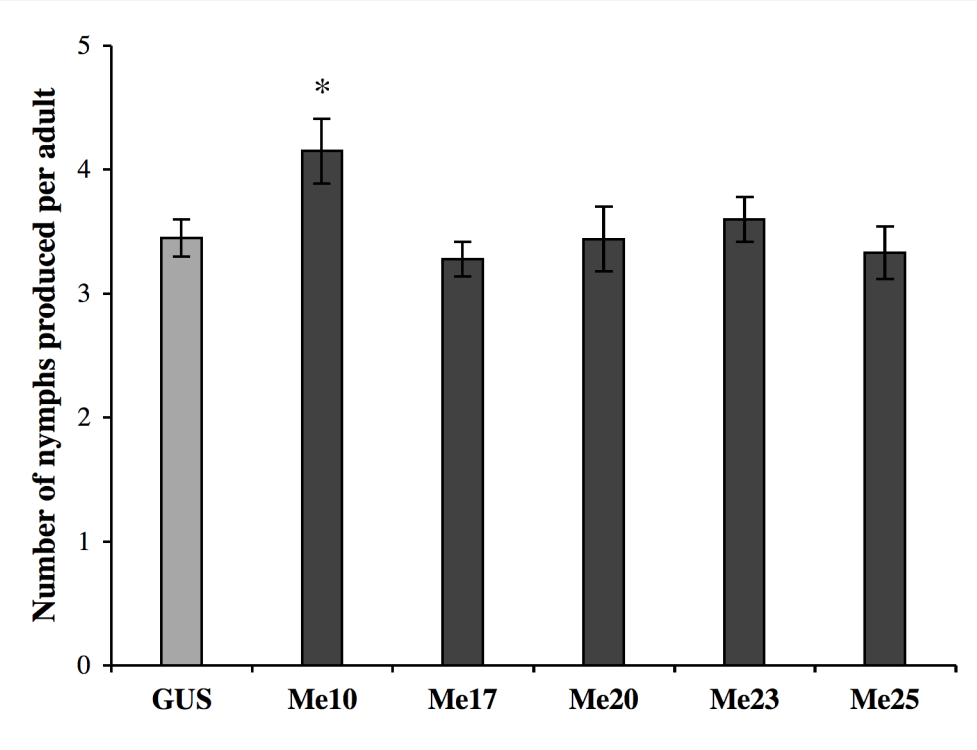


Fig. 4.4 *Macrosiphum euphorbiae* performance on tomato plants expressing *M. euphorbiae* candidate effectors. Five-week-old tomato plants were vacuum infiltrated with recombinant *Pseudomonas syringae* pv. *tomato* DC3000Δ*avrPto*/Δ*avrPtoB* (*Pst*) delivering individual aphid effector or GUS used as control. Five plants per treatment were vacuum infiltrated with recombinant *Pst* 1 x 10^3 CFU/ml and 24 h post infiltration each plant was infested with 9 age-synchronized one-day-old adult potato aphids. Aphid progeny was evaluated for 5 days. Graphs show the average number of nymphs produced per adult aphid (n=5). Data from one experiment is presented with error bars indicating the standard error. Asterisk indicates statistical significance compared to the GUS ($P = 0.021$). This experiment was performed three times with essentially identical results.

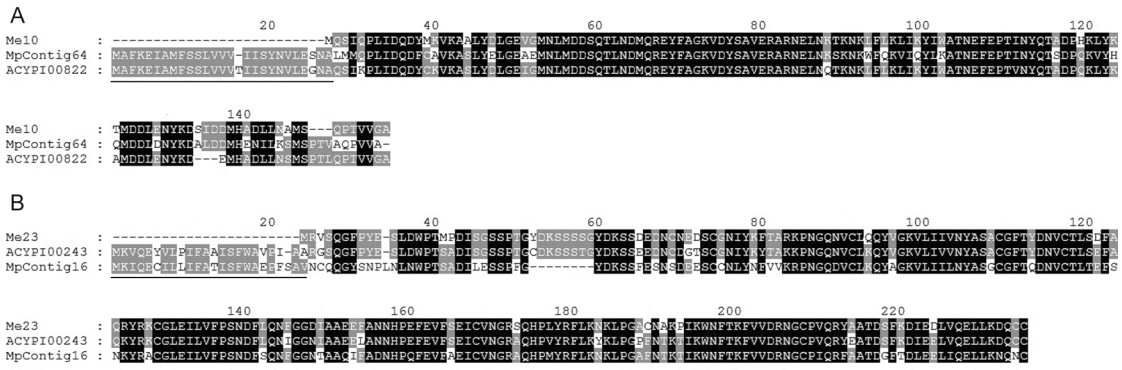


Fig. 4.5A-B Alignment of deduced amino acid sequences of *Macrosiphum euphorbiae* (*Me*) effectors *Me*10 (**A**) and *Me*23 (**B**), with putative orthologs from *Acyrthosiphon pisum* (ACYP) and *Myzus persicae* (*Mp*). Black and grey shades indicate identical and highly conserved amino acids, respectively. Putative secretion signal peptide sequences are underlined.

Table 4.1. Gateway primers for cloning in the expression vectors pEarleyGate100 and pVSP_PsSPdes

Contig	Forwrad primer (5'-3')	Reverse primer (5'-3')
Me5	GGGGACAAGTTTGTACAAAA AAGCAGGCTAACAATGAATA ATCAGGAATTTGTTCATTT	GGGGACCACTTTGTACAAGAAA GCTGGGTTCAGTTTCCGGTGCAT AGGCGAATAAGACG
Me10	GGGGACAAGTTTGTACAAAA AAGCAGGCTAGTGTACTAGA AGGAACAATGCAATCAATA	GGGGACCACTTTGTACAAGAAA GCTGGGTTTATGCTCCAACGACT GTTGGTTGGGAC
Me13	GGGGACAAGTTTGTACAAAA AAGCAGGCTCGGACAACTAG TAACAATGAGTTTGTTTCC	GGGGACCACTTTGTACAAGAAA GCTGGGTTTATCCGTTGTTGGTC GAAGATACATTTTT
Me14	GGGGACAAGTTTGTACAAAA AAGCAGGCTTGTTCCAAACA ATGCAGTTTCTAATAT	GGGGACCACTTTGTACAAGAAA GCTGGGTATATTACATAGGTAGT AATATATAAGCTTA
Me17	GGGGACAAGTTTGTACAAAA AAGCAGGCTACCAGCAGGTG TGCAACAATGGAACCGAC	GGGGACCACTTTGTACAAGAAA GCTGGGTGAGTTGGTGTATAACG TCAACG
Me20	GGGGACAAGTTTGTACAAAA AAGCAGGCTATAACAATGTG GCCGGGATTCACCGTGGTC	GGGGACCACTTTGTACAAGAAA GCTGGGTATCGGTTCGTACTTCC TGGGATTTTGAGTG
Me23	GGGGACAAGTTTGTACAAAA AAGCAGGCTAACAATGAGGG TTTCCCAAGGATTTCCCTA	GGGGACCACTTTGTACAAGAAA GCTGGGTGTTTACGTAACATAAT TAGCAACATTGGTCC
Me25	GGGGACAAGTTTGTACAAAA AAGCAGGCTGCAACAATGTC ACCAAGTCACGGAGAAAACG	GGGGACCACTTTGTACAAGAAA GCTGGGTTCAGAATAGAGCTAG AGTAGCGAGTTTG
GUS	GGGGACAAGTTTGTACAAAA AAGCAGGCTCCATGGTCCGT CCTGTAGAAAC	CACCACTTTGTACAAGAAAGCTG GGTCTTATTGTTTGCCTCCCTGCT

Table 4.2. Primers used for gene expression analysis.

Contig	Forwrad primer (5'-3')	Reverse primer (5'-3')
Me5	CATTCAAGGCCTAGACAGGAC	TCTTCTCAAGTCGTTTGTCCC
Me10	CAAGACCAAAAATAAGTTGTTCCT	GACATTGCGTTAAGAAGATCG
Me13	GAAGTCAATTGGCCATTAAACAA	CCATTTCTGCATTATTTCTGG
Me14	GTCATTCGAGAGCGTACCAT	CAAAAGGACTCGATCGGTTG
Me17	GTTTGACGAGATGGACTTGCTTA	GTATCTCAACCTTCTTTTCAGCC
Me20	GAACCGAAGACGACCTCATTCT	CAAGTCTCCATGGGCATCG
Me23	GGAAATTTGTGTAAATGGCAGA	CATTGGTCCTTTAACAGTTCTTG
Me25	CATGGACAACTAACACACAATTC	GAGCTAGAGTAGCGAGTTTGCAAC
L27	CCGAAAAGCTGTCATAATGAAGAC	GGTGAAACCTTGTCTACTGTTACATCTT
MeC002	GAGCAGGAAGAAGCGTCTGT	CTTGGTGGGAGCATTGGTTA
MeSuc	GAGATCGATCCTATTTATGGC	CATTCCATTCCCACGGAGATC

General Conclusions

Plant cultivation is challenged by diverse biotic agents ranging from insects to viruses, which also depend on plants for their survival. The coexistence of plants with these biotic agents, is the outcome of a complex and continuous change, adaptation, and counter-adaptation among the interacting organisms (Schneider and Collmer 2010). Plant adaptation includes physical barriers, chemical weapons and pattern-triggered immunity (PTI) that limit the growth of pathogen or pest using different mechanisms (Segonzac and Zipfel 2011; Walters 2011; Wittstock and Gershenzon 2002). To adapt to this new conditions, microbial organisms, pests, and infectious agents evolved more complex life cycles, avoidance mechanisms, and counter-chemical substances that affect components of the innate immune system or metabolic pathways of their hosts (Dobler et al. 2011; Walling 2008). Some of these substances, called effector molecules, suppress PTI resulting in effector-triggered susceptibility (ETS) (Jones and Dangl 2006). Plants in turn have evolved resistance proteins to counteract the ETS resulting in effector-triggered immunity (ETI) (Eitas and Dangl 2010). Several biotic agents have evolved mechanisms to evade or suppress the ETI continuing the evolutionary arms race for survival. Plants being essential for human survival, humans side with plants in this arms race, applying various strategies to improve the chance of crops to gain the upper hand.

One environmentally safe approach for crop improvement involves exploiting available natural resistance mechanisms in wild relatives of respective cultivated plant species and introducing them into crops. The *Mi-1.2* gene has been introgressed from

the wild tomato species *Solanum peruvianum* into cultivated tomato (*Solanum lycopresicum*) (Smith 1944). This gene confers resistance to potato aphid (*Macrosiphum euphorbiae*), root-knot nematode (RKN) (*Meloidogyne* sp.), whitefly (*Bemisia tabaci*), and tomato psyllid (*Bactericera cockerelli*) (Casteel et al. 2006; Dropkin 1969; Nombela et al. 2003; Roberts and Thomason 1986; Rossi et al. 1998). All these pests have broad host ranges and host resistance genes have not been identified against these pests. It would be desirable to transfer the *Mi-1.2*-mediated resistance into plant species that are attacked by these pests.

Some *R*-genes, like the tobacco *N*-gene conferring resistance against tobacco mosaic virus (TMV) and the pepper *Bs2*-gene conferring resistance against *Xanthomonas campestris* pv. *vesicatoria* (Xcv), are functional against TMV and Xcv when introduced into tomato (Tai et al. 1999; Whitham et al. 1996). However, introduction of *Mi-1.2* into several solanaceous and non-solanaceous crops were unsuccessful. Introduction only into eggplant conferred resistance against RKN but not potato aphid (Goggin et al. 2006). Thus, successful incorporation of *Mi-1.2*-mediated resistance into other plant species requires thorough understanding of the resistance-signaling pathway.

Towards this goal, in Chapter one of this dissertation, a role was identified for *SlWRKY70* transcription factor (TF) in *Mi-1.2*-mediated resistance against potato aphids and RKN. Using virus-induced gene silencing (VIGS), the transcript level of the *SlWRKY70* gene was specifically knocked-down. These tomato plants had compromised *Mi-1.2*-mediated resistance against potato aphid and RKN. Moreover

gene expression analysis showed that certain aspects of the regulation of this TF are conserved between tomato and *Arabidopsis thaliana* (Arabidopsis). This suggests that through identification and incorporation of the missing components of the tomato *Mi-1.2*-signaling pathway into Arabidopsis, will likely make *Mi-1*-mediated defense functional in Arabidopsis. Arabidopsis with its outstanding resources for functional genomics made possible identification of the functions of a large number of genes providing scientists with much deeper understanding of plant development, immunity, and environmental responses. Studying *Mi-1.2*-mediated resistance in Arabidopsis will speed up the process of comprehensive characterization of this resistance, providing valuable new strategies to combat these economically important pests.

The *WRKY* family transcription factors are found throughout the green lineage, green algae to terrestrial plants. During their long evolutionary history, the WRKY family greatly expanded, as demonstrated by the increased number of *WRKY* genes in higher plants (Rushton et al. 2010; Zhang and Wang 2005). The availability of plant whole genome sequences resulted in the identification of WRKY TF from several species. However the orthologous relationship of the WRKY TF members among the different plant species has not been established. To address this issue, Chapter two of this dissertation reports a comprehensive phylogenetic analysis of plant WRKY TFs. This analysis established orthologous relationships among the WRKY TFs in diverse plant species. One of the accomplishments of this study was the development of a modified analysis pipeline that made use of conserved motif (CM) search followed by excluding of the very diverse sequences for efficient alignment. Previous efforts were

unsuccessful in assessing orthology relationships among the WRKY sequences from multiple plant species because they relied on direct alignment of full-length group specific sequences to construct phylogenetic trees. The approach employed in Chapter two used initial detailed CM analysis prior to sequence alignment. This enabled the identification of WRKY sequences that were considerably diverged from the majority of sequences within a specific group. Current alignment programs are trained to align all provided sequences to the best of their ability and cannot discard sequences that disallow proper alignments. Through exclusion of highly divergent sequences prior to aligning sequences, helped the alignment program to output much better alignments which consequently allowed for better-resolved trees and proper inferences of orthology among each of the group phylogenies. This analysis was performed using sequences from 15 different plant species. Finally, defining signatures were identified among each orthologous members. The significance of this work is the ability to establish putative orthologus relationships of WRKY TFs from newly sequenced plant genomes with those already available, by referring to the signature motifs provided in this study. Moreover, this analysis allowed the design of a systemic nomenclature for the WRKY TF family to include the inferred orthology relationships. The proposed nomenclature provides an example and might encourage a community effort aiming at developing specific guidelines for the annotation of both present WRKY TFs and those that will be identified in the future.

Aphids are economically important pests that cause damage to wide range of plant species. The unconventional aphid biology that includes complex life cycle,

phenotypic plasticity, and “telescoping of generations”, which result in build up of immense populations very quickly are some of the aphid adaptations that have contributed to their success as pests. Moreover aphids have complex interactions with their hosts. These interactions are thought to be analogous to the well-characterized plant-microbial pathogen interactions as aphid feeding induces similar host defense responses including callose deposition and defense gene induction. Unlike chewing insects that cause extensive tissue damage, aphids penetrate the host with their modified mouthparts, called stylets, moving intercellularly and cause minimal tissue damage. During feeding, aphids secrete two types of saliva, gelling and liquid saliva. There is evidence that aphid feeding can manipulate plants to their advantage by converting sink tissue to source tissue, improving nutrient composition of the phloem sap to enhance their growth, and possibly suppressing plant immunity (Giordanengo et al. 2010). However, some aphid salivary components are recognized by the plant defense surveillance system and act as elicitors of immune responses (De Vos et al. 2009). Despite their importance, the genome sequence of only one aphid species, *Acyrthosiphon pisum*, is publicly available (International Aphid Genomic Consortium, 2010).

To better understand aphid biology, the transcriptome of the potato aphid was sequenced and annotated in Chapter three. Several RNA-Seq libraries were prepared from different aphid developmental stages and aphids exposed to biotic (*Mi-1.2* resistance) and abiotic (starvation) stresses to enrich the diversity of expressed genes. The generated transcriptome provides a platform for functional genomic research in

this aphid species to better understand mechanisms of aphid adaptation and plant's counter adaptation. The availability of a potato aphid transcriptome will make possible identification of potato aphid genes differentially regulated after feeding on resistant (*Mi*/*Mi*) compared to susceptible (*mi*/*mi*) tomato plants using RNA-Seq technologies and sequencing short-reads. Further functional characterization of these genes will identify the mode of action or target(s) of the *Mi-1.2*-mediated resistance in potato aphid. These might represent novel targets that could be manipulated for controlling a wide range of insect pests.

Moreover, the potato aphid transcriptome added to the available transcriptome and EST sequences from other aphid species and allowed the performance of comparative sequence analysis among insects. Through this analysis, putative aphid-specific clusters were identified which might contribute to an aphid's adaptation to their host and environment. Among the identified clusters could be genes indispensable for aphid's ability to reproduce parthenogenetically or genes central for perceiving environmental cues and development of winged forms. Targeting such genes represents powerful measures to counteract this devastating pest by reducing the aphid population buildup or inhibiting the ability of aphids to quickly disperse to neighboring fields. Characterization of one aphid-specific gene (Me29764) with an unknown function was pursued. The gene structure is not yet complete but approximately 15 kb was sequenced. Expression analysis showed it is preferentially expressed in the gut. Functional characterization of this gene might identify roles in

detoxification of plant secondary compounds, breakdown of ingested phloem sap, or protection against parasites.

For more detailed understanding of the aphid interactions with its host, and identification of aphid effector proteins that may contribute to aphid colonization of a host plants, the potato aphid salivary gland transcriptome was sequenced and results reported in Chapter four. A total of 200 aphid salivary glands were dissected and used to prepare an RNA-Seq library and sequenced using Illumina technology. By utilizing the *A. pisum* genome, the putative full-length *A. pisum* orthologs of the *M. euphorbiae* salivary gland transcripts were identified. Using bioinformatics, secretion signal peptide was identified among a subset of these genes. Identification of the complete repertoire of effector proteins possessed by aphids represents a first step toward understanding how aphids manipulate plant cellular functions during infestation. For eight putative secreted proteins, the full-length cDNAs were cloned and sequenced from potato aphid. *In planta* functional characterization of these eight putative potato aphid secreted salivary gland proteins was conducted using transient expression in *N. benthamina* and delivery in tomato using Pseudomonas type-three secretion system. Roles for two proteins were identified in manipulating the host plant to their advantage. Aphids had significantly higher fecundity on plants expressing Me23 or Me10. Me23, encoding glutathione peroxidase, may suppress the plant oxidative burst induced by aphid feeding. No hypothesis can be drawn regarding the possible role of Me10 since it encodes uncharacterized protein. Further analysis will identify the mechanisms of action of these proteins.

References

Casteel, C. L., Walling, L. L., and Paine, T. D. 2006. Behavior and biology of the tomato psyllid, *Bactericerca cockerelli*, in response to the *Mi-1.2* gene. Entomol. Exp. Appl. 121:67-72.

De Vos, M., and Jander, G. 2009. *Myzus persicae* (green peach aphid) salivary components induce defence responses in *Arabidopsis thaliana*. Plant Cell Environ. 32:1548-1560.

Dobler, S., Petschenka, G., and Pankoke, H. 2011. Coping with toxic plant compounds--the insect's perspective on iridoid glycosides and cardenolides. Phytochemistry 72:1593-1604.

Dropkin, V. H. 1969. Cellular responses of plants to nematode infections. Annu. Rev. Phytopathol. 7:101-122.

International Aphid Genomic Consortium. 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biol. 8:e1000313.

Eitas, T. K., and Dangl, J. L. 2010. NB-LRR proteins: pairs, pieces, perception, partners, and pathways. Curr. Opin. Plant Biol. 13:472-477.

Giordanengo, P., Brunissen, L., Rusterucci, C., Vincent, C., van Bel, A., Dinant, S., Girousse, C., Faucher, M., and Bonnemain, J. L. 2010. Compatible plant-aphid interactions: how aphids manipulate plant responses. C. R. Biol. 333:516-523.

Goggin, F. L., Jia, L., Shah, G., Hebert, S., Williamson, V. M., and Ullman, D. E. 2006. Heterologous expression of the *Mi-1.2* gene from tomato confers resistance against nematodes but not aphids in eggplant. Mol. Plant-Microbe Interact. 19:383-388.

Hogenhout, S. A., and Bos, J. I. 2011. Effector proteins that modulate plant--insect interactions. Curr. Opin. Plant Biol. 14:422-428.

Lindeberg, M., Cunnac, S., and Collmer, A. 2012. *Pseudomonas syringae* type III effector repertoires: last words in endless arguments. Trends Microbiol. 20:199-208.

Nombela, G., Williamson, V., M, and Muñiz, M. 2003. The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. Mol. Plant-Microbe Interact. 16:645-649.

Roberts, P. A., and Thomason, I. J. 1986. Variability in reproduction of isolates of *Meloidogyne incognita* and *M. javanica* on resistant tomato genotypes. Plant disease 70:547-551.

Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E., and Williamson, V. M. 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc. Natl. Acad. Sci. U.S.A. 95:9750-9754.

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. 2010. WRKY transcription factors. Trends Plant Sci. 15:247-258.

Schneider, D. J., and Collmer, A. 2010. Studying plant-pathogen interactions in the genomics era: beyond molecular Koch's postulates to systems biology. Annu. Rev. Phytopathol. 48:457-479.

Segonzac, C., and Zipfel, C. 2011. Activation of plant pattern-recognition receptors by bacteria. Curr. Opin. Microbiol. 14:54-61.

Smith, P. G. 1944. Embryo culture of a tomato species hybrid. Am. Soc. Hortic. Sci. 44:413-416.

Tai, T. H., Dahlbeck, D., Clark, E. T., Gajiwala, P., Pasion, R., Whalen, M. C., Stall, R. E., and Staskawicz, B. J. 1999. Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. Proc. Natl. Acad. Sci. U.S.A. 96:14153-14158.

Walling, L. L. 2008. Avoiding effective defenses: strategies employed by phloem-feeding insects. Plant Physiol. 146:859-866.

Walters, D. 2011. Plant Defense: Warding off attack by pathogens, herbivores and parasitic plants. Blackwell Publishing Ltd. London.

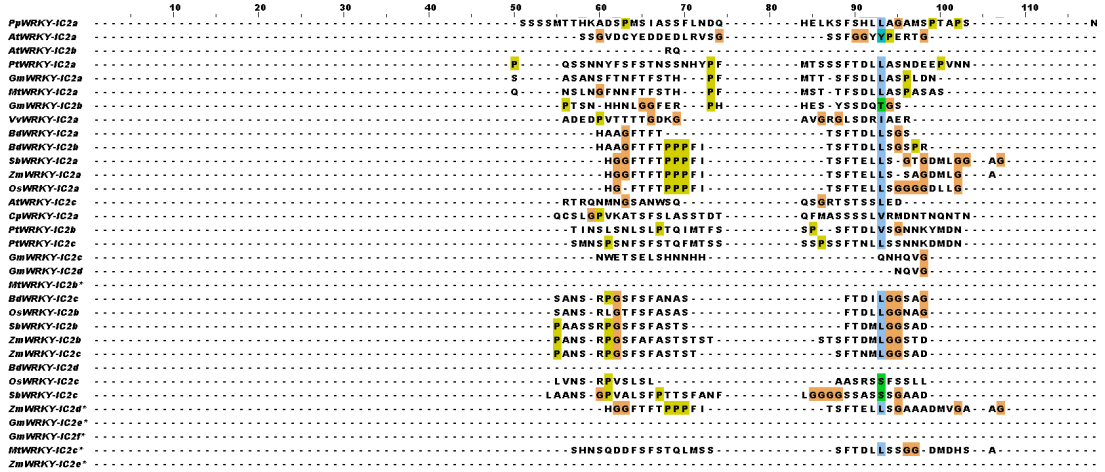
Whitham, S., McCormick, S., and Baker, B. 1996. The *N* gene of tobacco confers resistance to tobacco mosaic virus in transgenic tomato. Proc. Natl. Acad. Sci. U.S.A. 93:8776-8781.

Wittstock, U., and Gershenzon, J. 2002. Constitutive plant toxins and their role in defense against herbivores and pathogens. Curr. Opin. Plant Biol. 5:300-307.

Zhang, Y., and Wang, L. 2005. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol. Biol. 5:1.

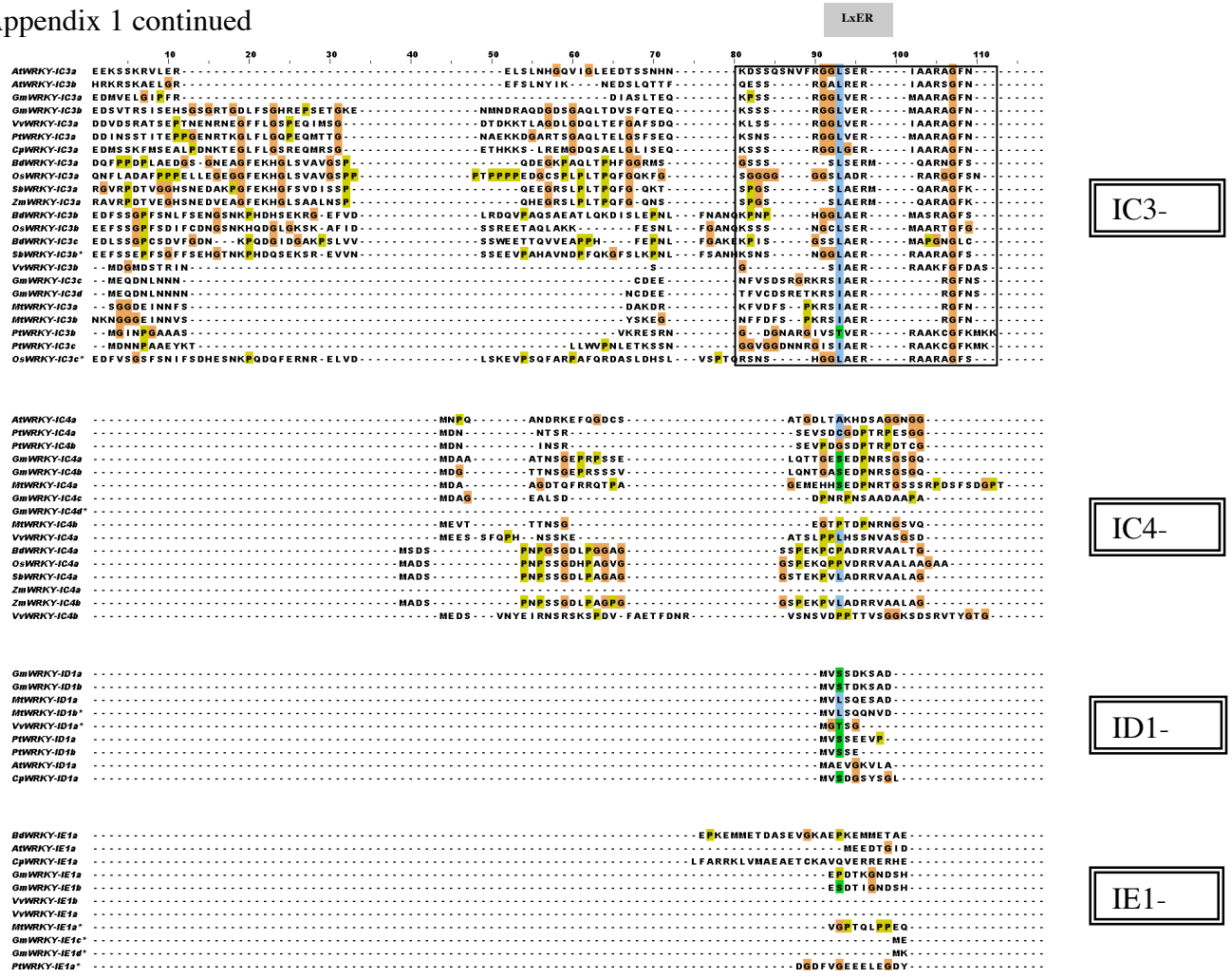
Appendix 1

Appendix 1 continued

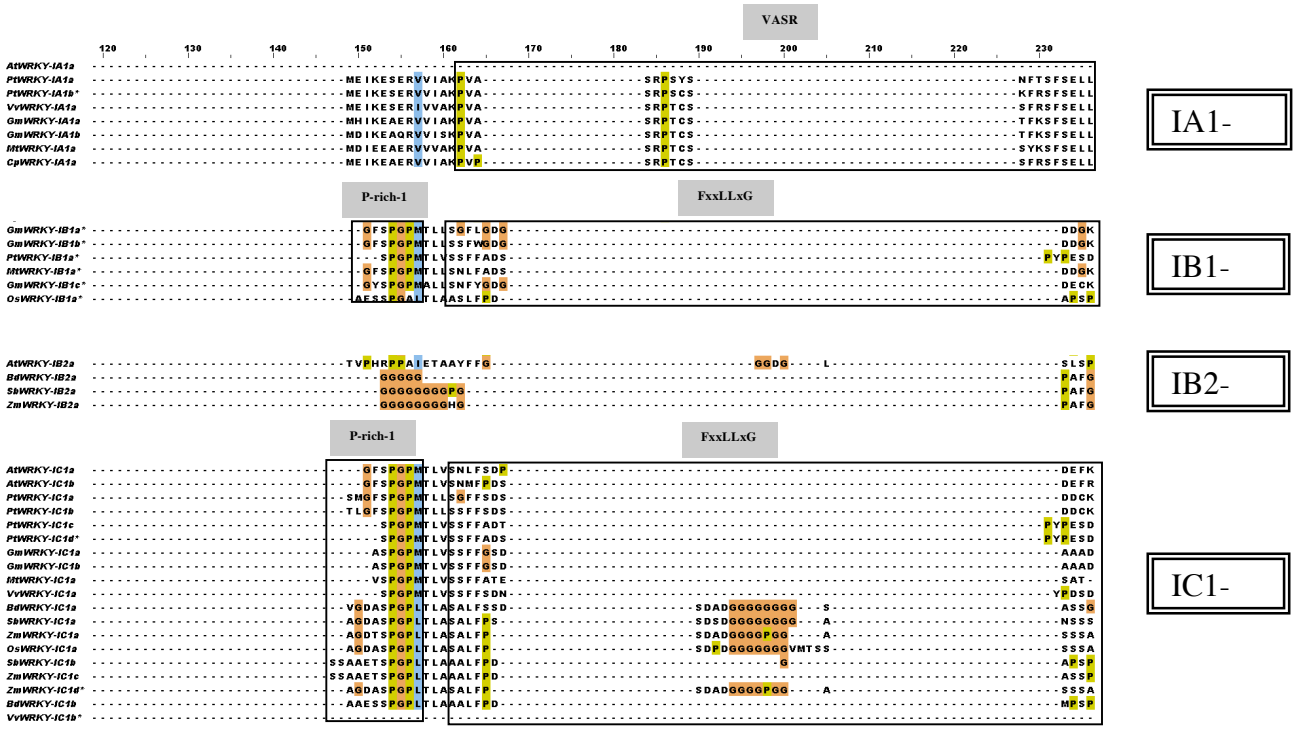


IC2-

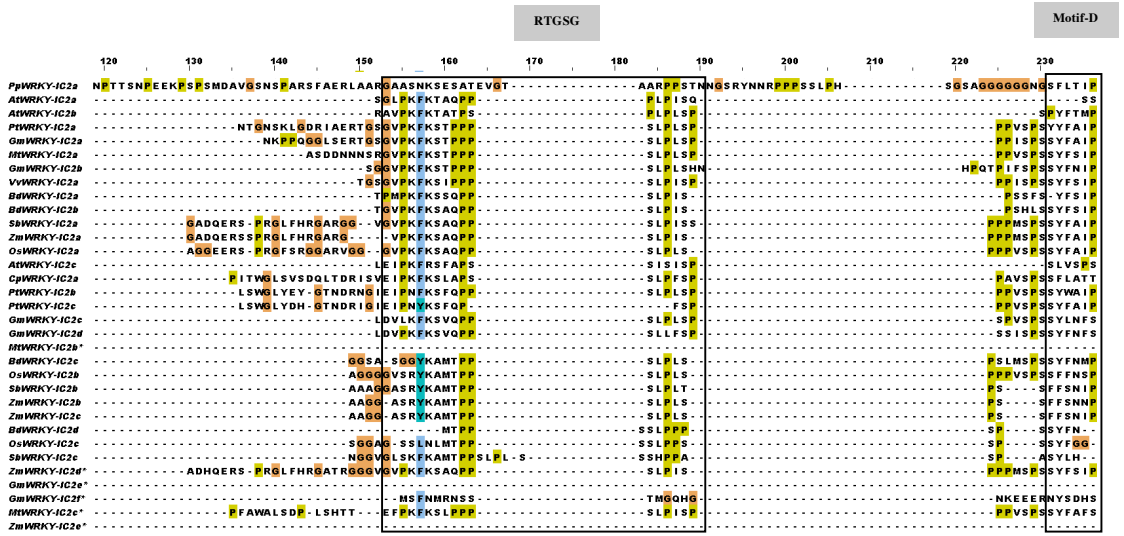
Appendix 1 continued



Appendix 1 continued

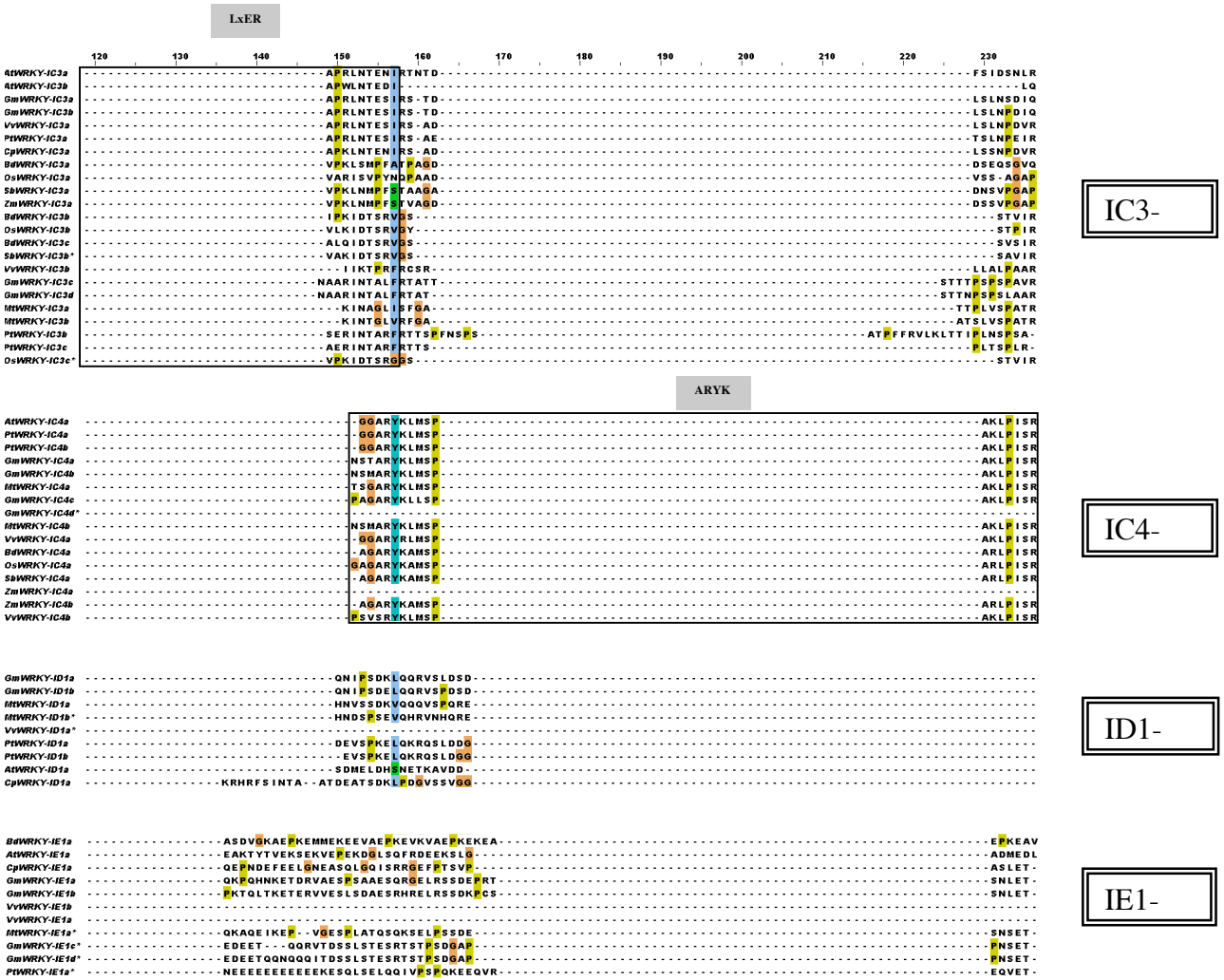


Appendix 1 continued

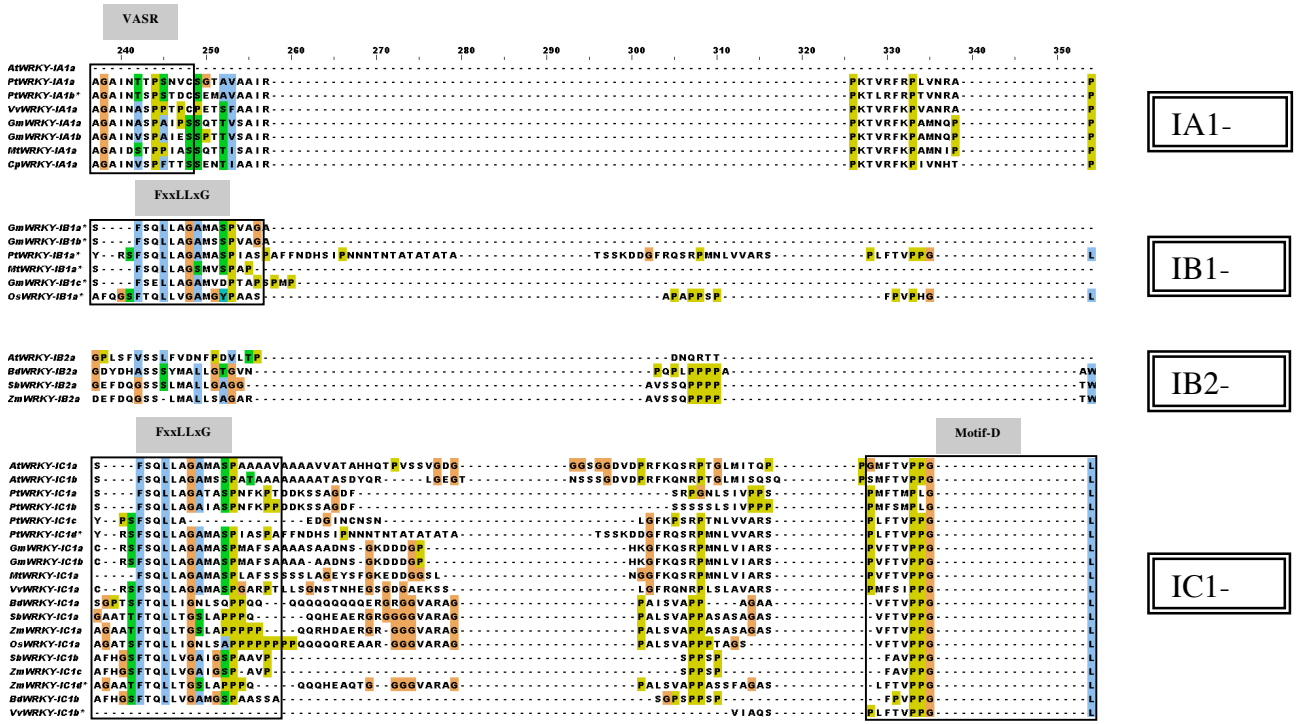


IC2-

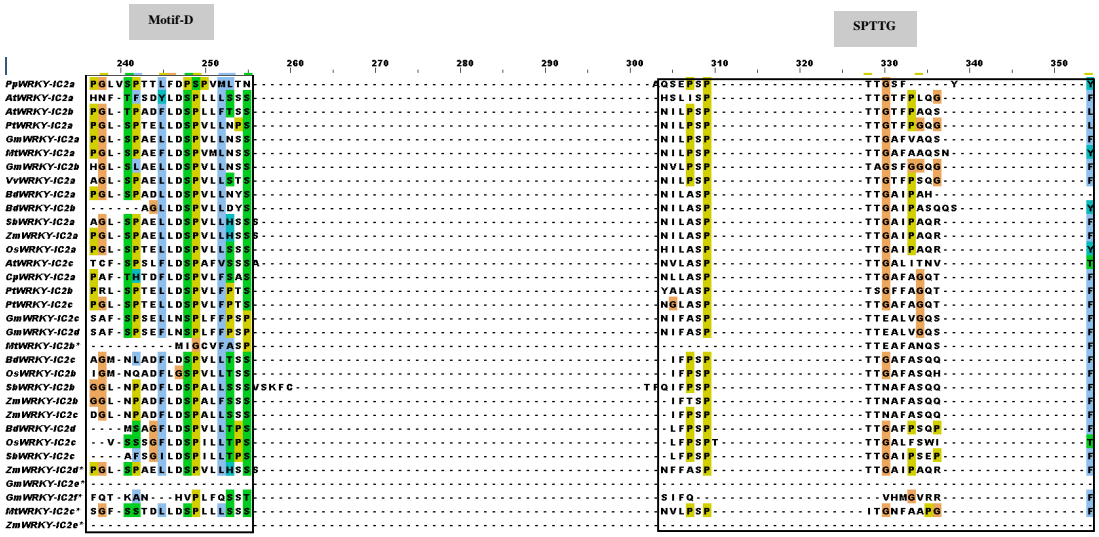
Appendix 1 continued



Appendix 1 continued

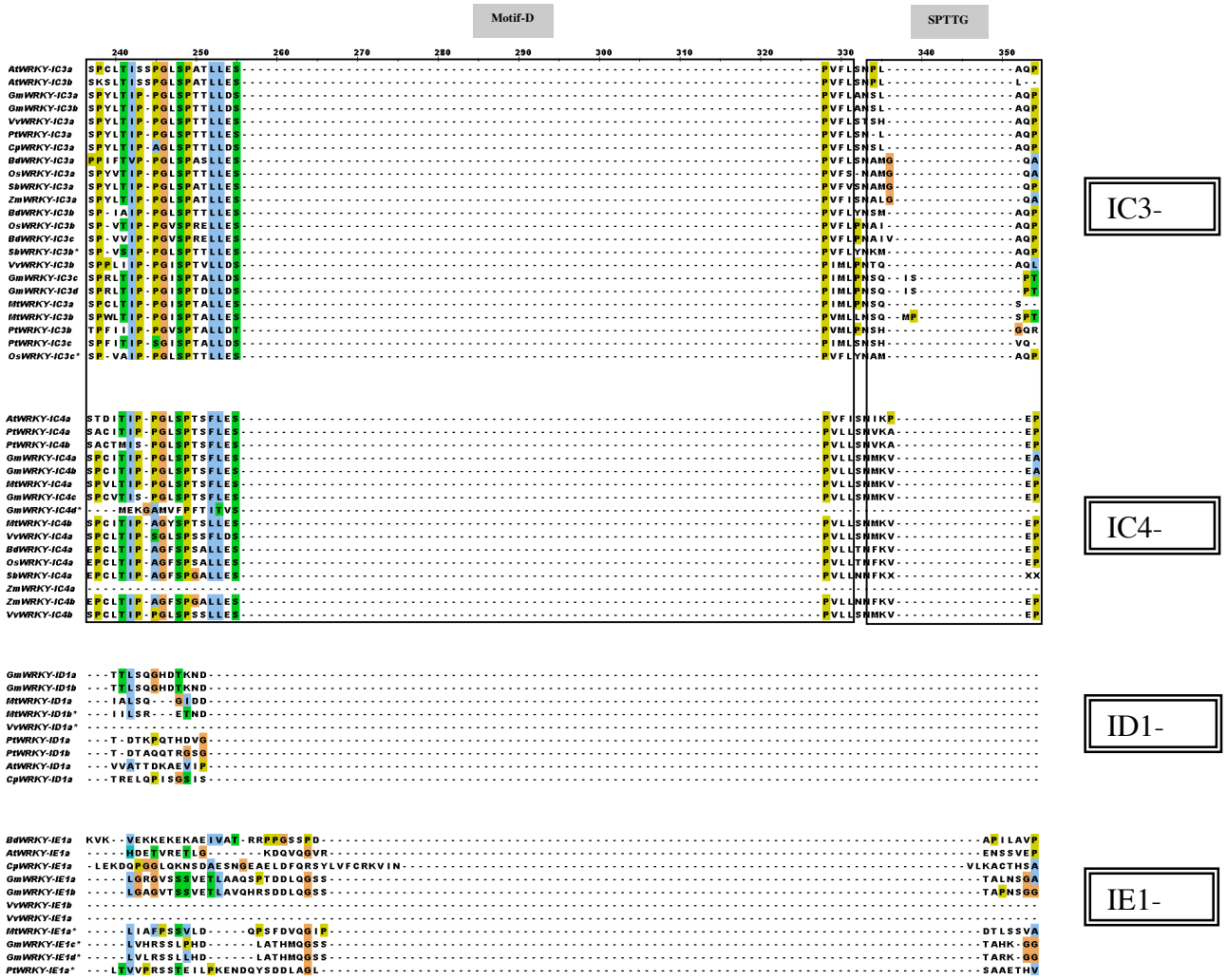


Appendix 1 continued

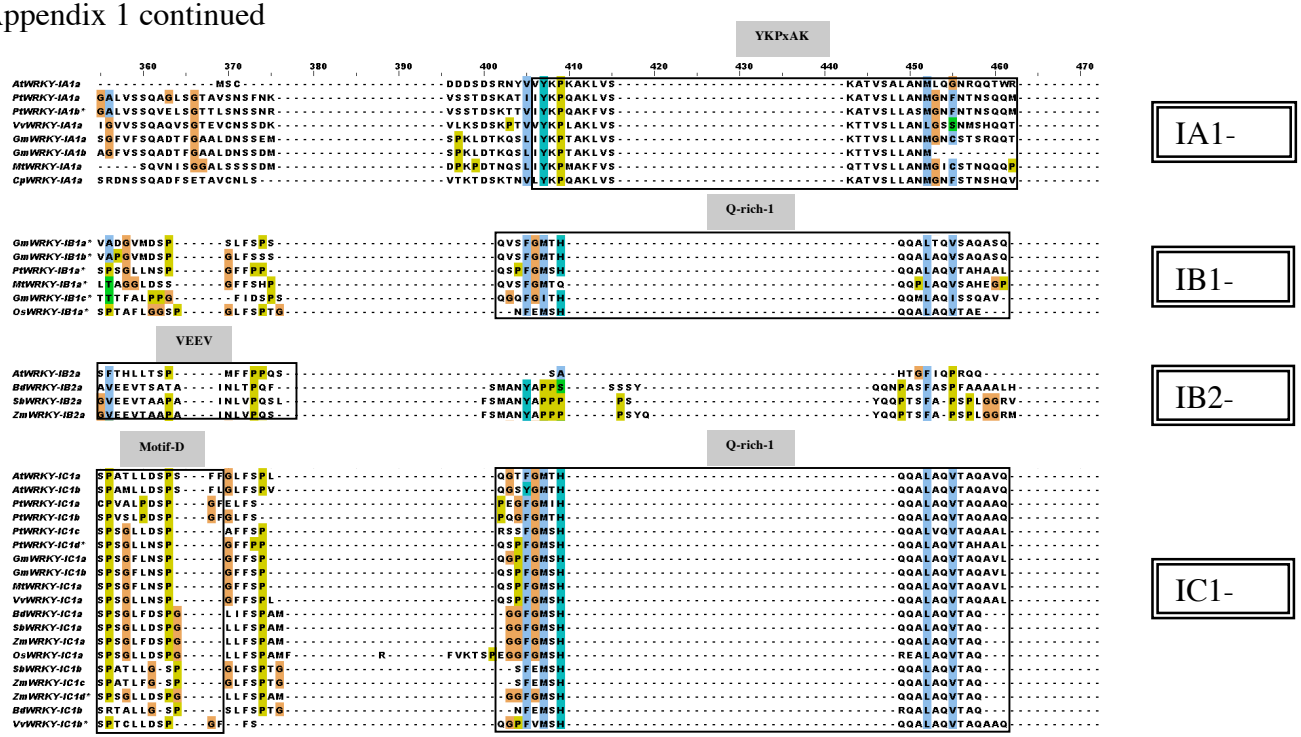


IC2-

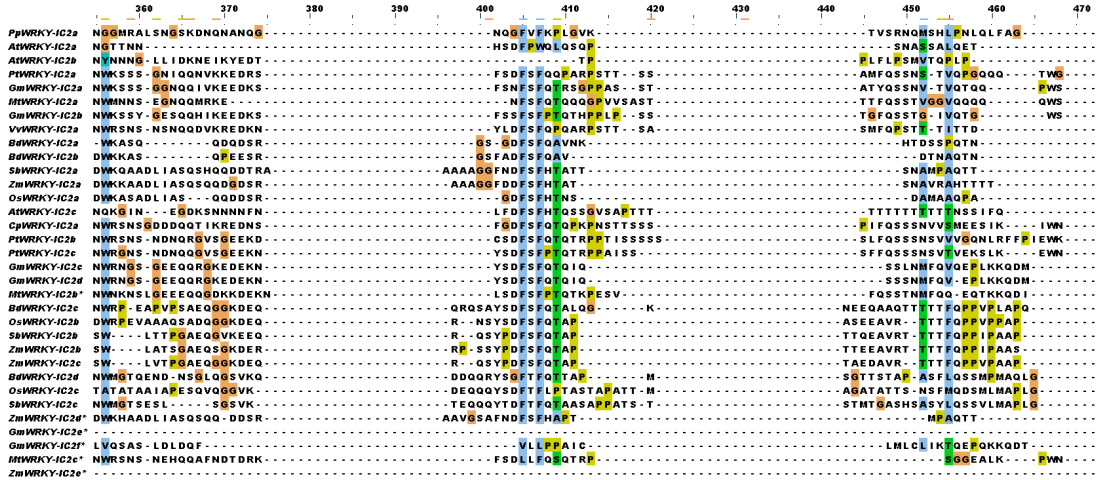
Appendix 1 continued



Appendix 1 continued

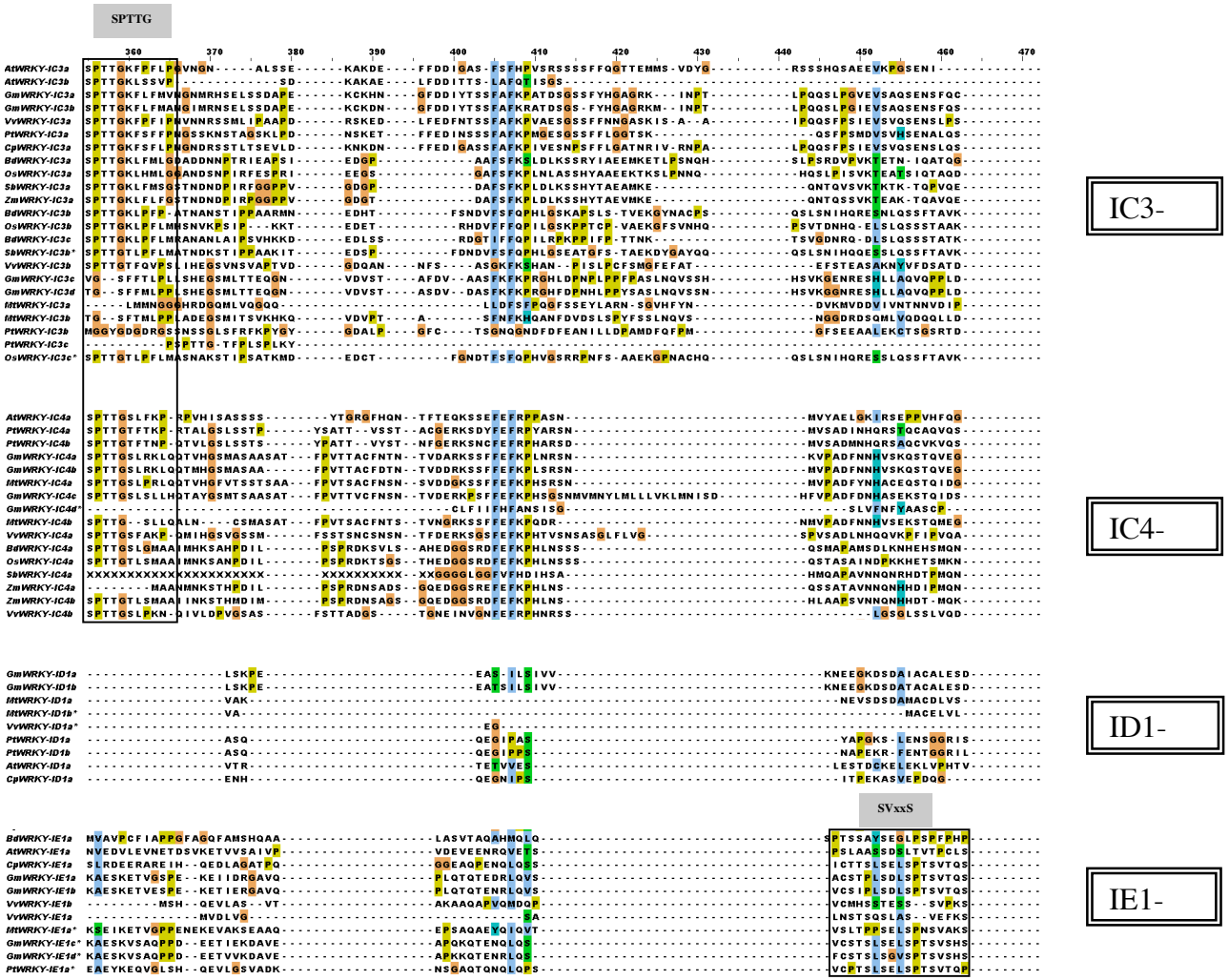


Appendix 1 continued

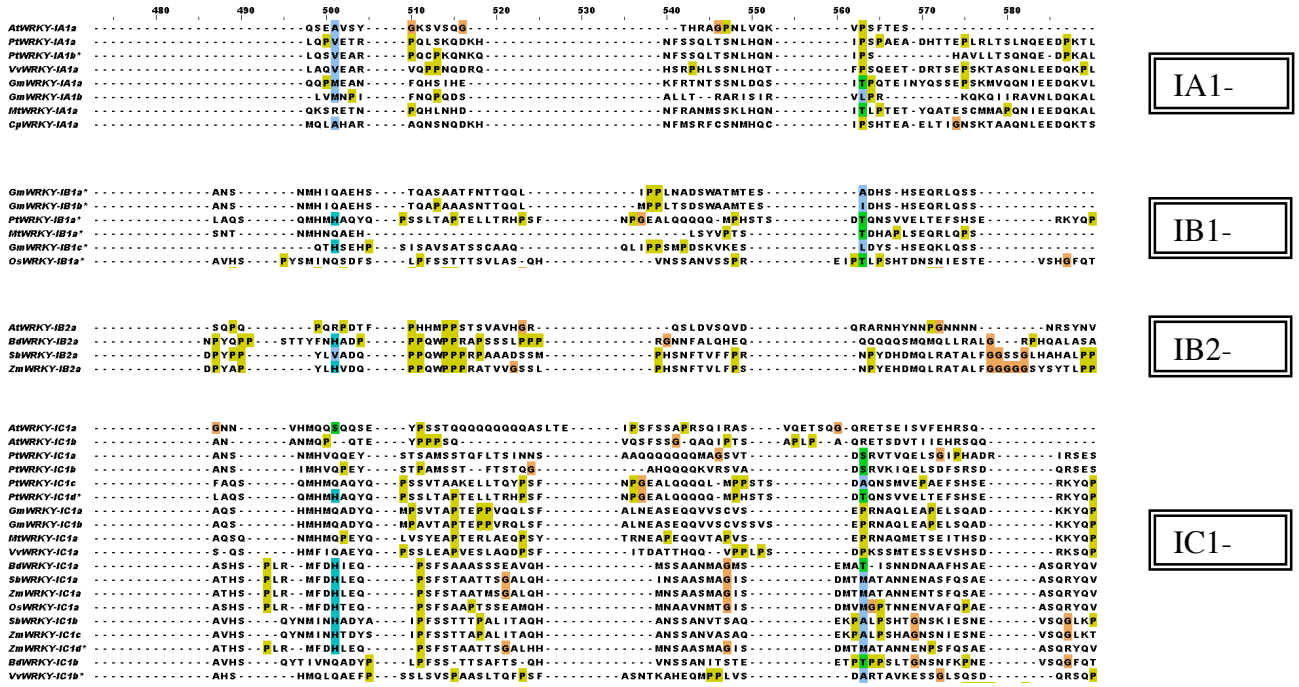


IC2-

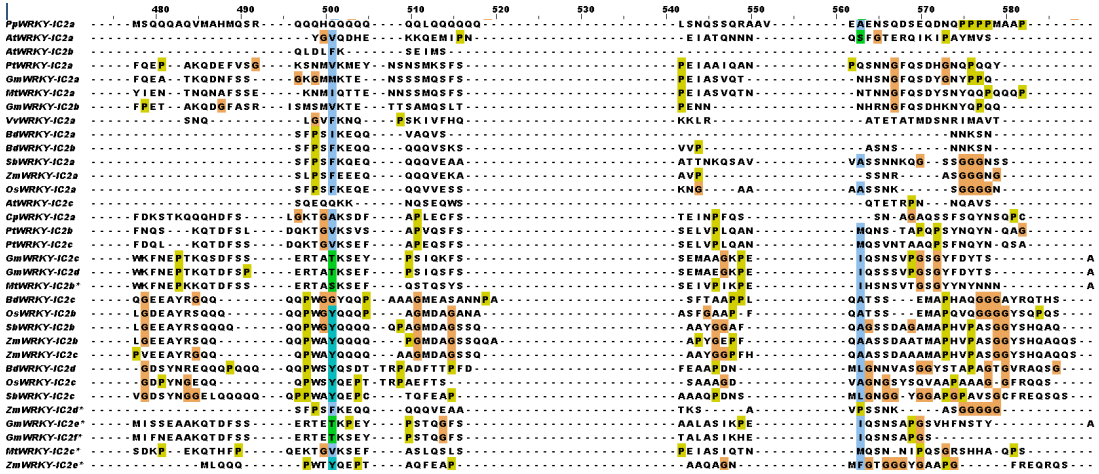
Appendix 1 continued



Appendix 1 continued

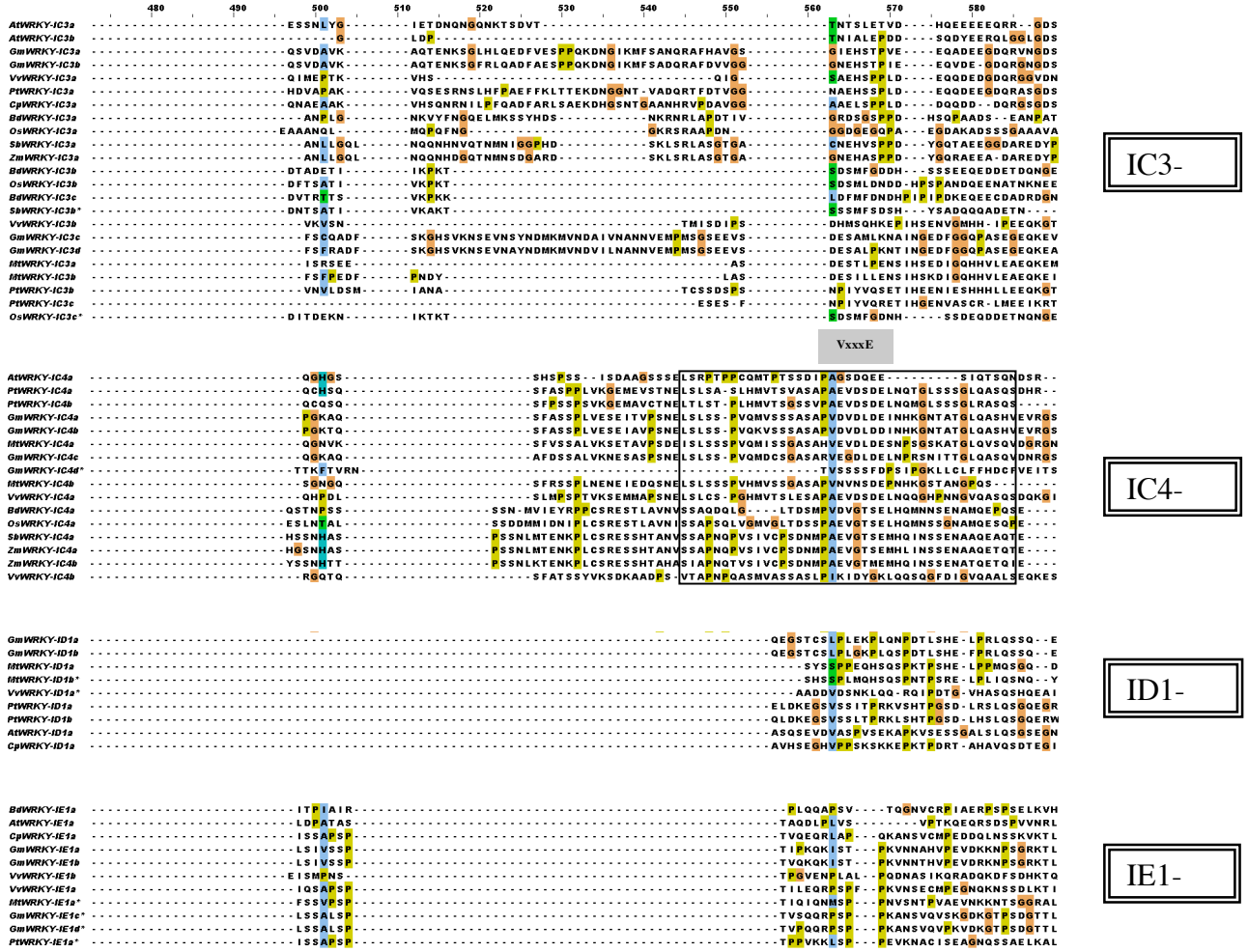


Appendix 1 continued

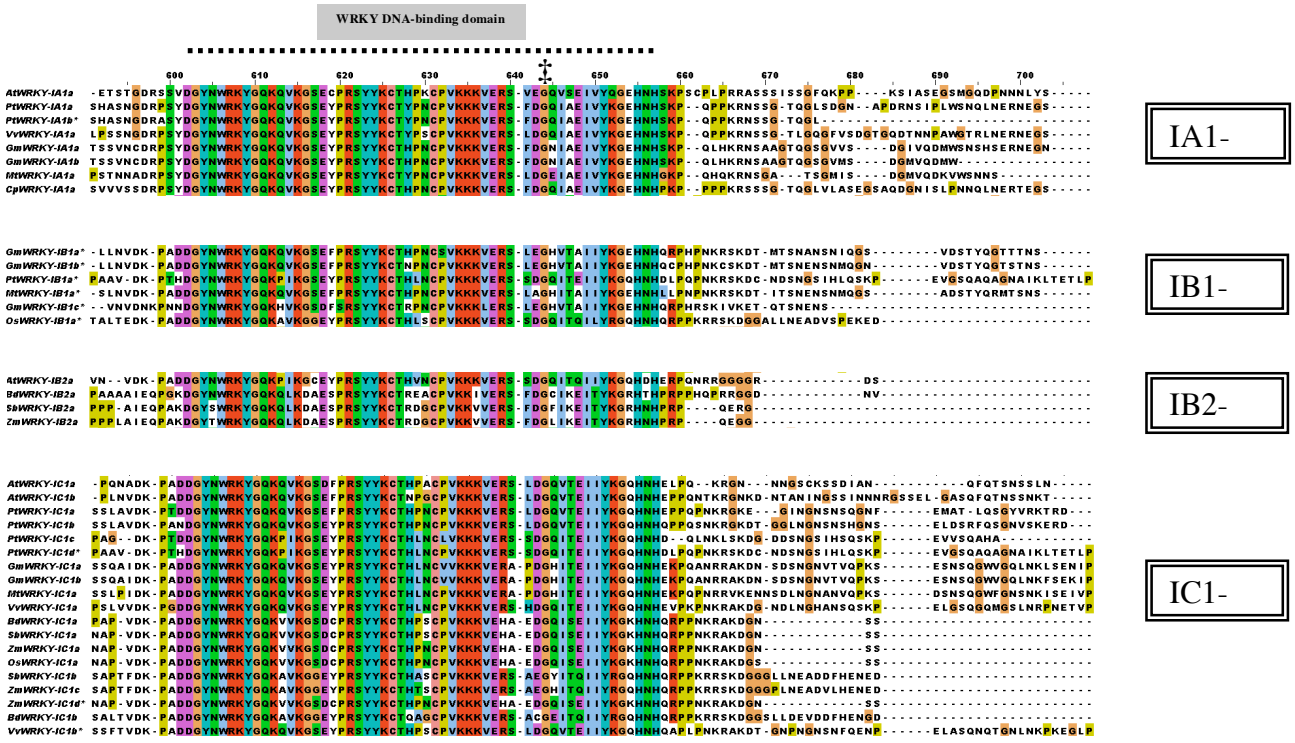


IC2-

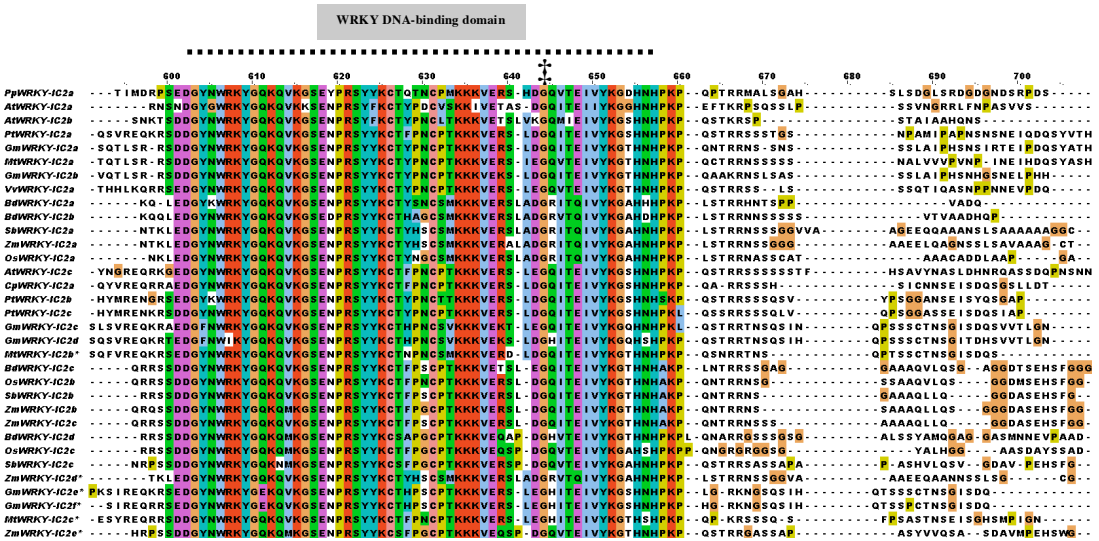
Appendix 1 continued



Appendix 1 continued



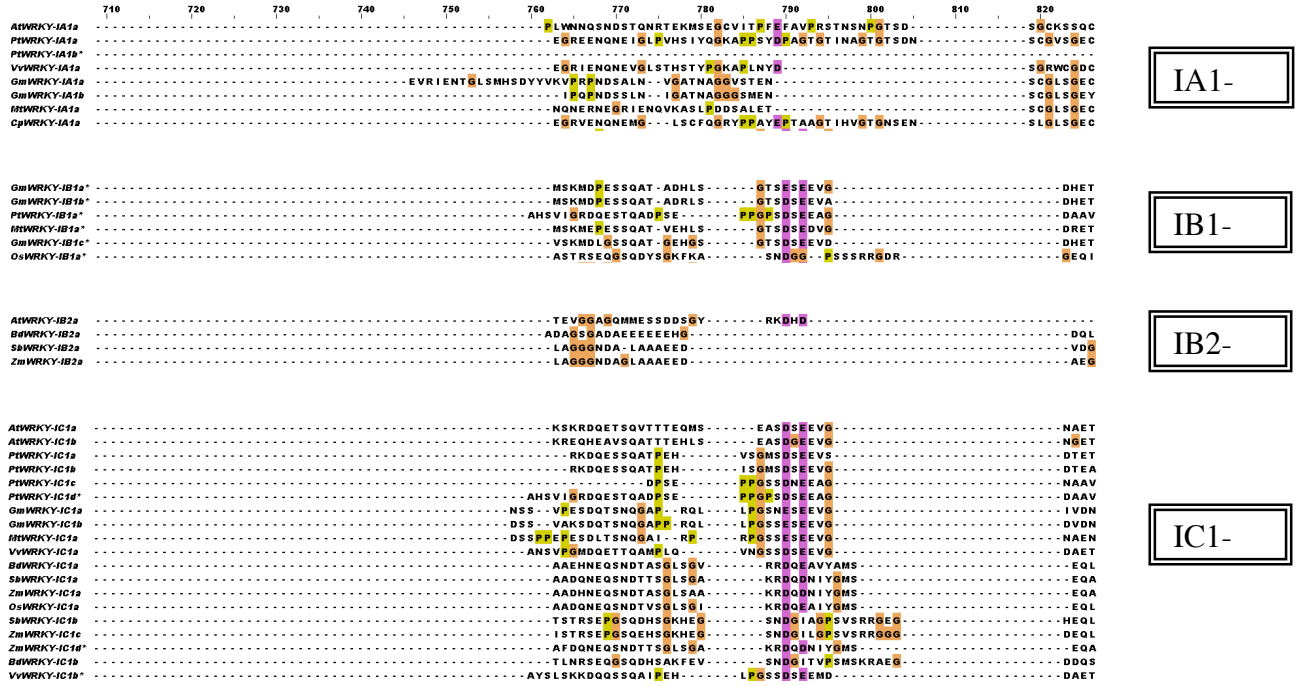
Appendix 1 continued



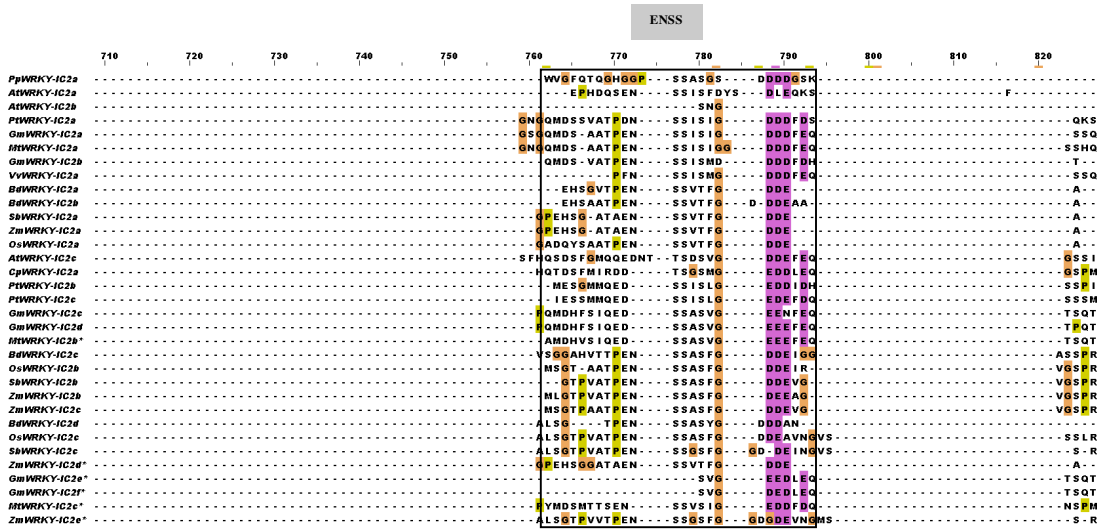
IC2-

Appendix 1 continued

Appendix 1 continued

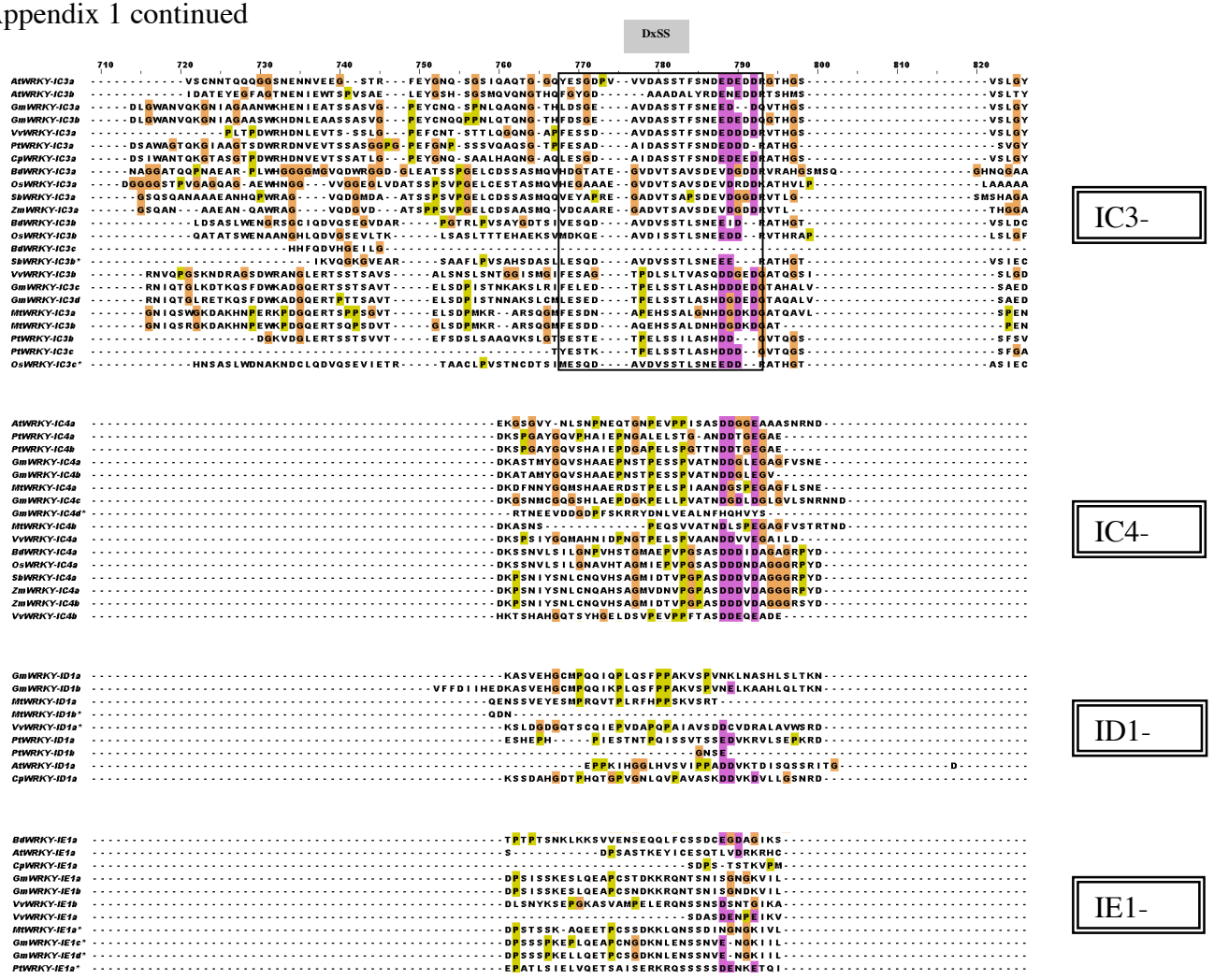


Appendix 1 continued

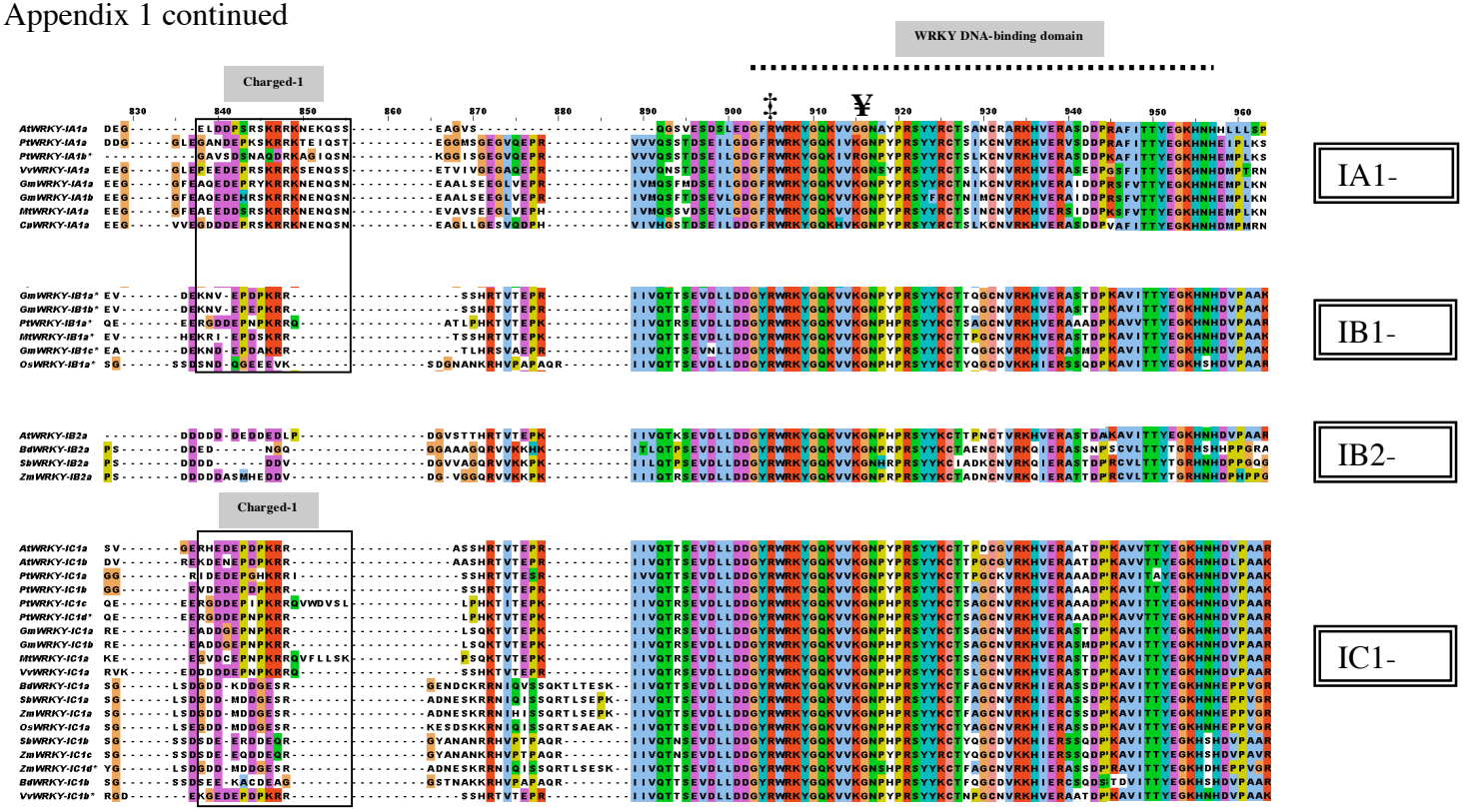


IC2-

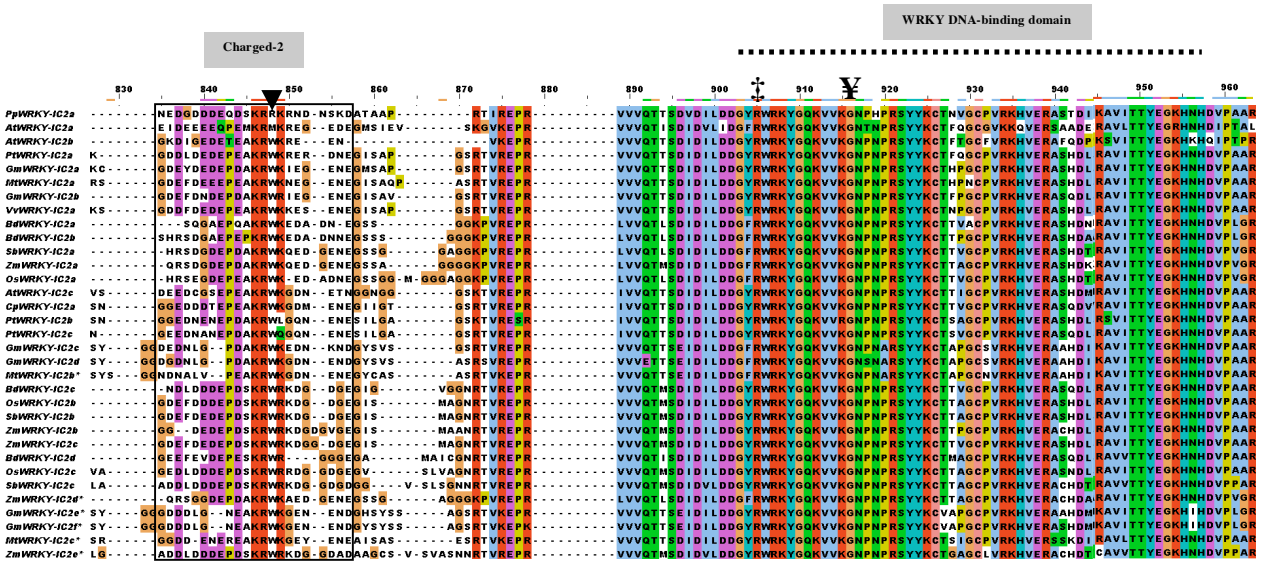
Appendix 1 continued



Appendix 1 continued

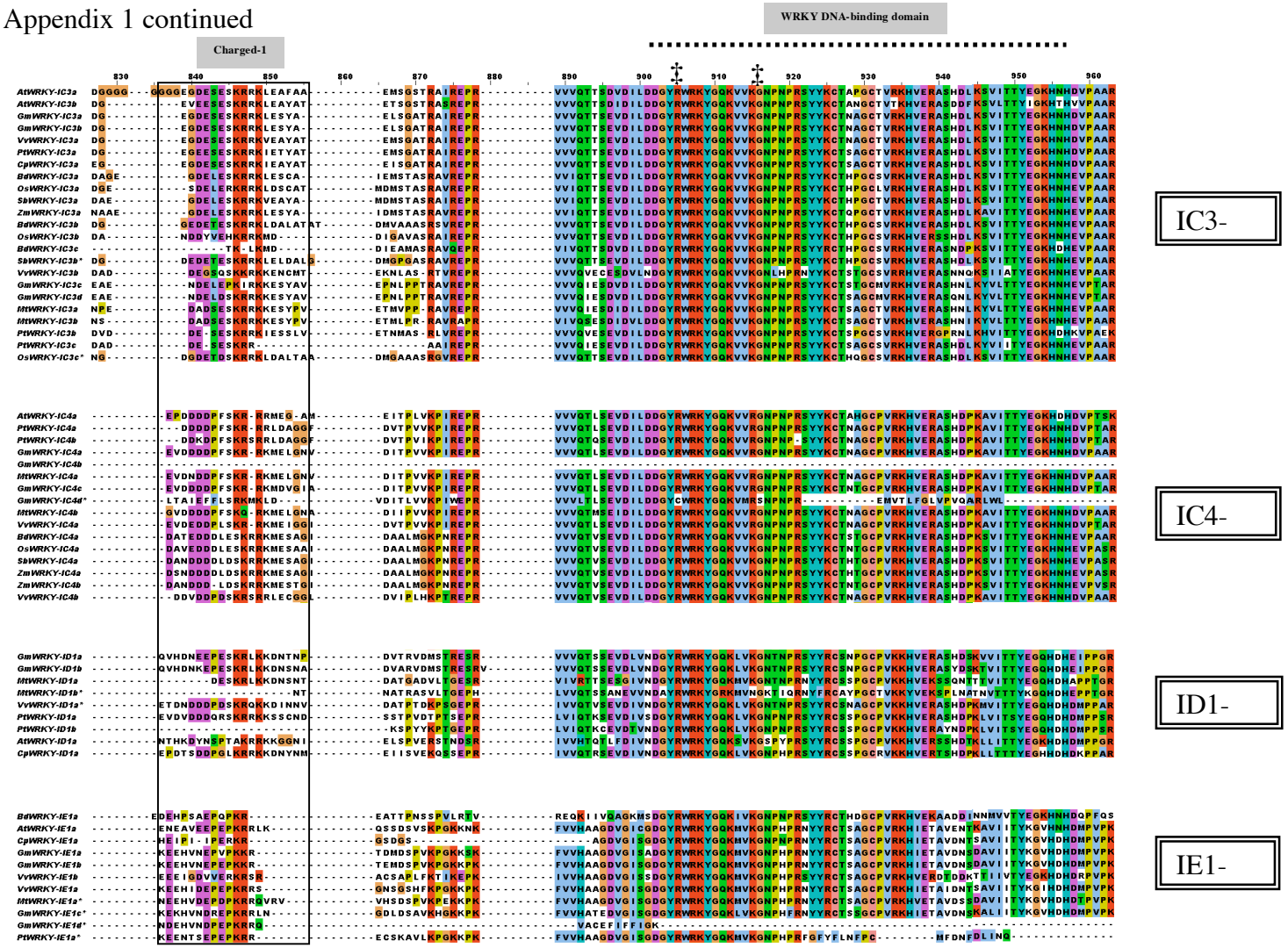


Appendix 1 continued

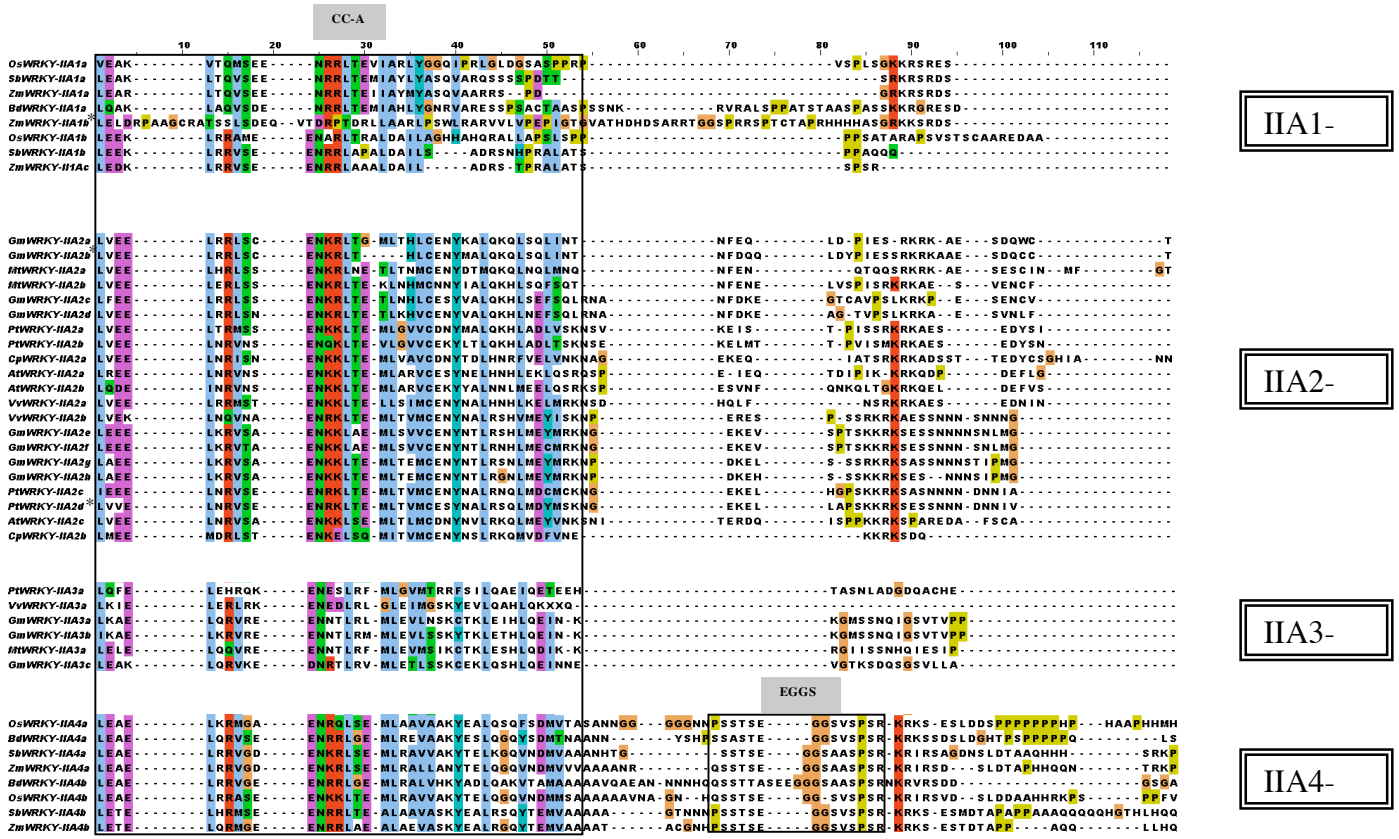


IC2-

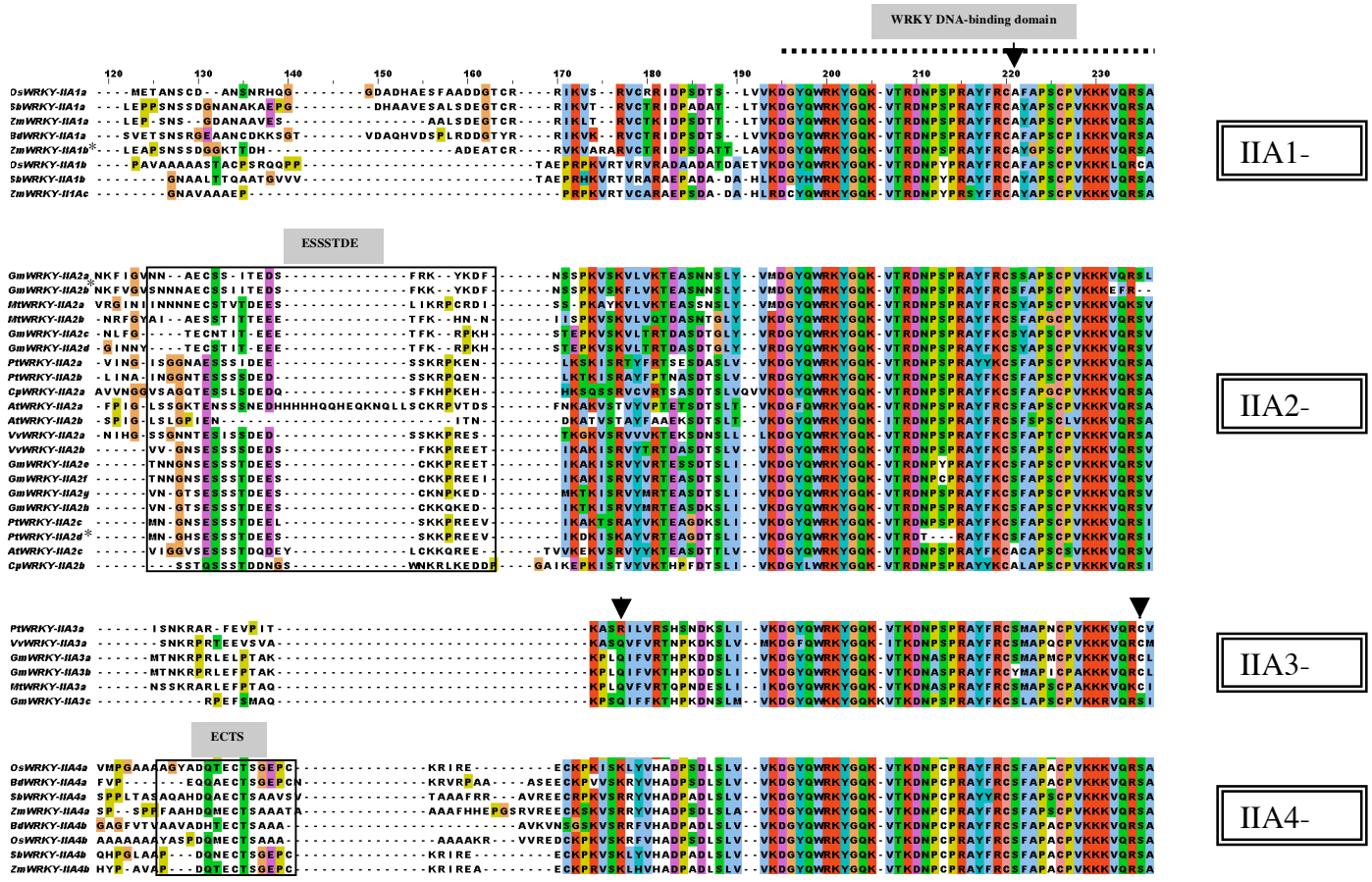
Appendix 1 continued



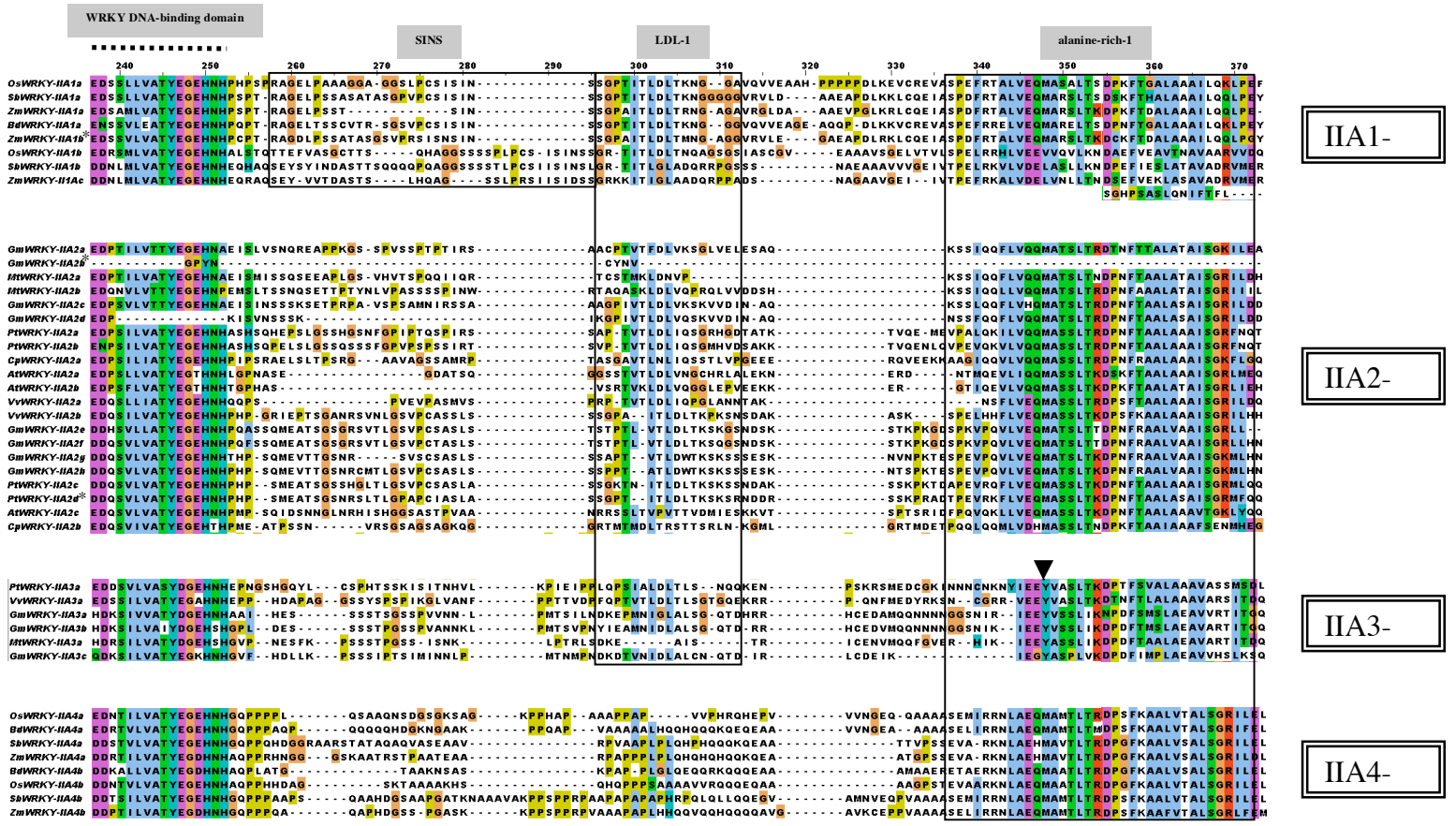
Appendix 2



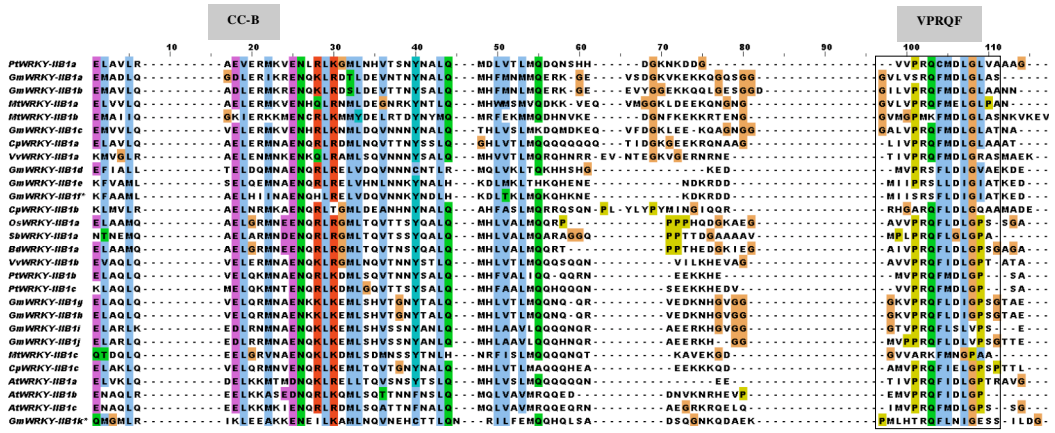
Appendix 2 continued



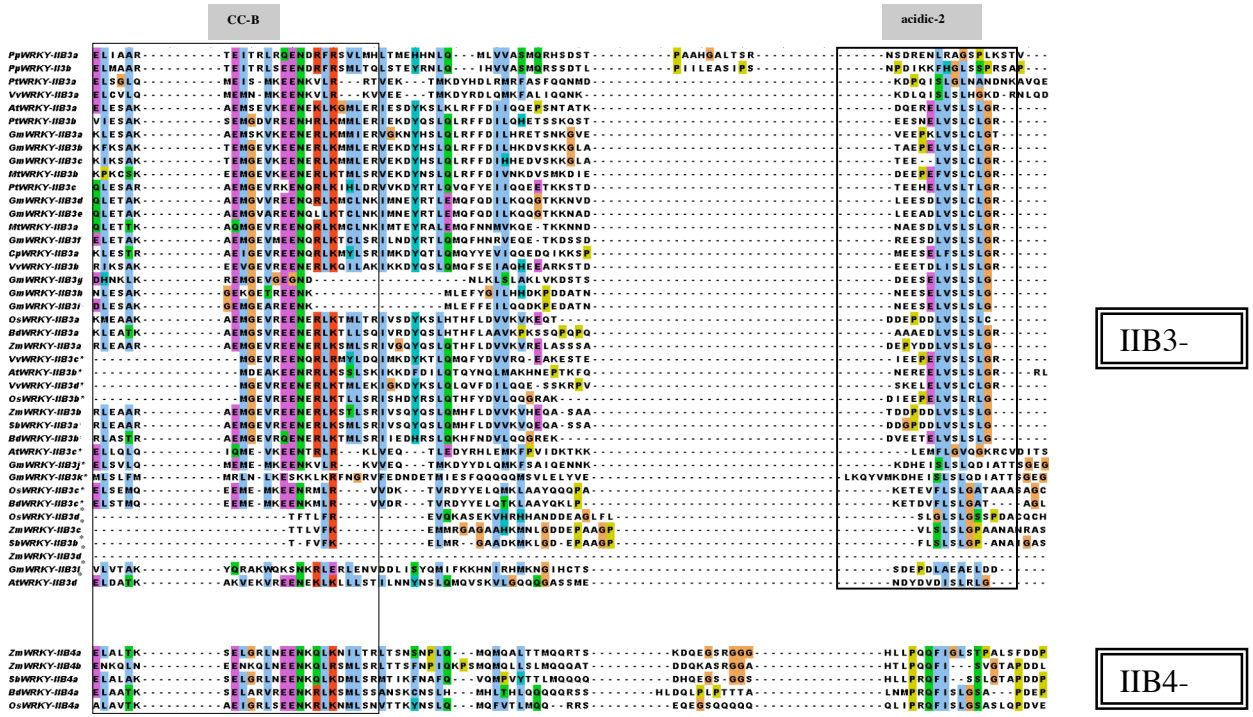
Appendix 2 continued



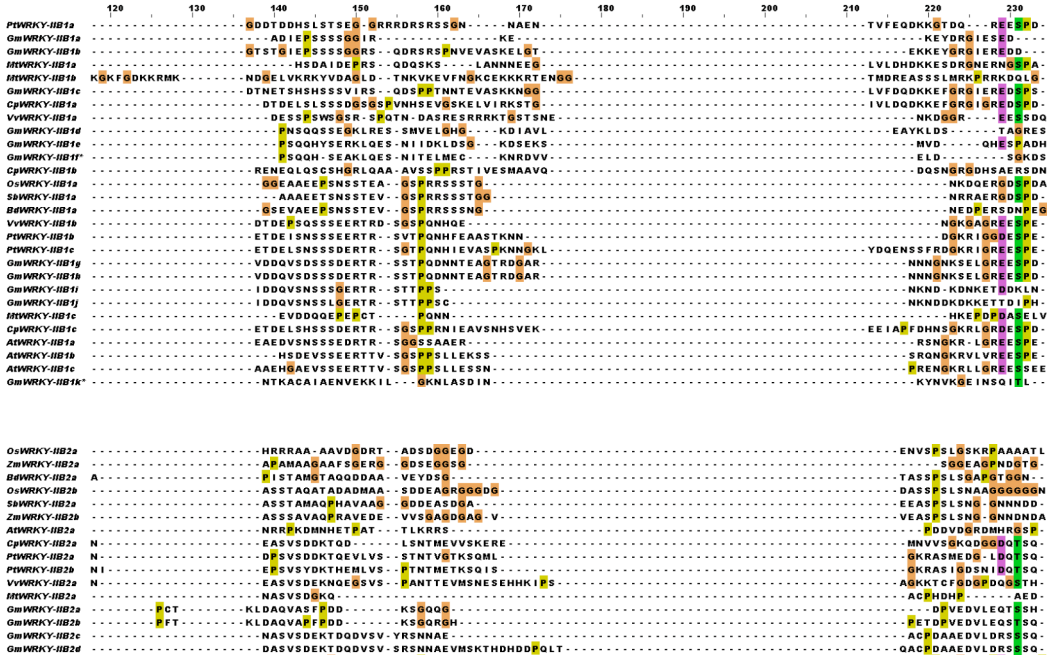
Appendix 3



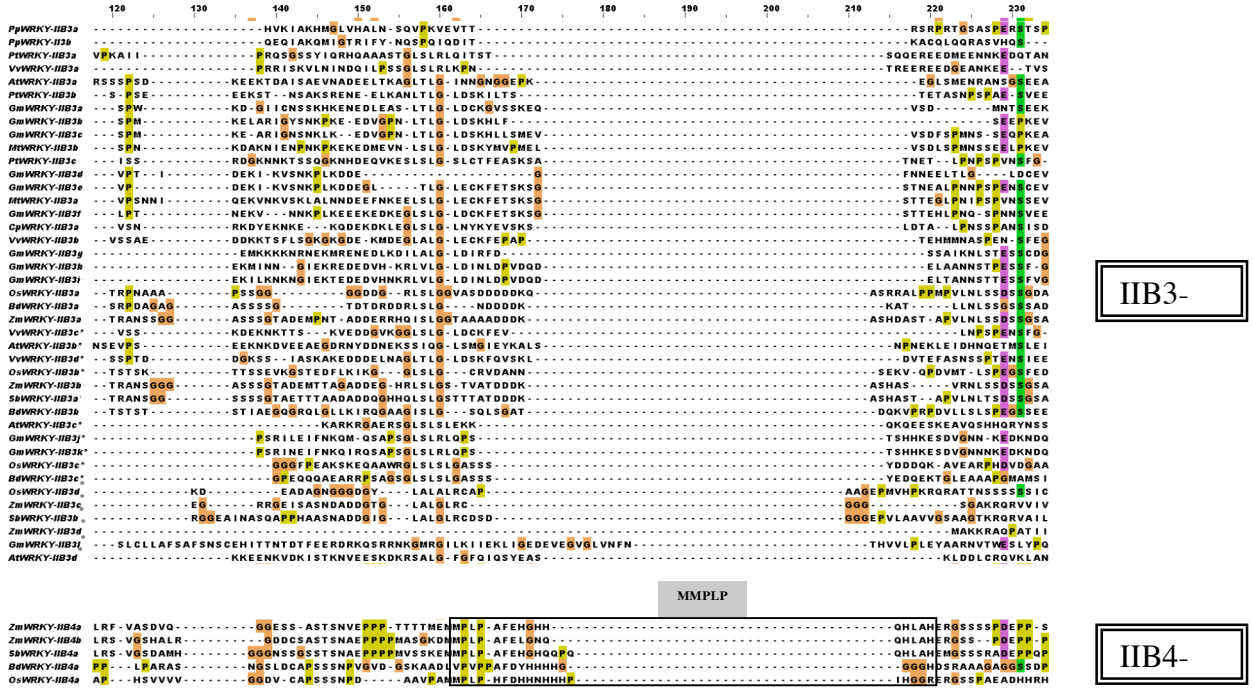
Appendix 3 continued



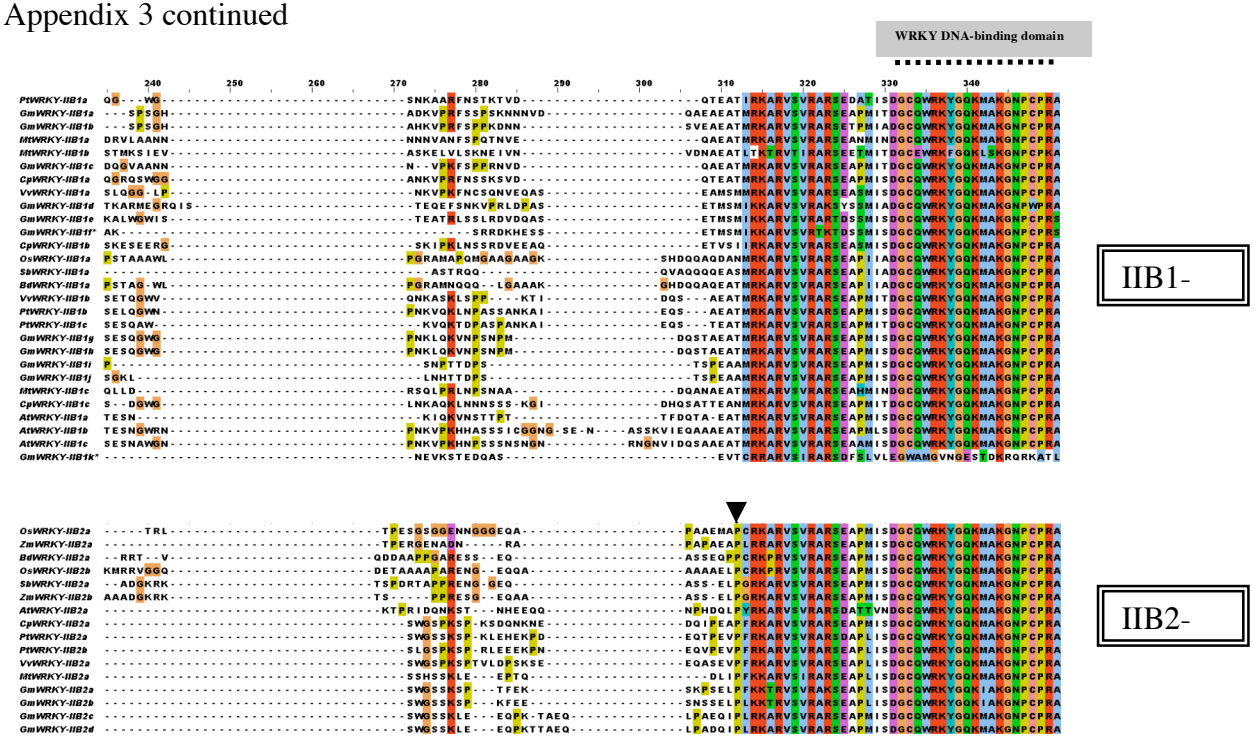
Appendix 3 continued



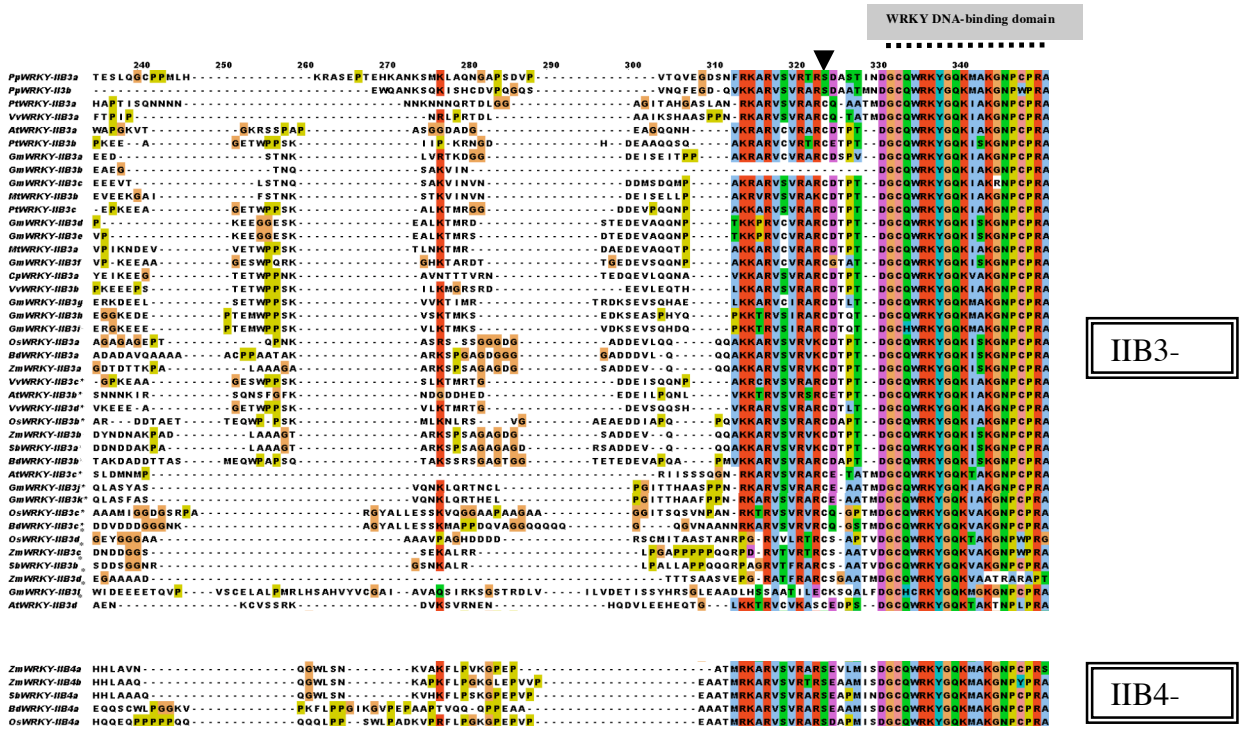
Appendix 3 continued



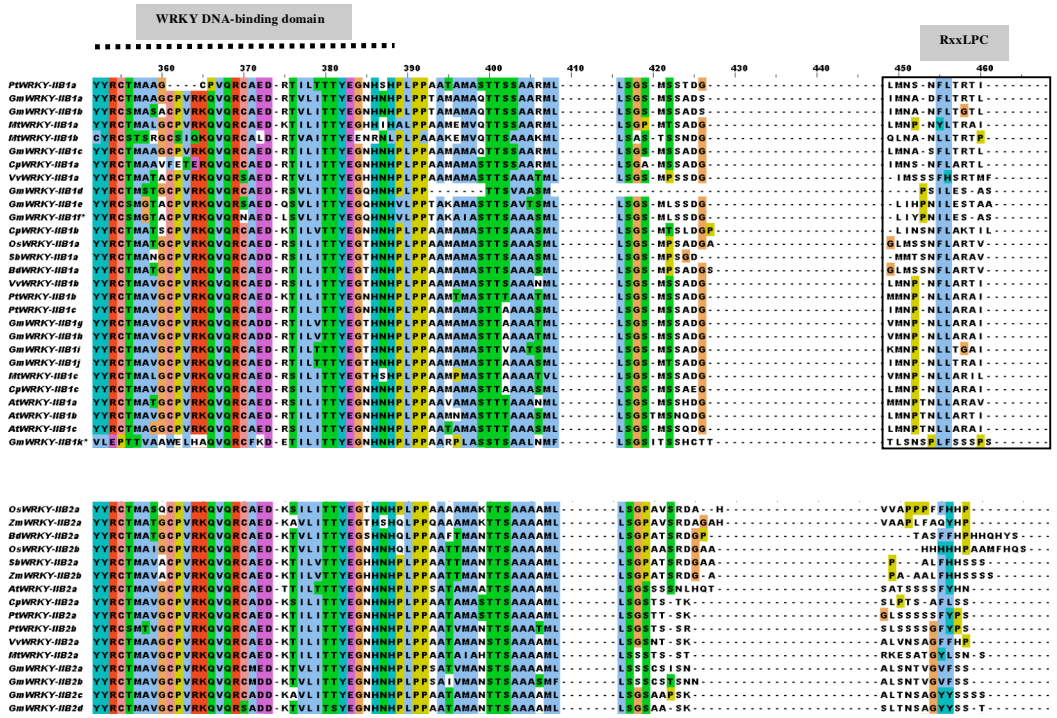
Appendix 3 continued



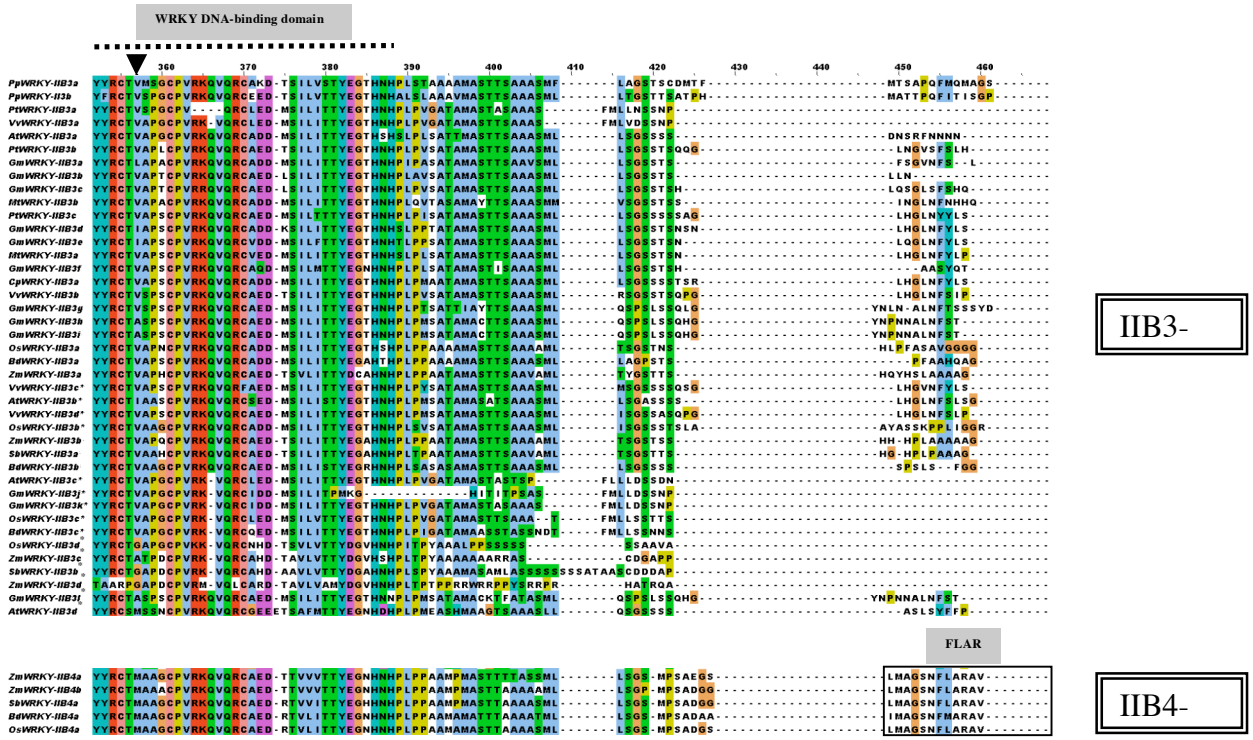
Appendix 3 continued



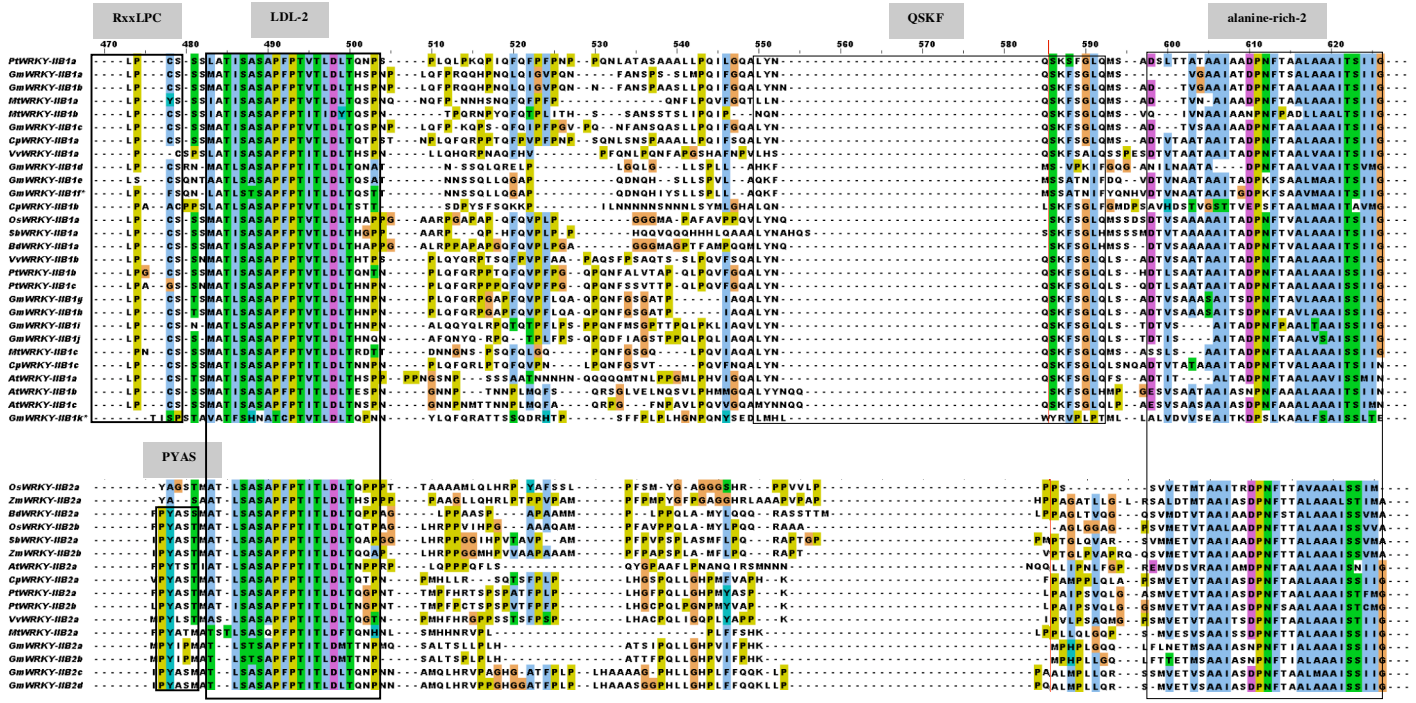
Appendix 3 continued



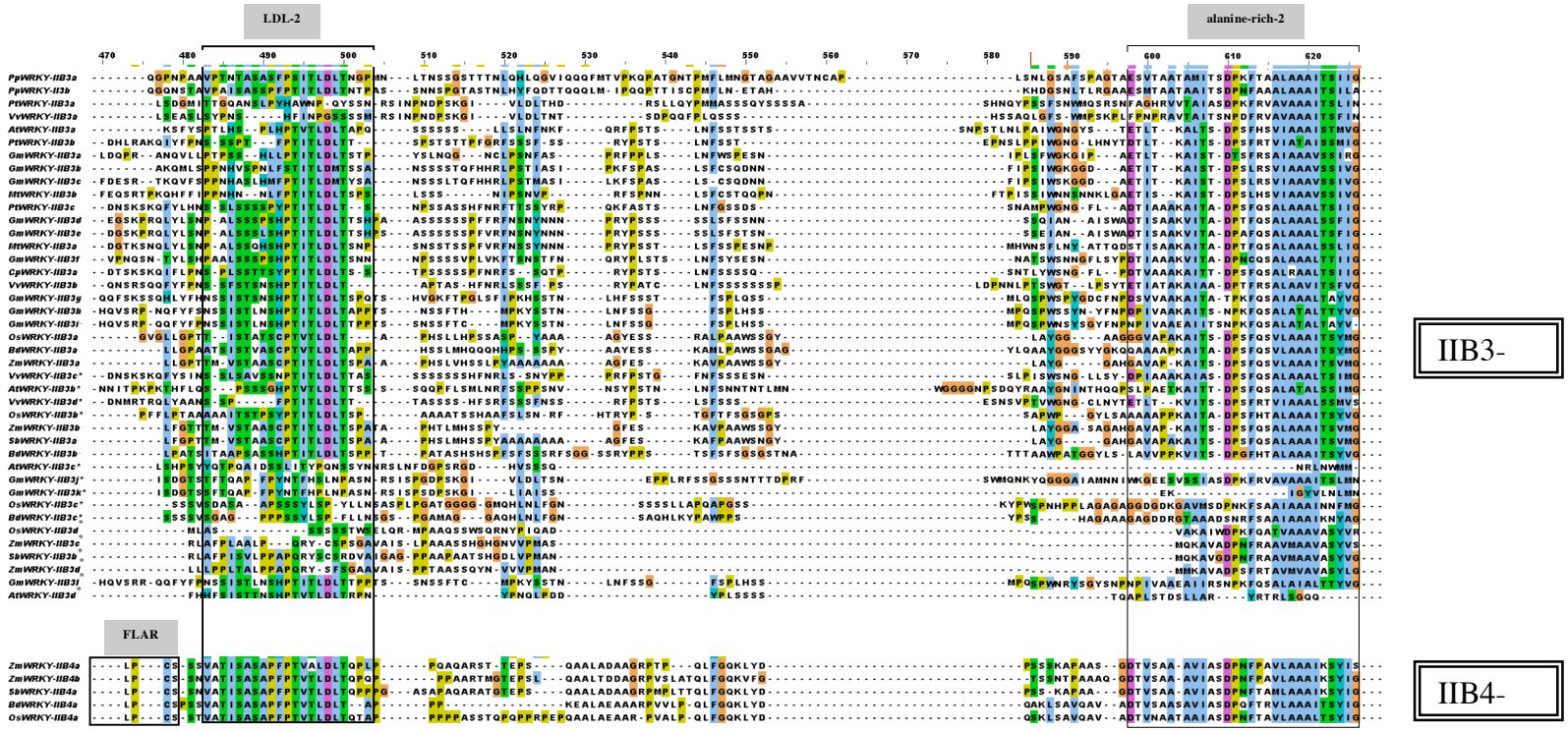
Appendix 3 continued



Appendix 3 continued

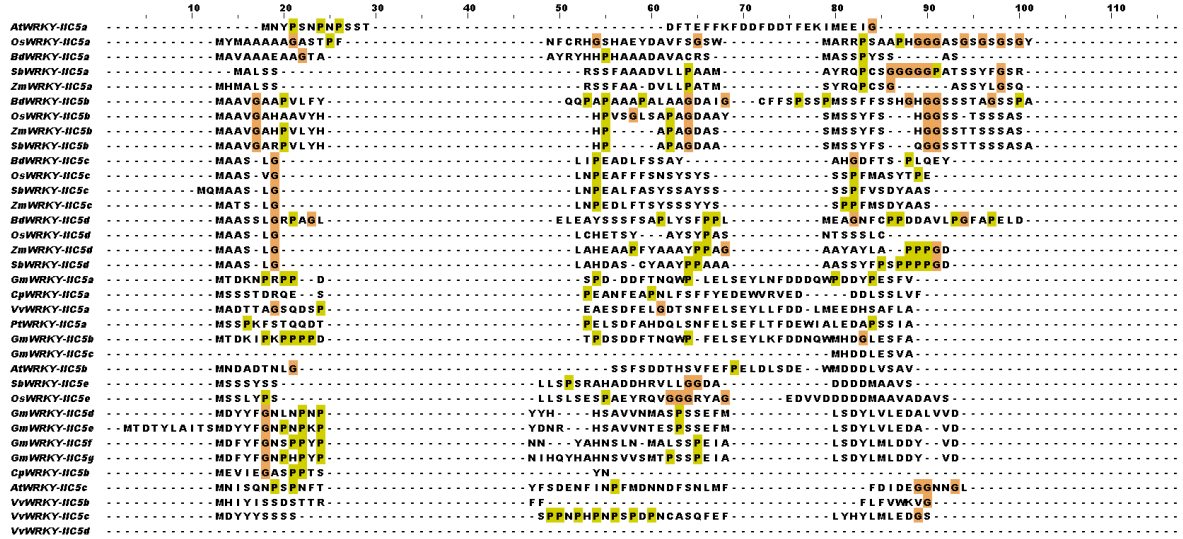


Appendix 3 continued



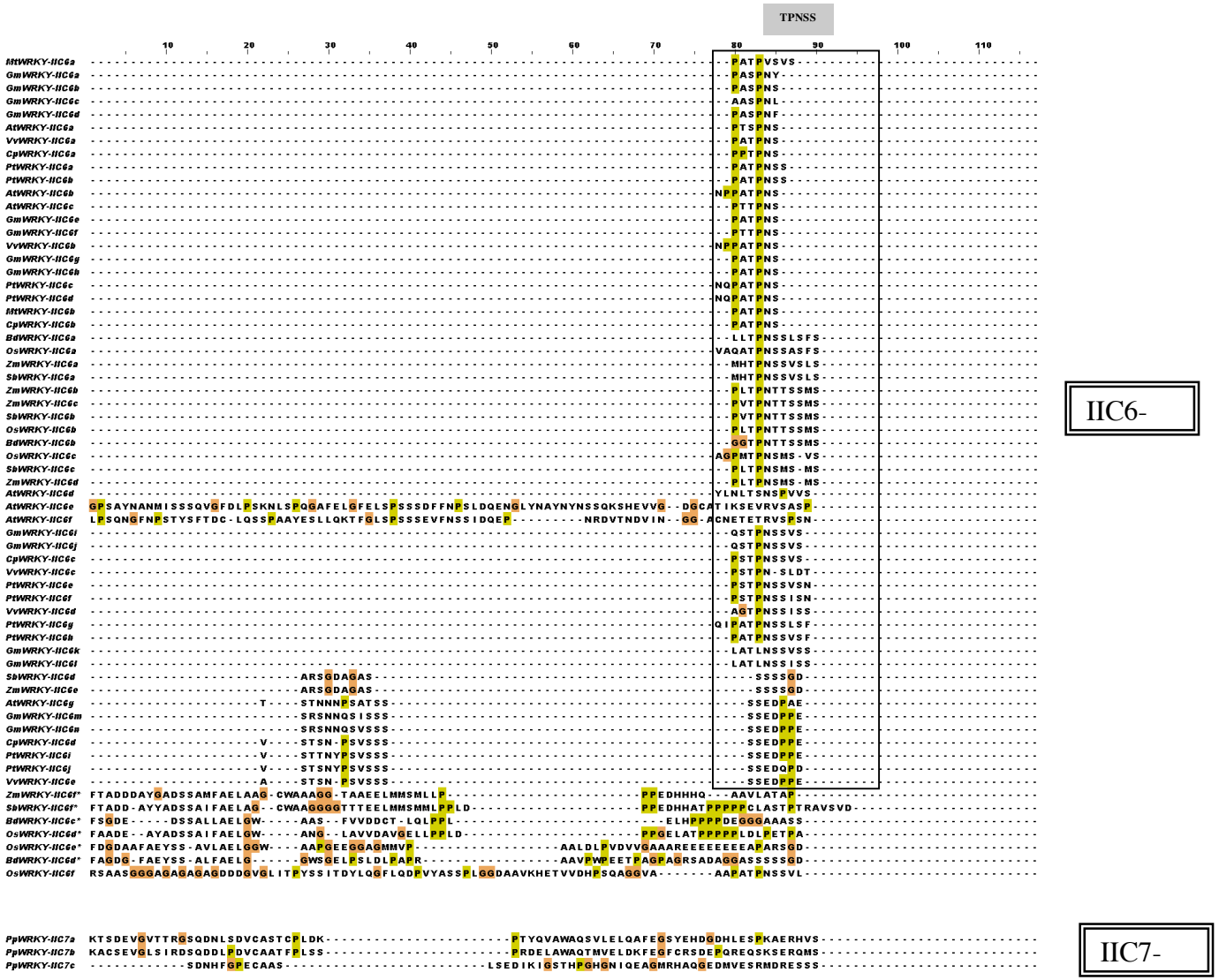
Appendix 4

Appendix 4 continued

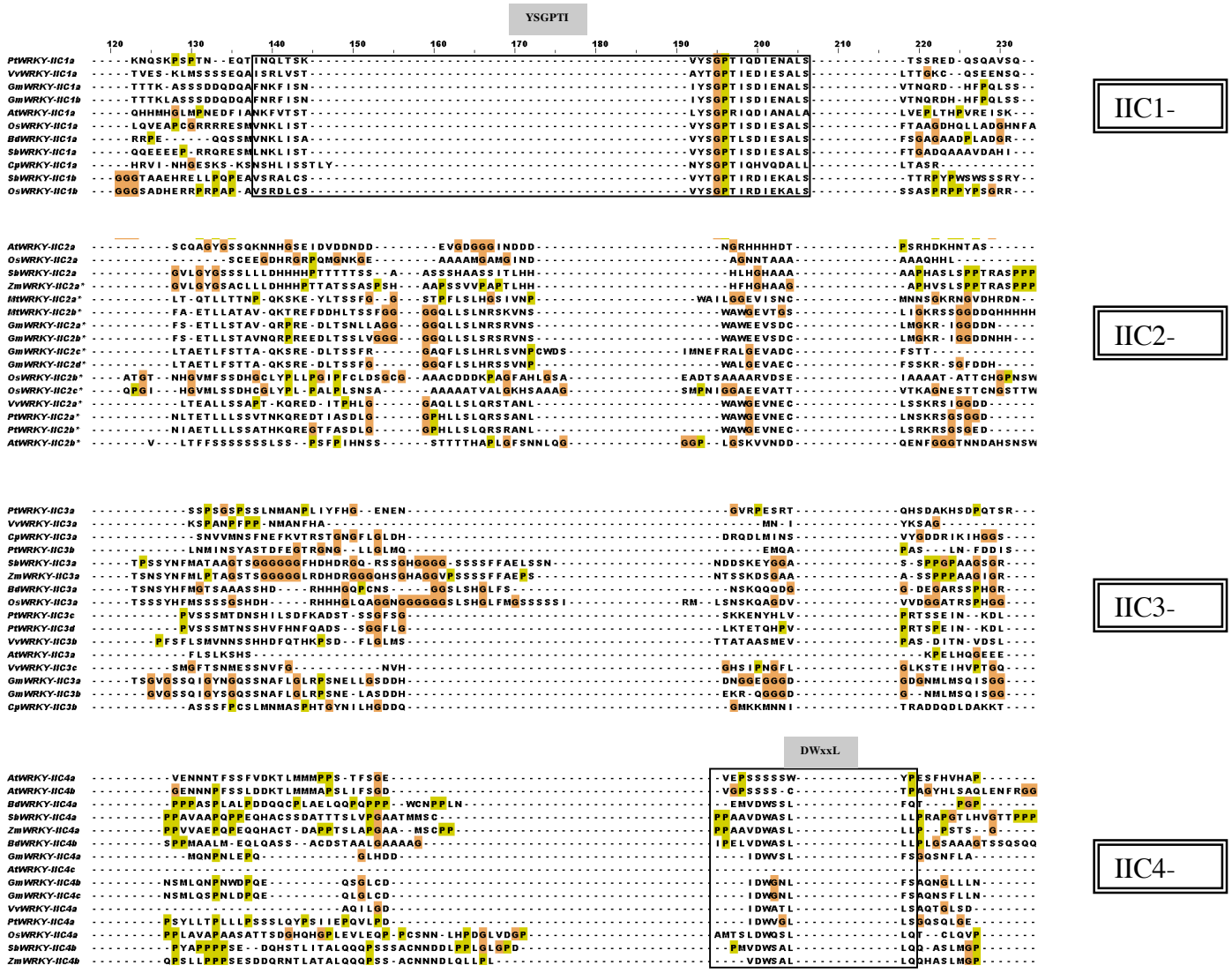


IIC5-

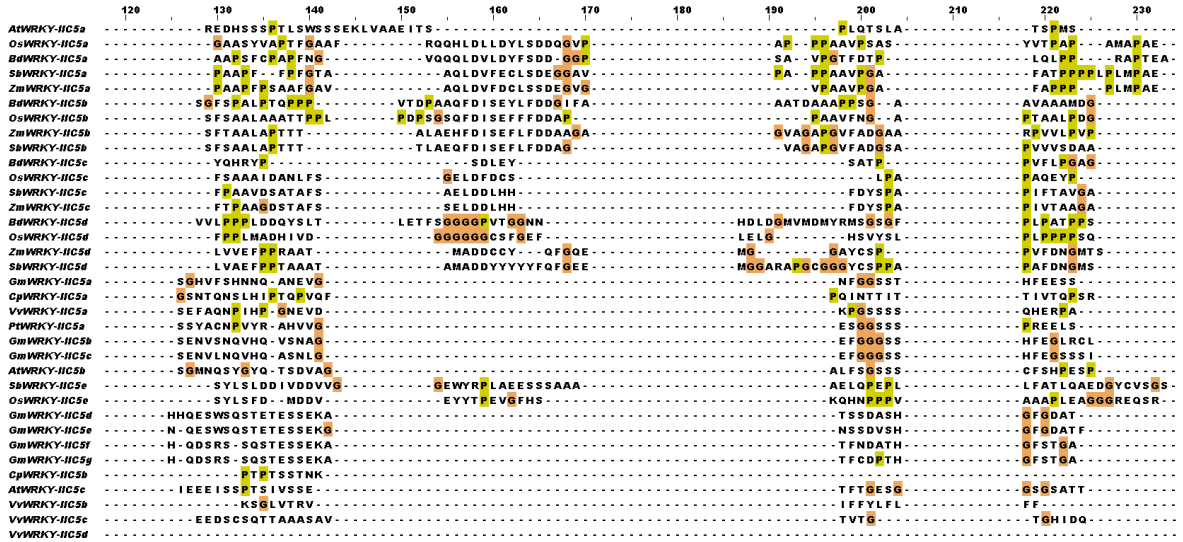
Appendix 4 continued



Appendix 4 continued

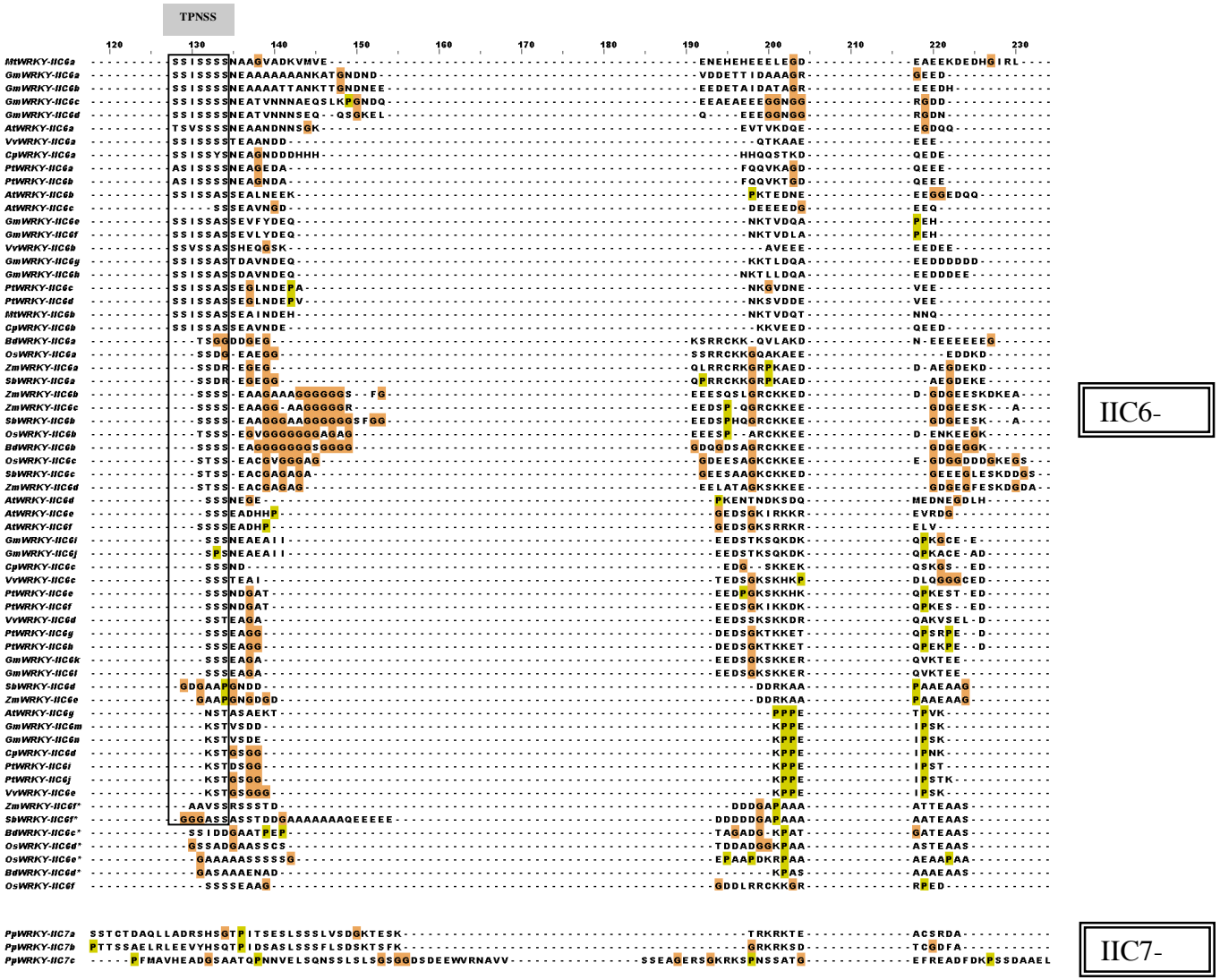


Appendix 4 continued

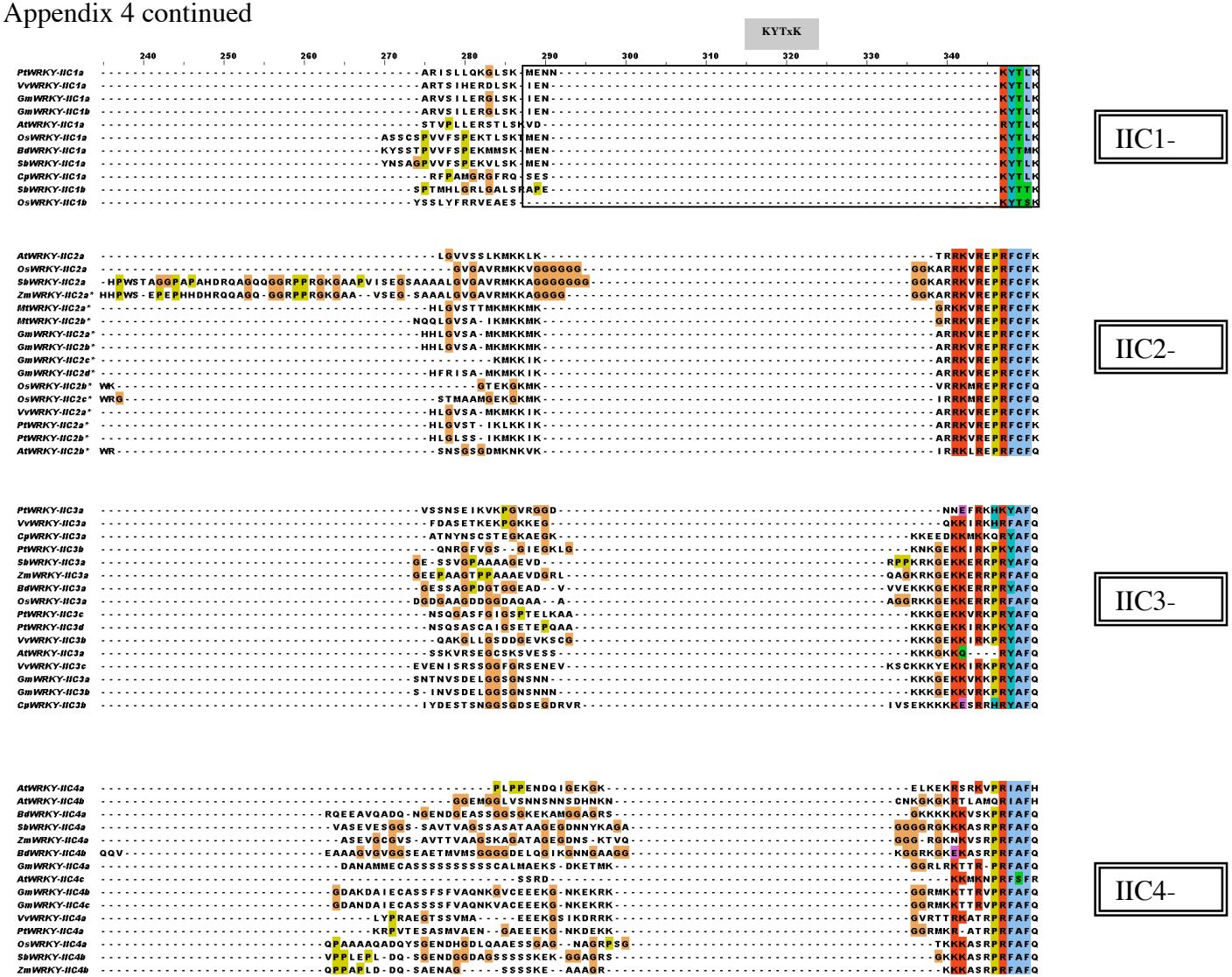


IIC5-

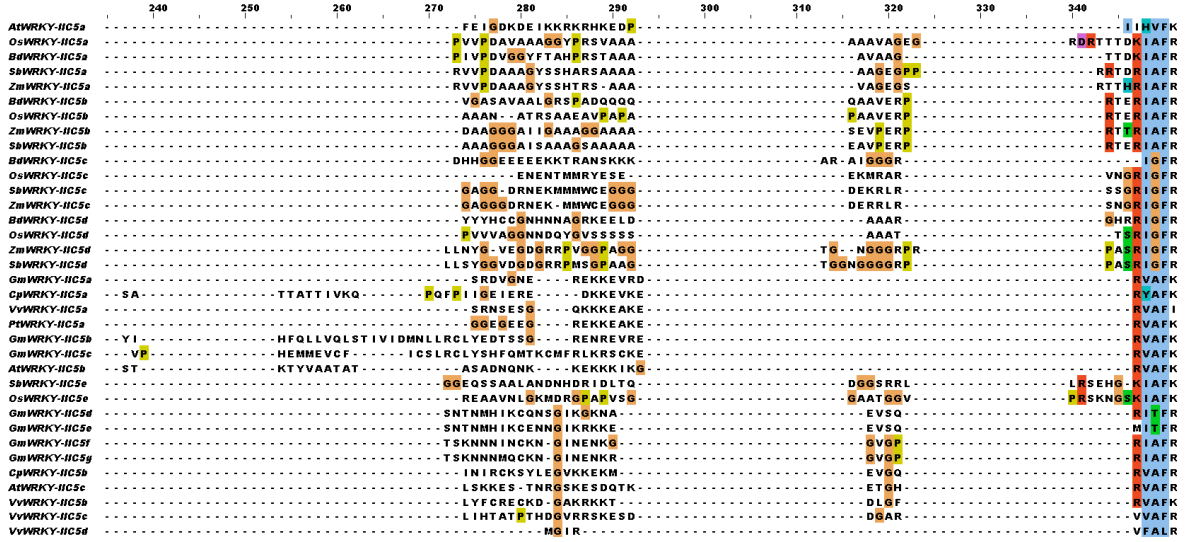
Appendix 4 continued



Appendix 4 continued

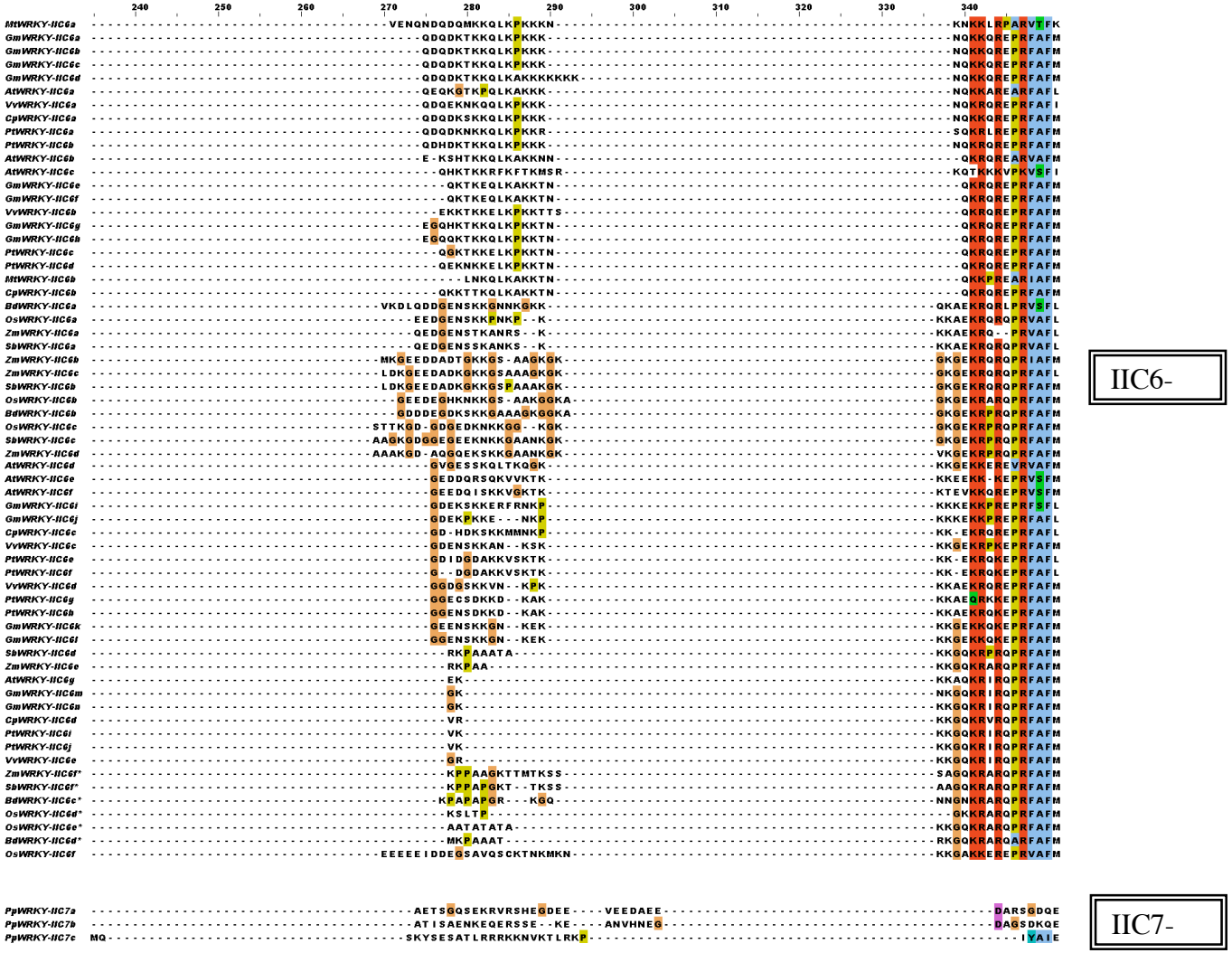


Appendix 4 continued

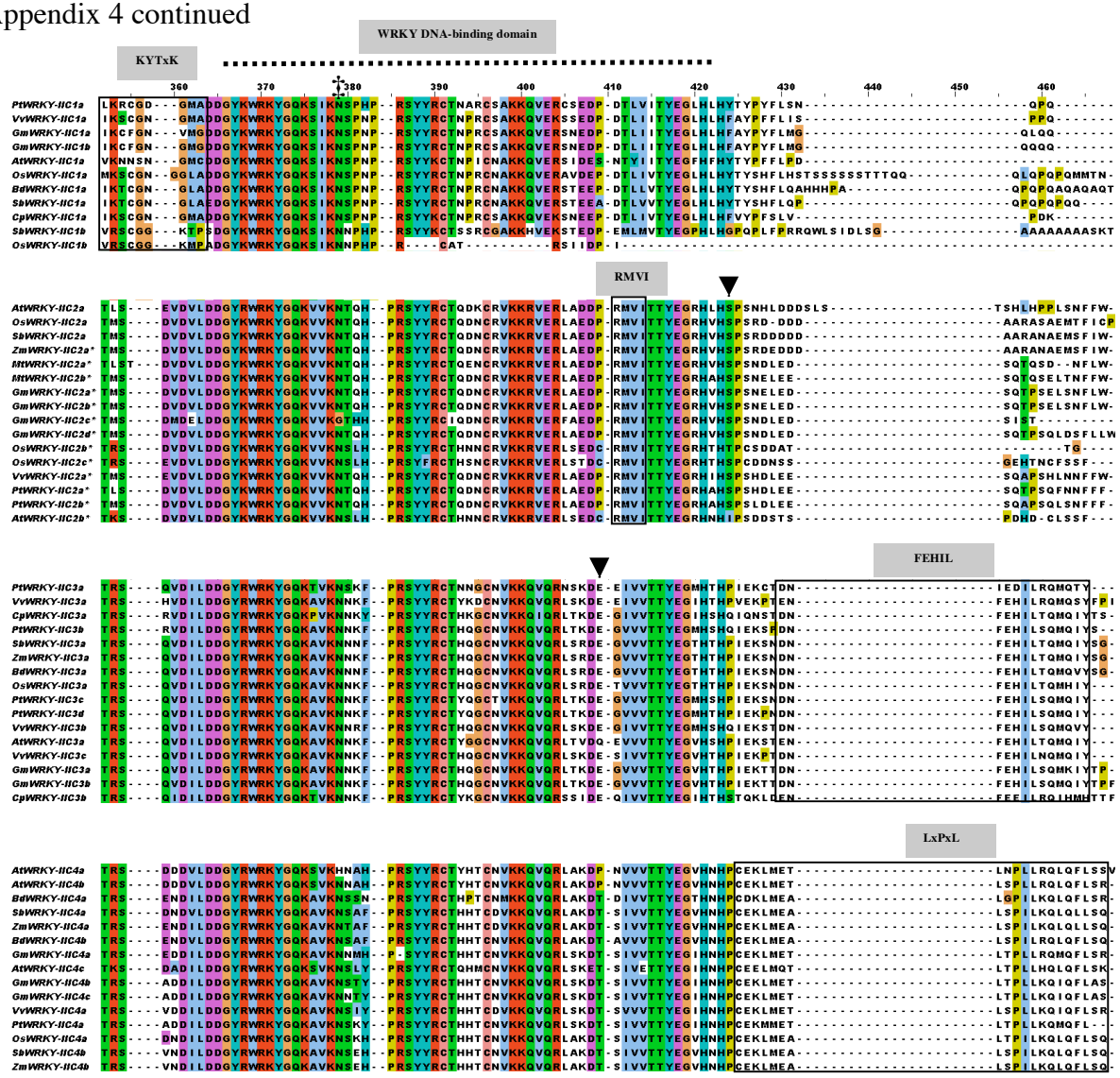


IIC5-

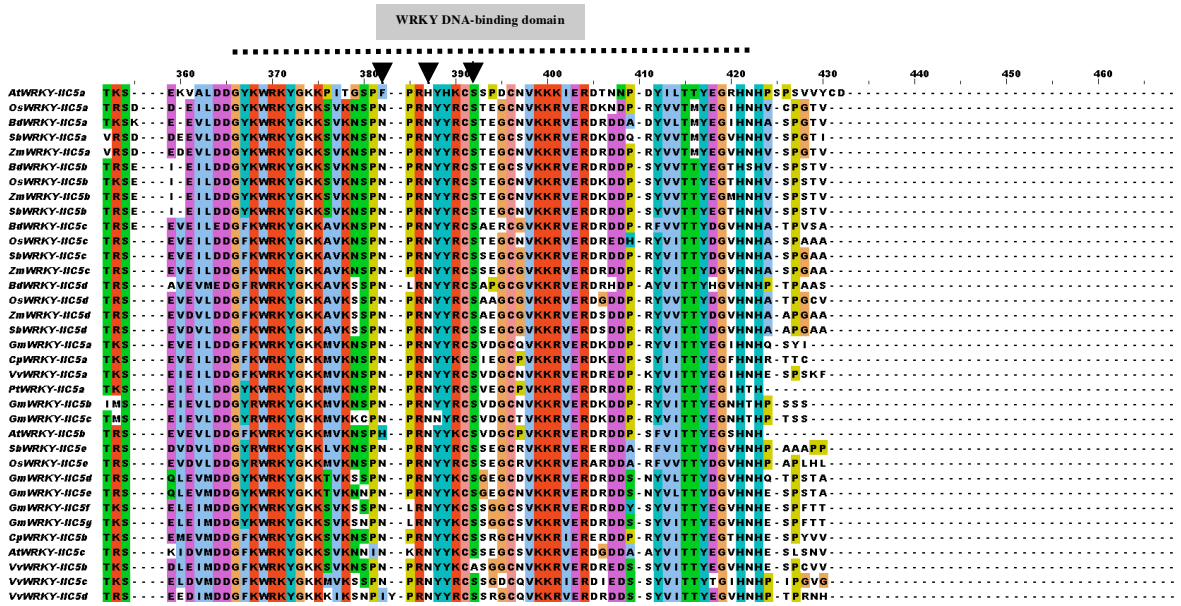
Appendix 4 continued



Appendix 4 continued

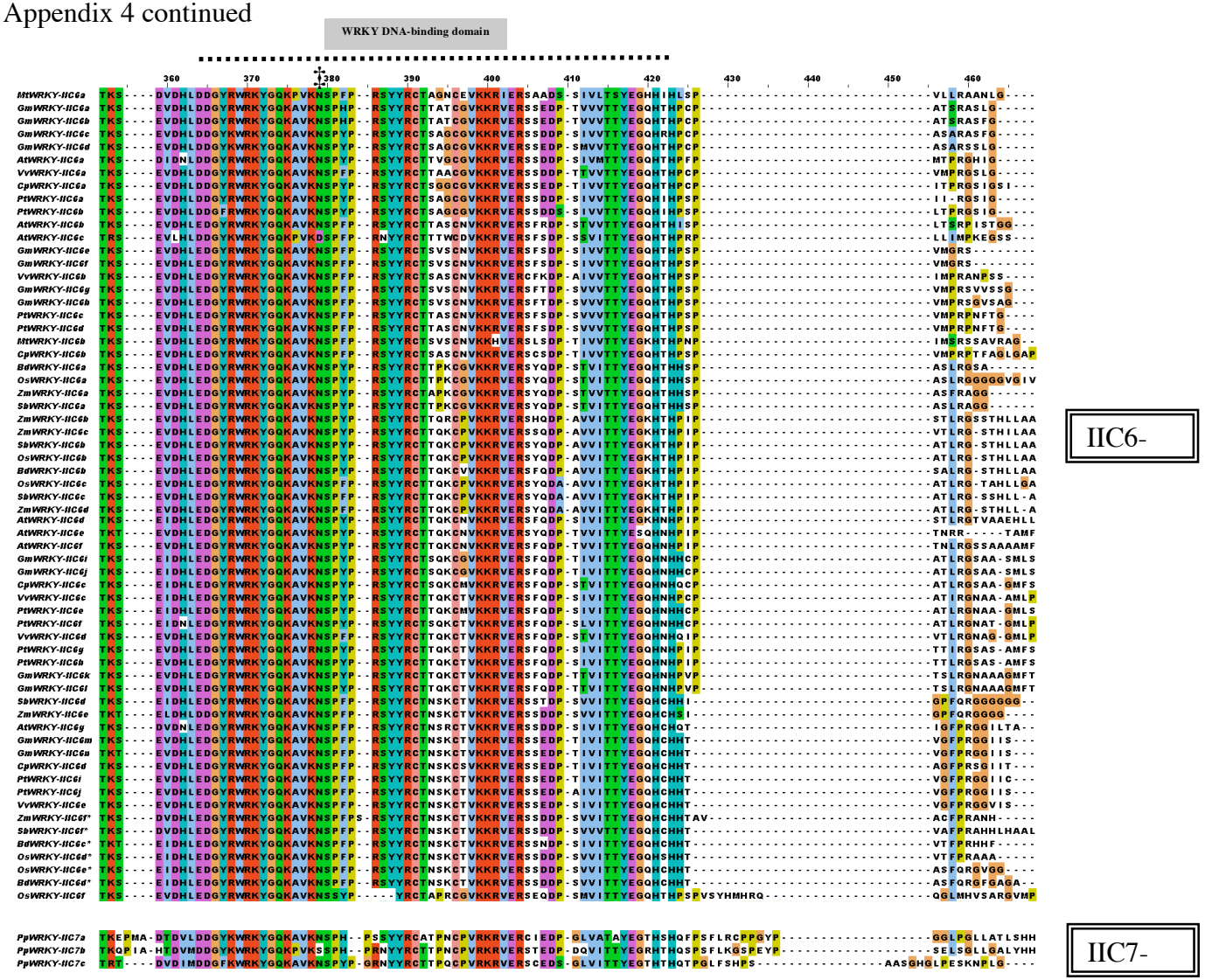


Appendix 4 continued

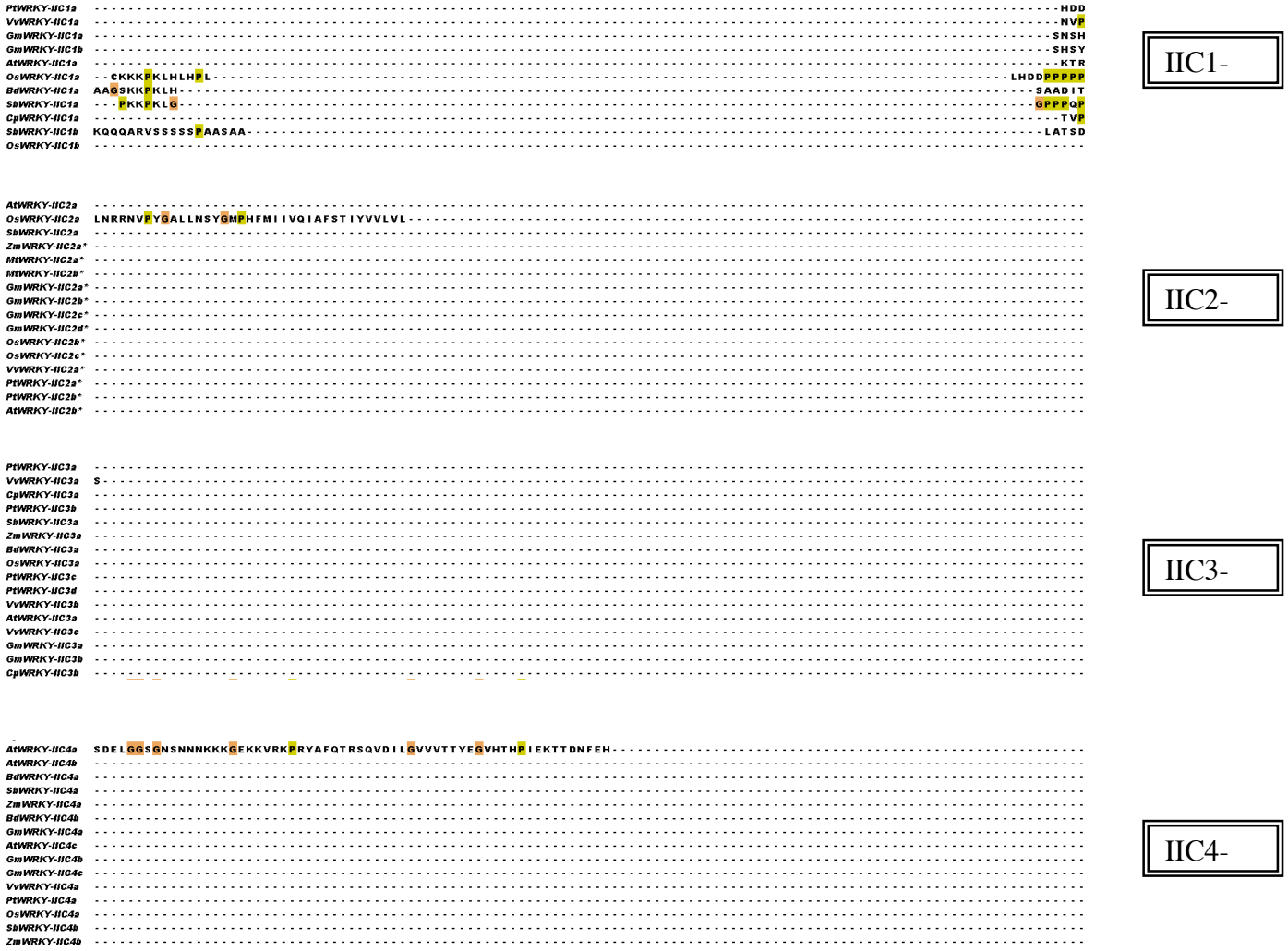


IIC5-

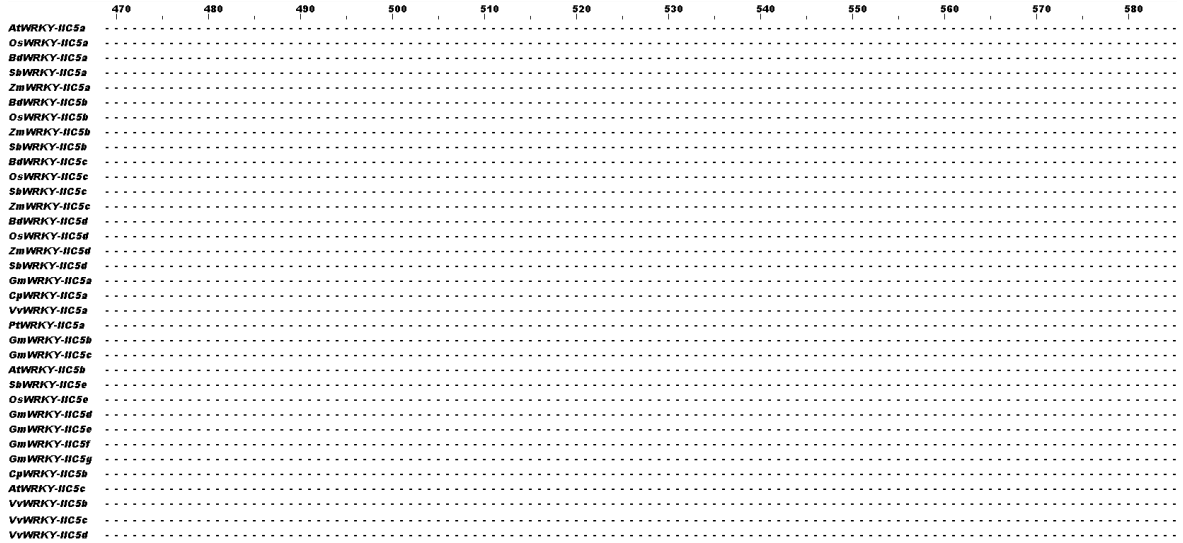
Appendix 4 continued



Appendix 4 continued

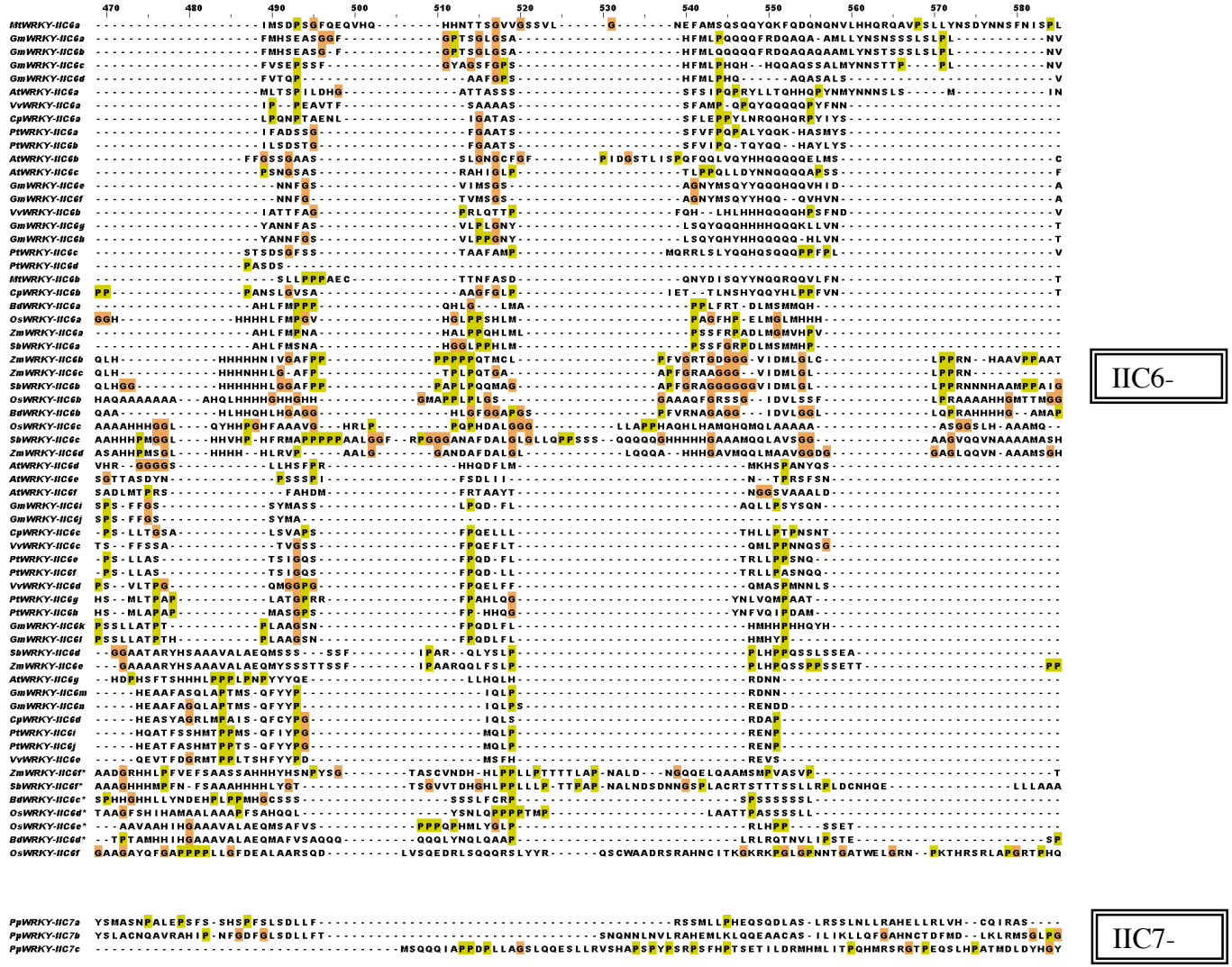


Appendix 4 continued

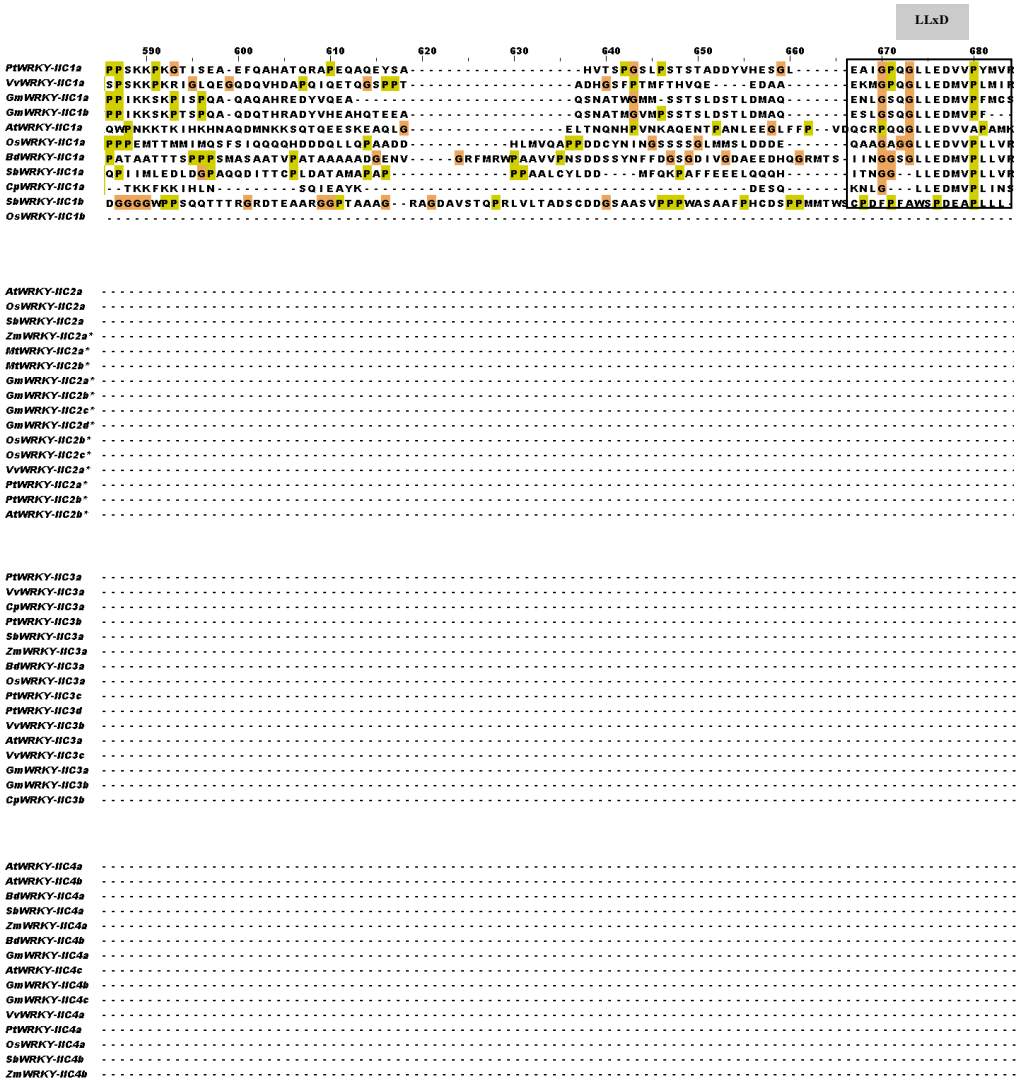


IIC5-

Appendix 4 continued



Appendix 4 continued



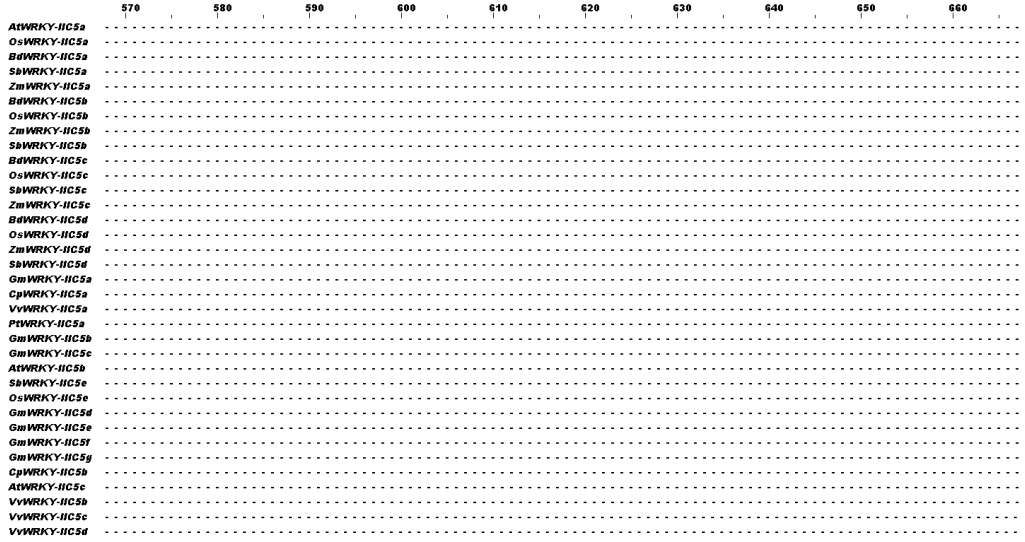
IIC1-

IIC2-

IIC3-

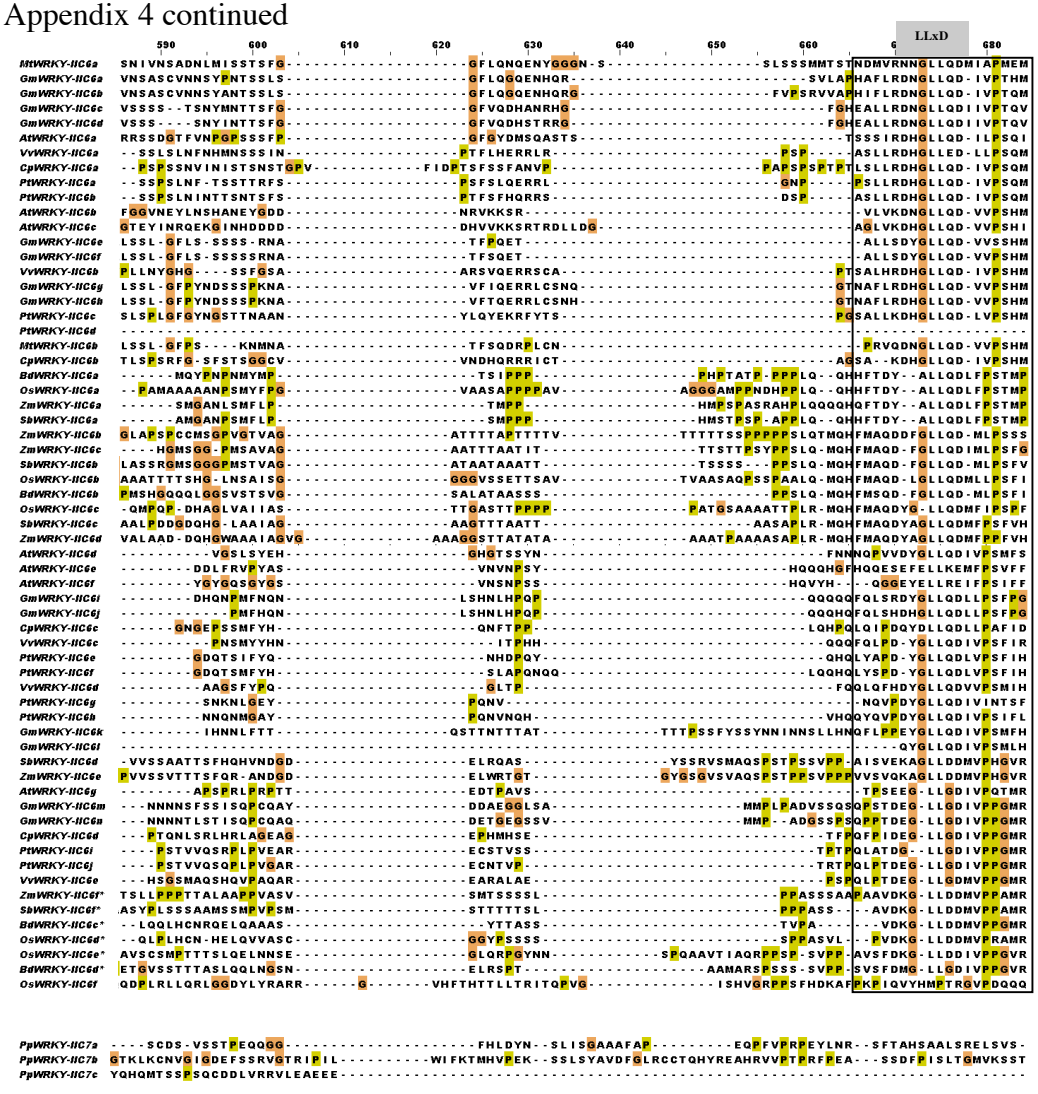
IIC4-

Appendix 4 continued

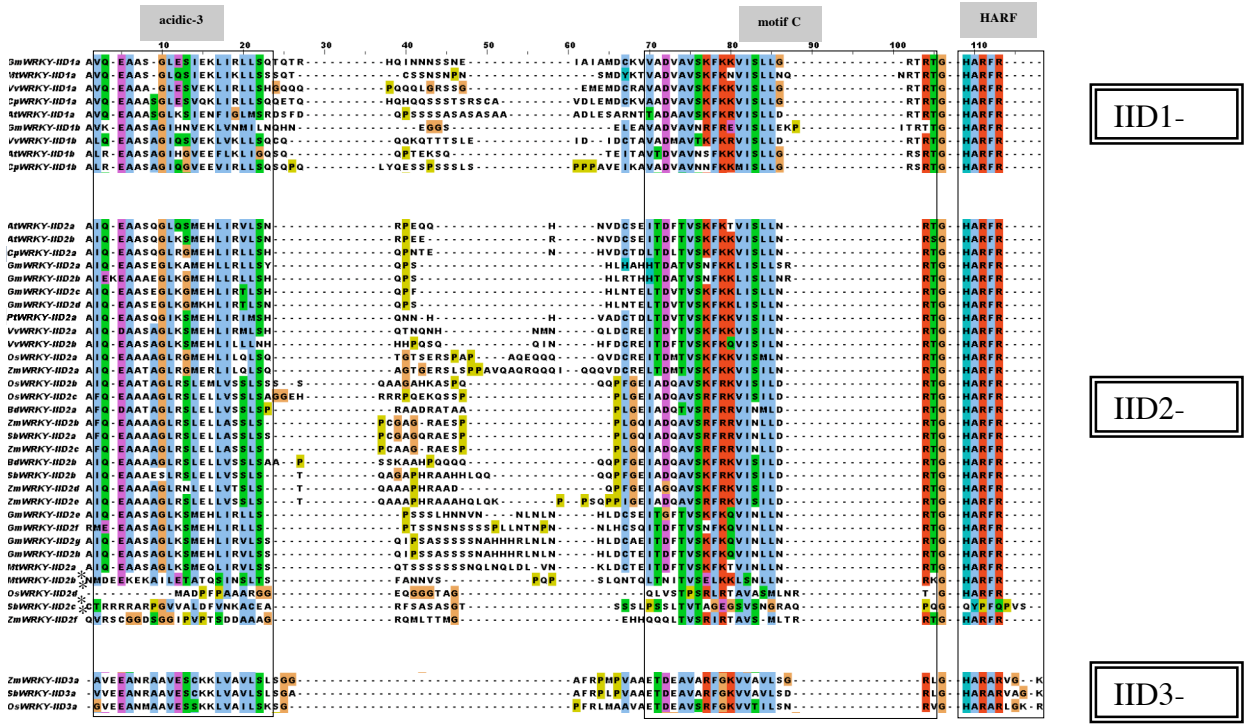


IIC5-

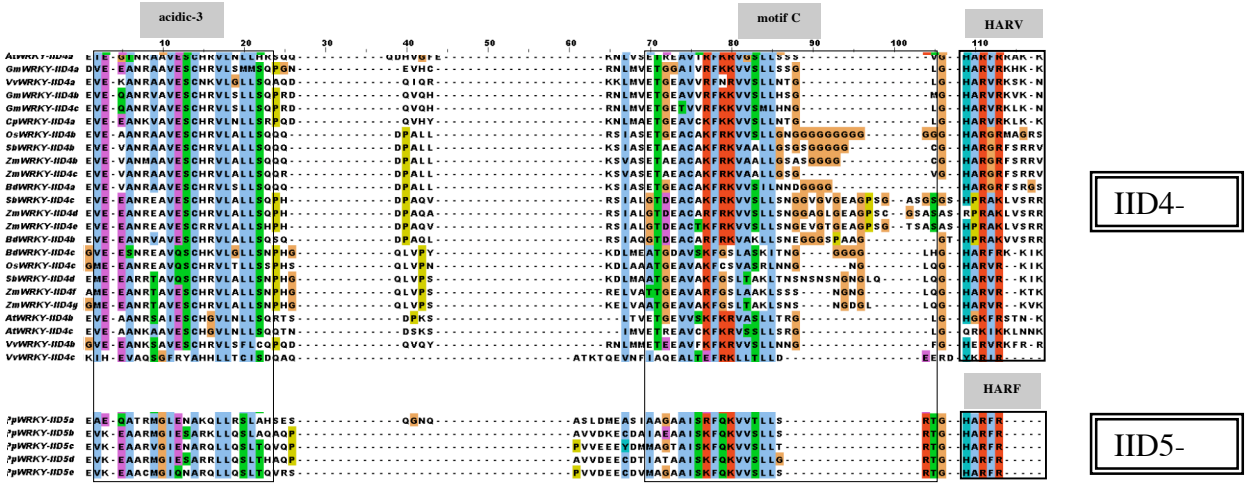
Appendix 4 continued



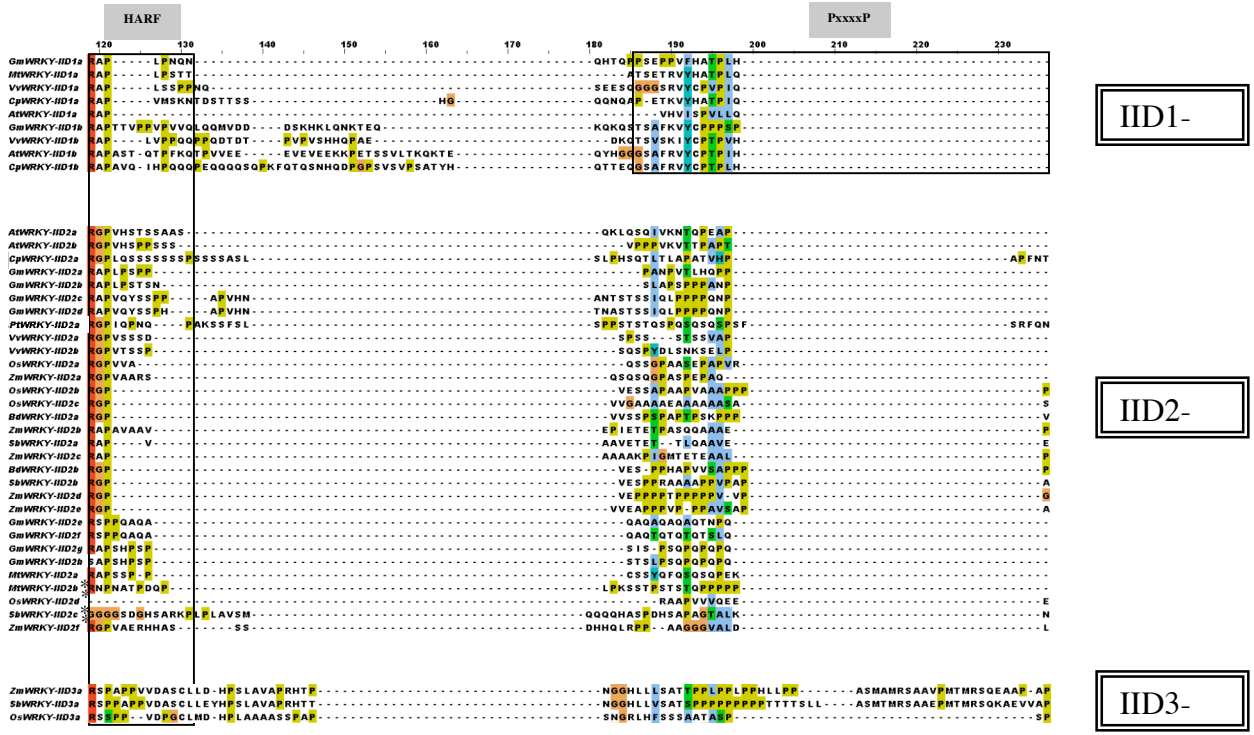
Appendix 5



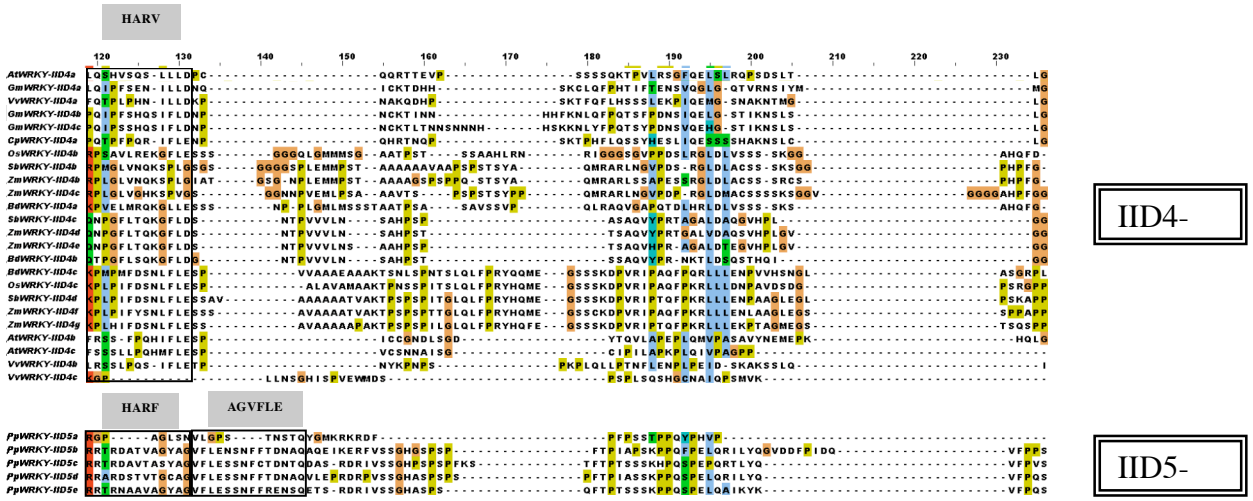
Appendix 5 continued



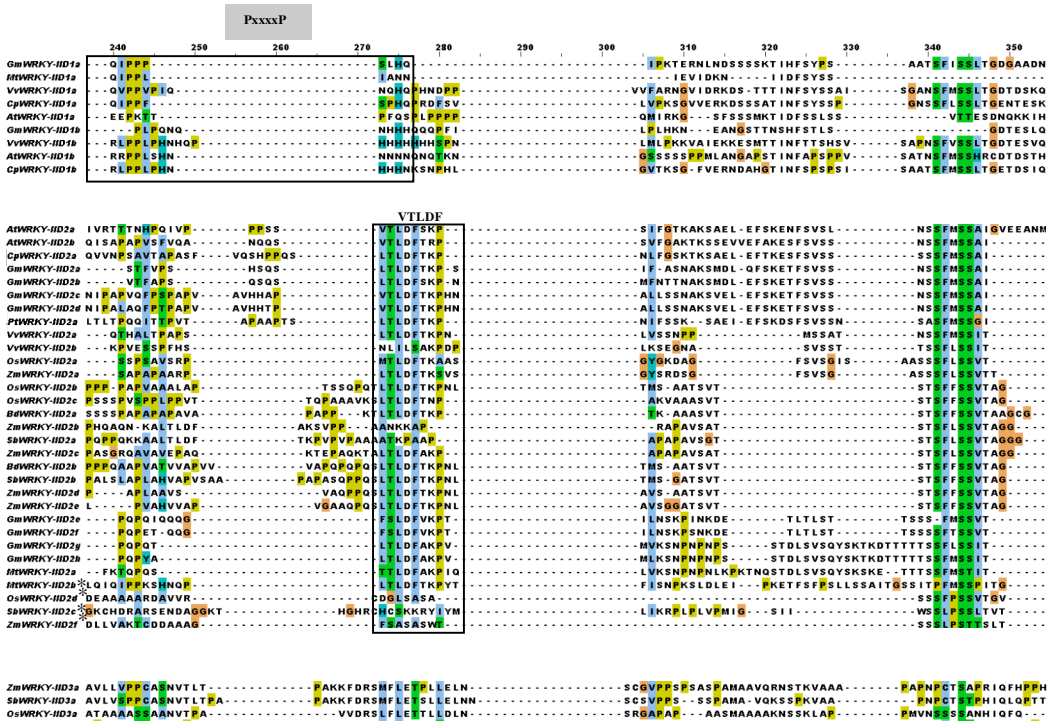
Appendix 5 continued



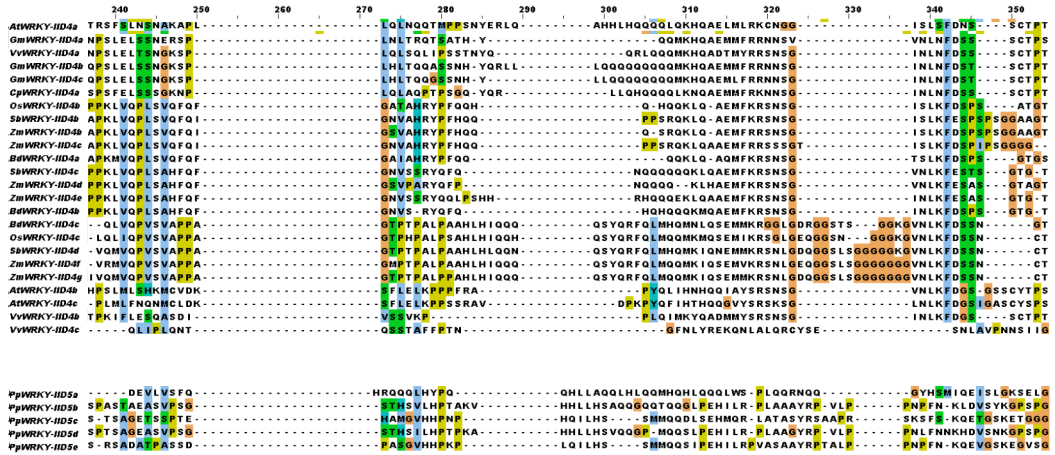
Appendix 5 continued



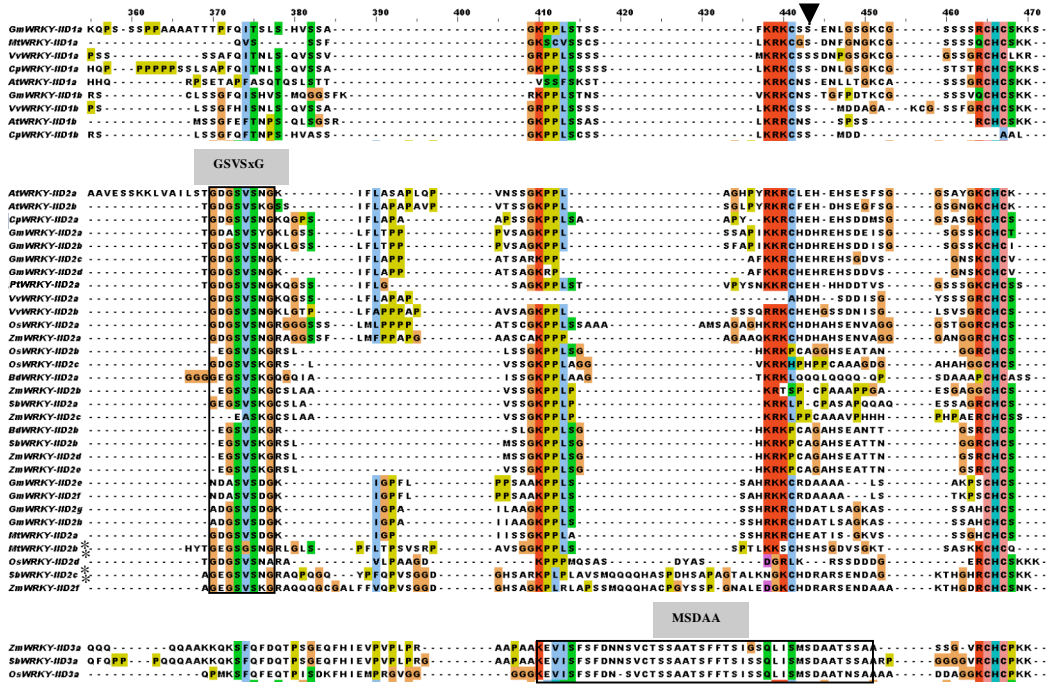
Appendix 5 continued



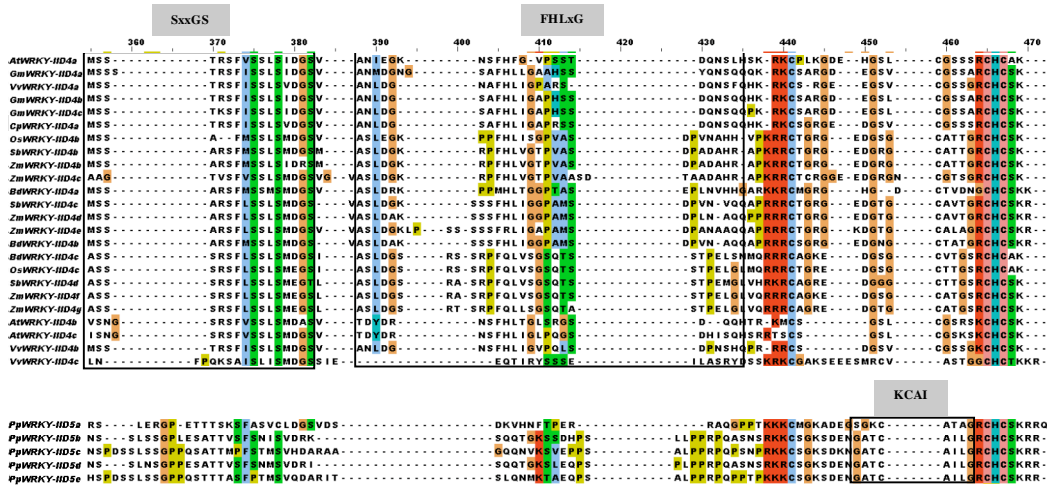
Appendix 5 continued



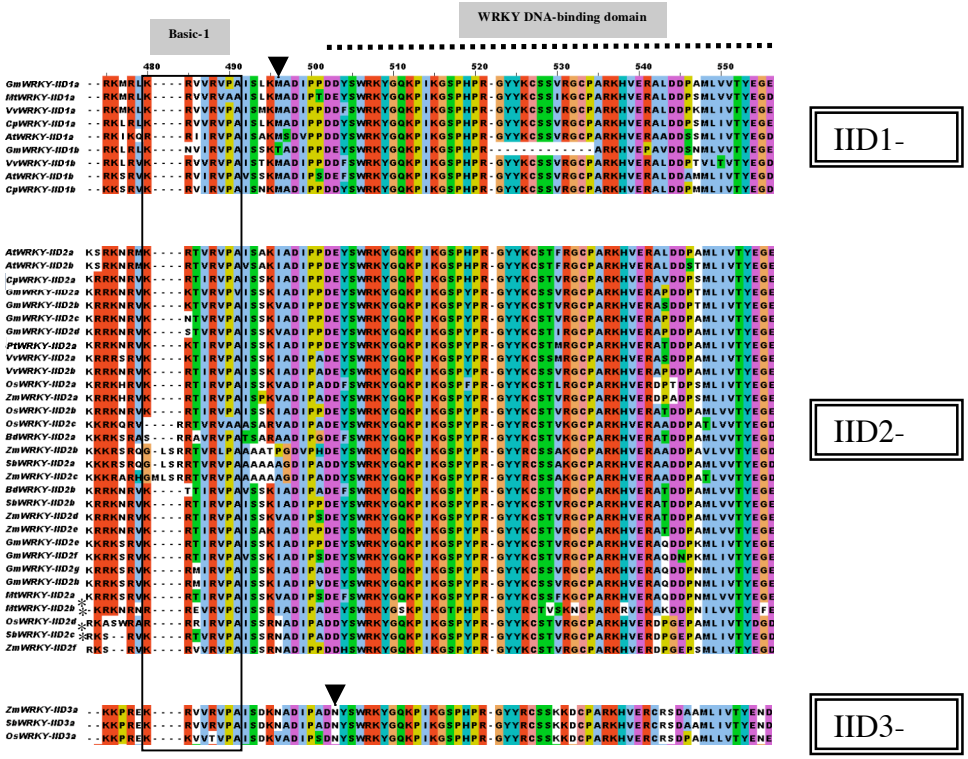
Appendix 5 continued



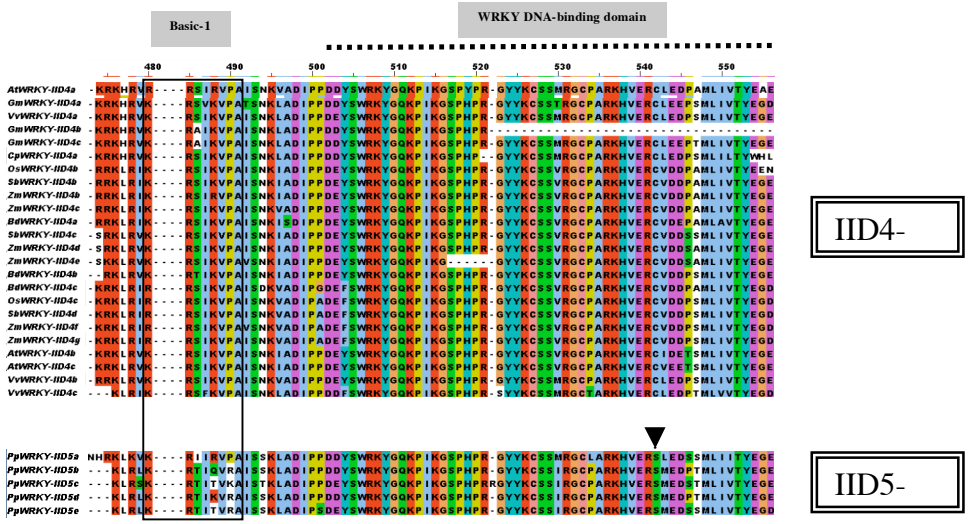
Appendix 5 continued



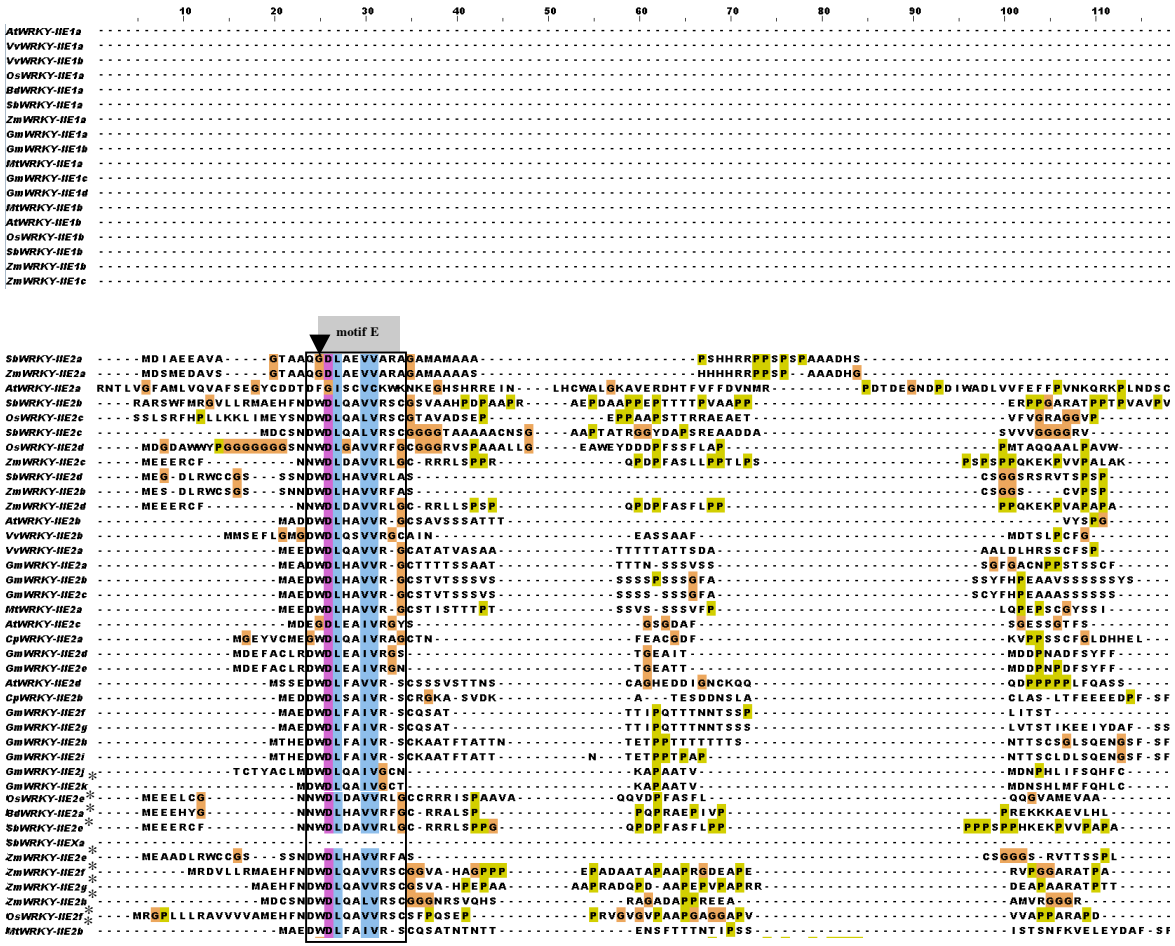
Appendix 5 continued



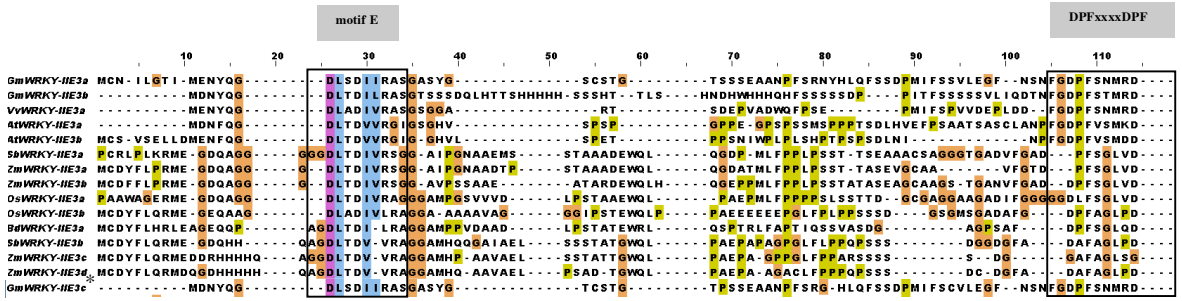
Appendix 5 continued



Appendix 6

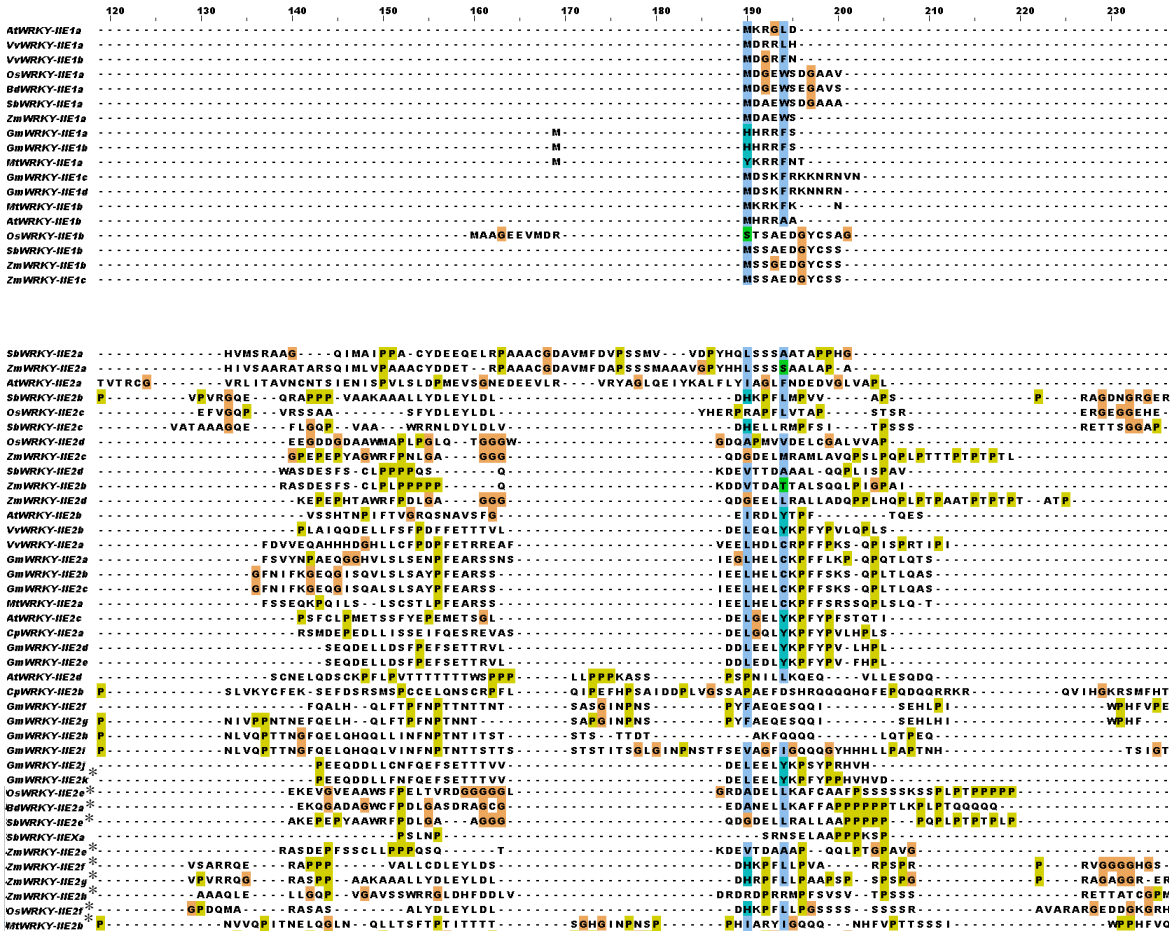


Appendix 6 continued

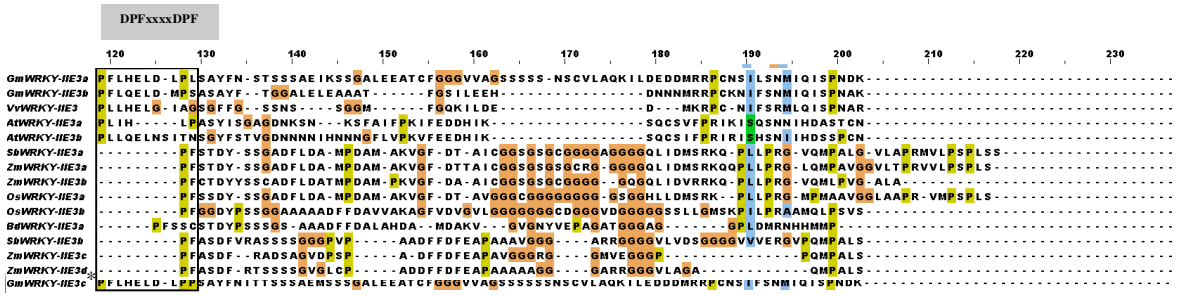


IIE3-

Appendix 6 continued

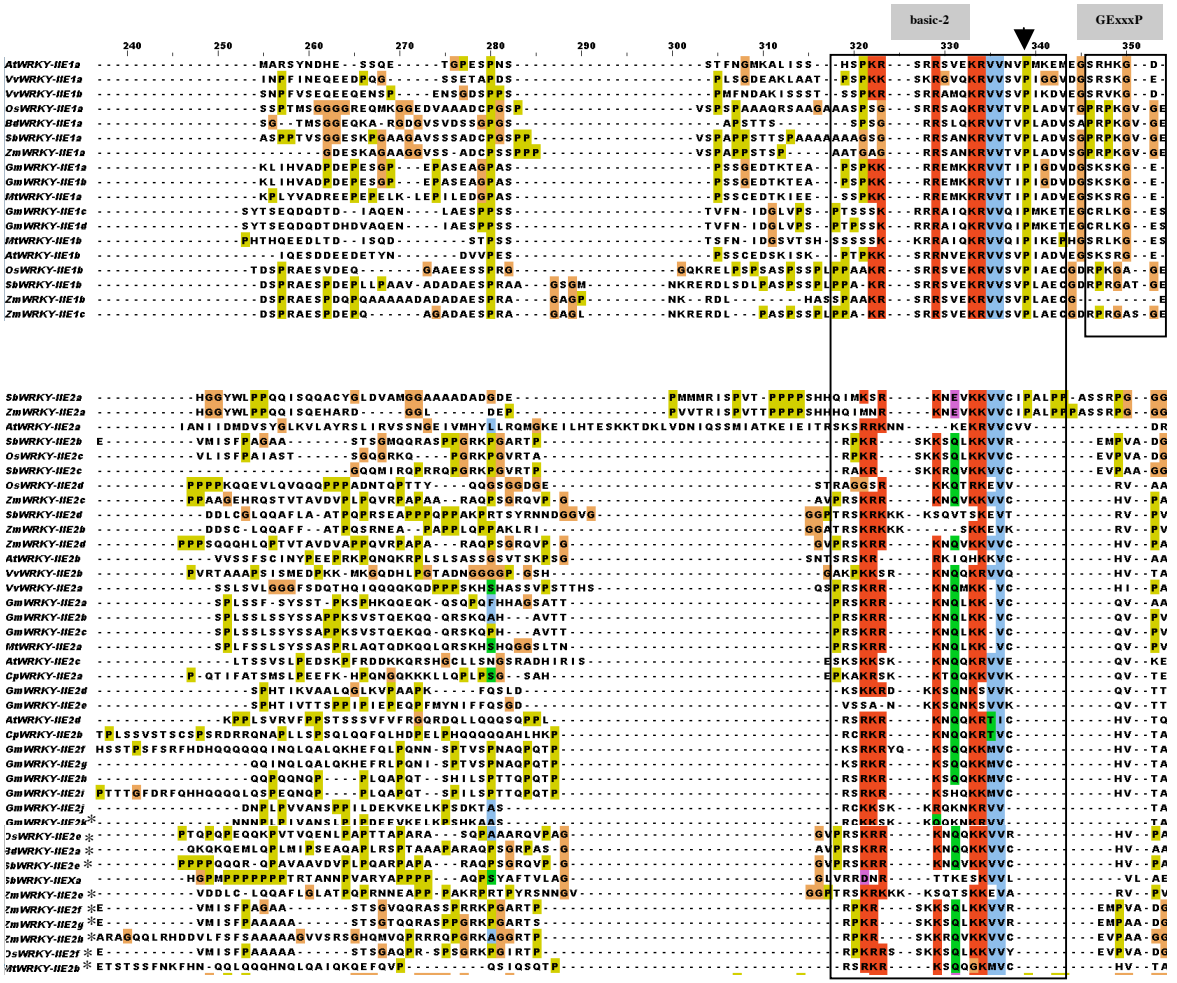


Appendix 6 continued

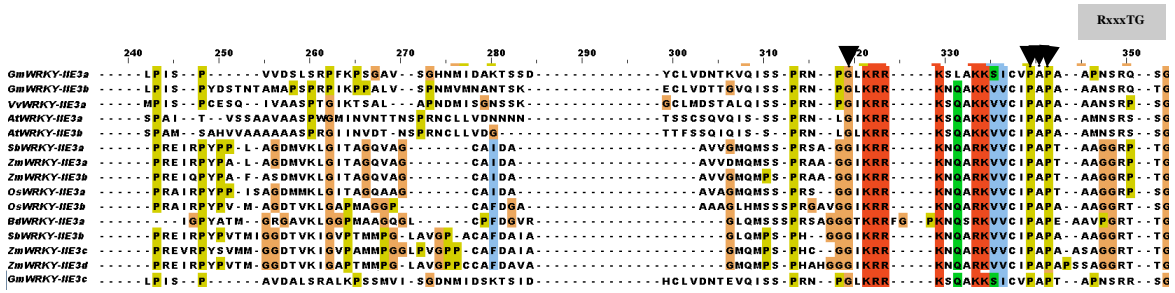


IIE3-

Appendix 6 continued

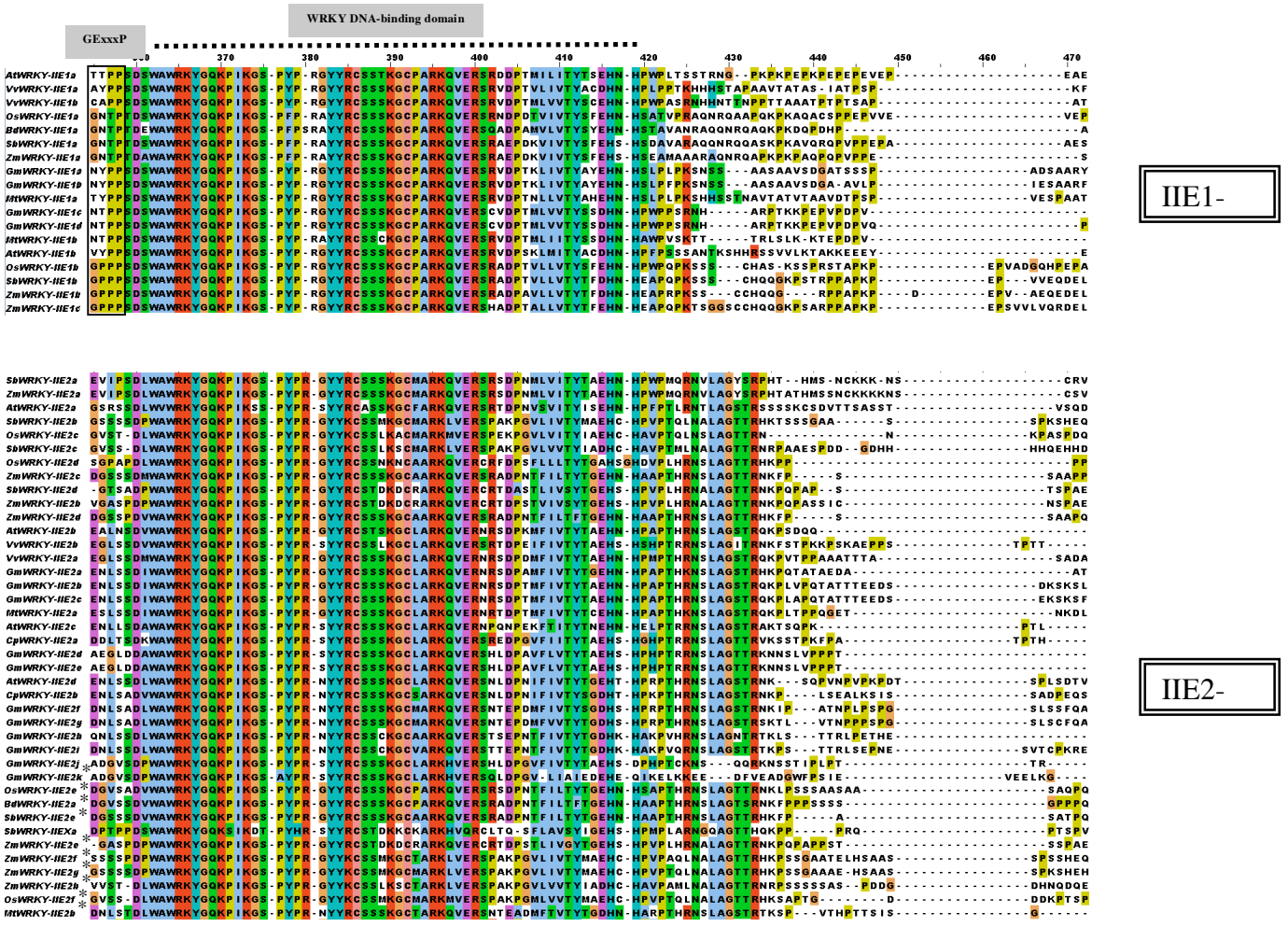


Appendix 6 continued

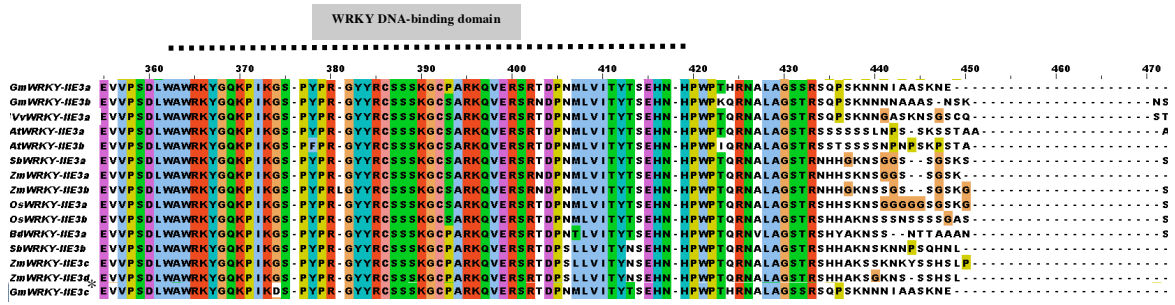


IIE3-

Appendix 6 continued

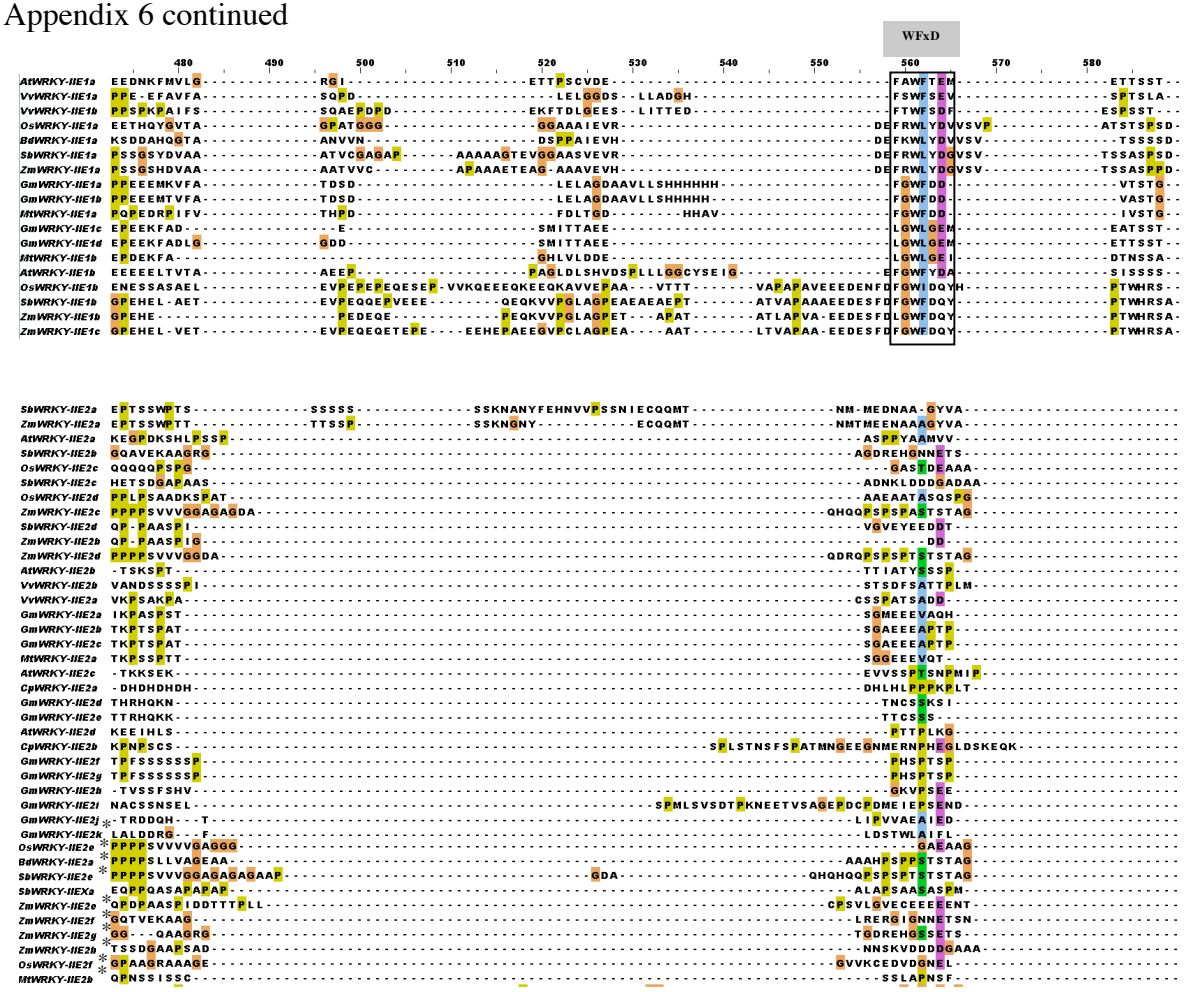


Appendix 6 continued

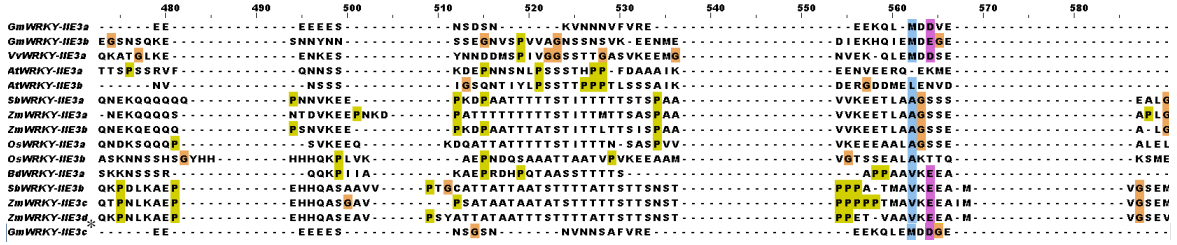


IIE3-

Appendix 6 continued

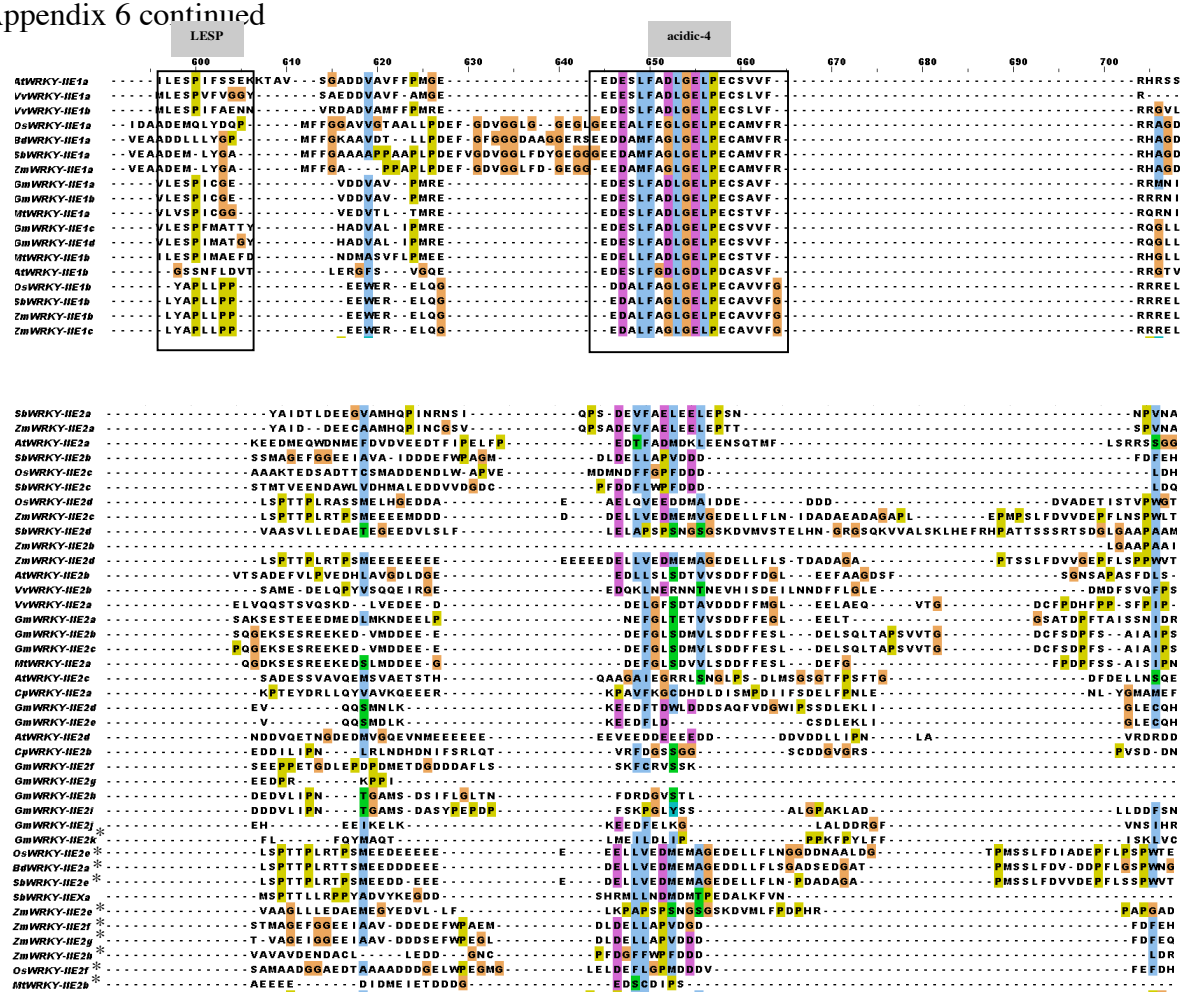


Appendix 6 continued

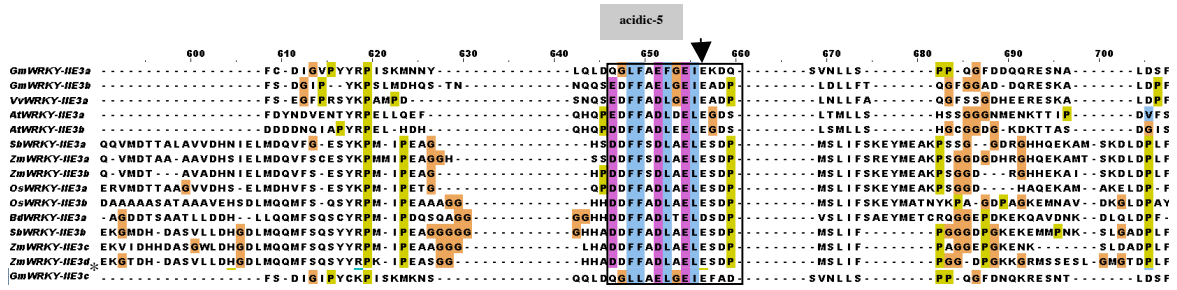


IIE3-

Appendix 6 continued

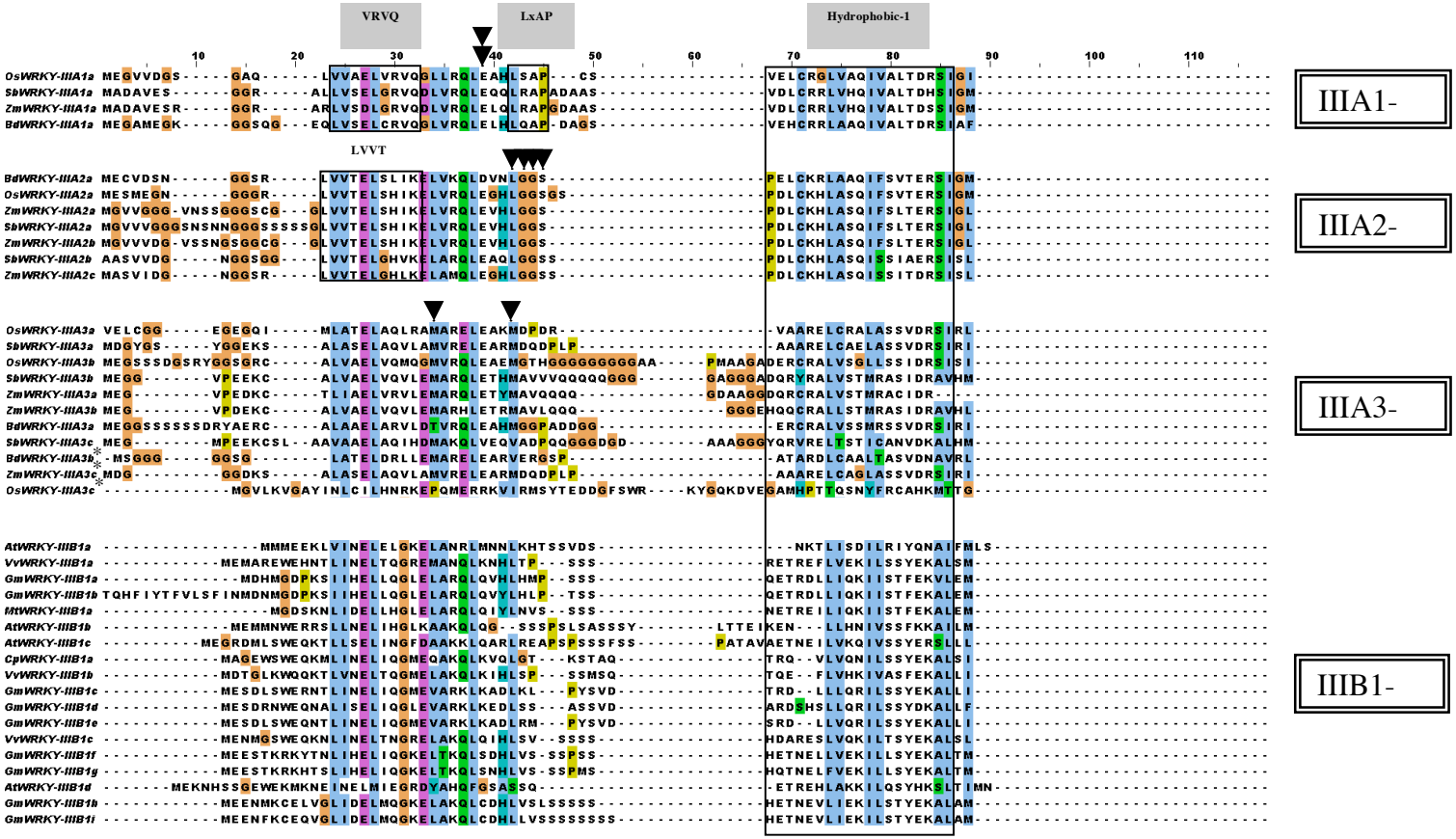


Appendix 6 continued

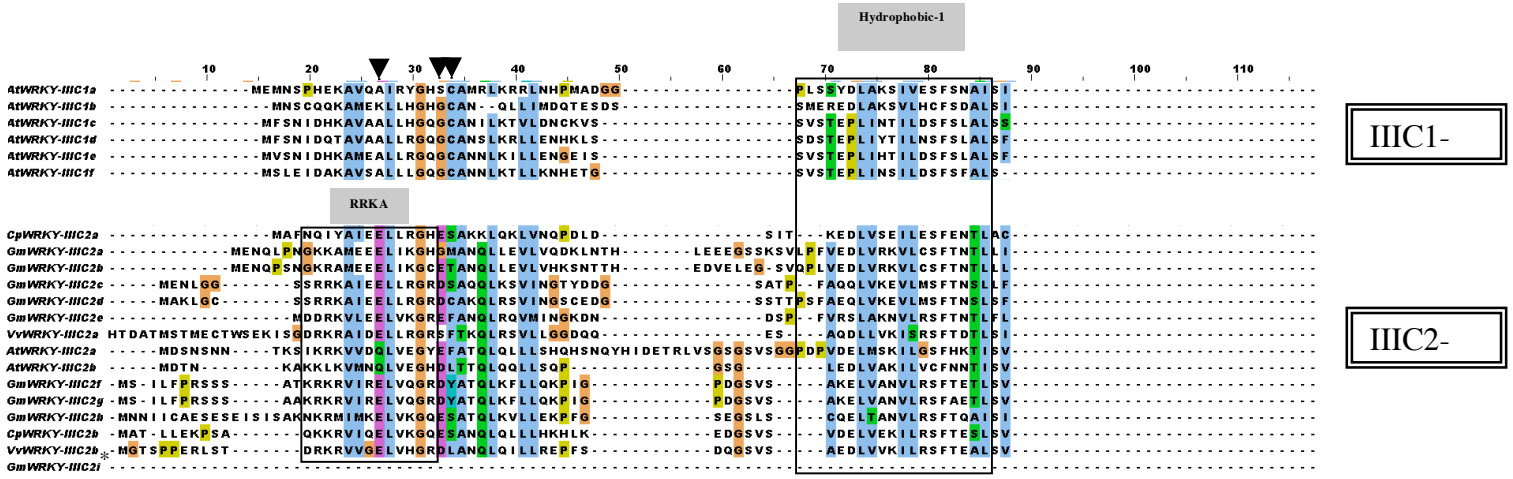


IIE3-

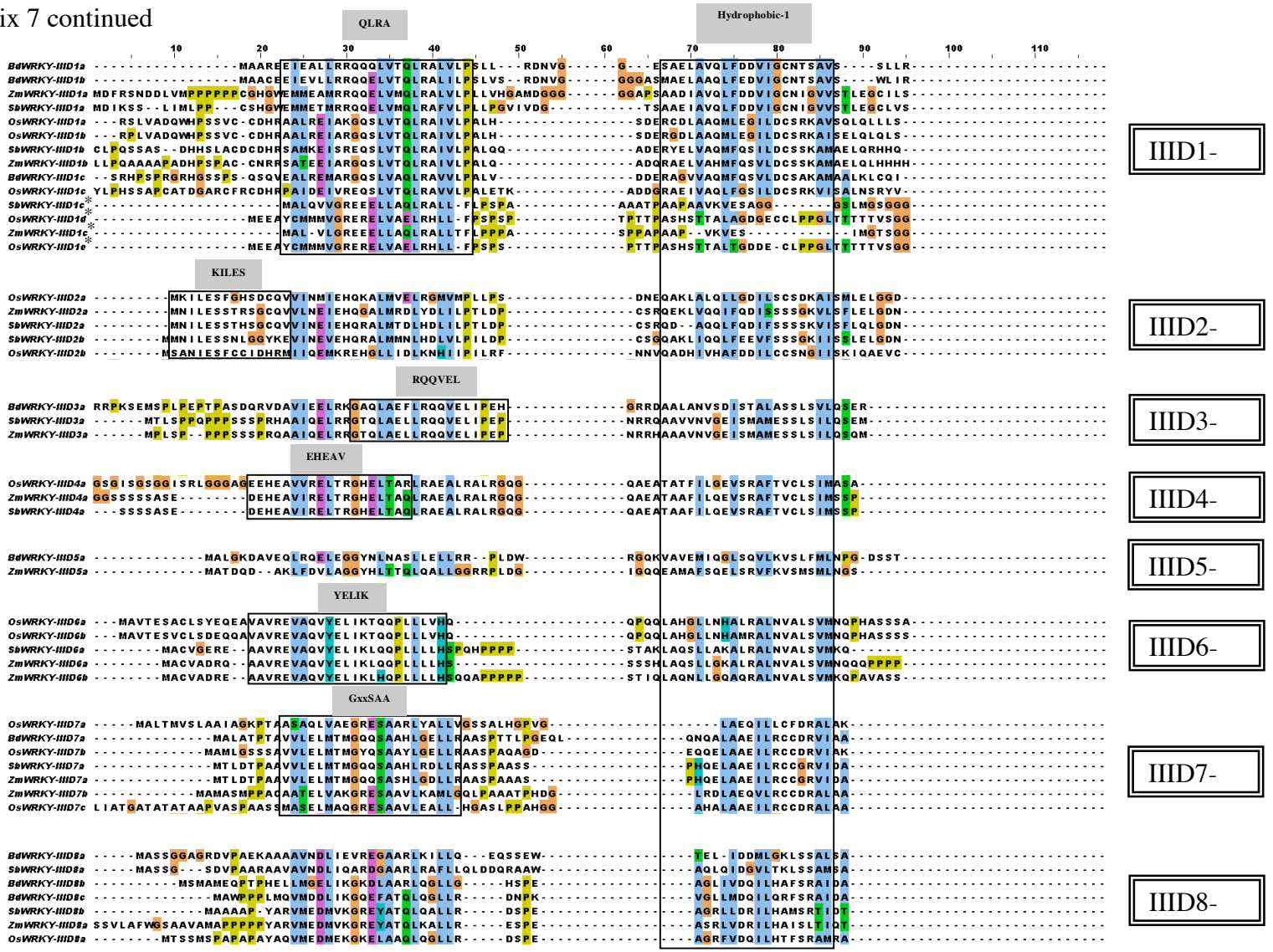
Appendix 7



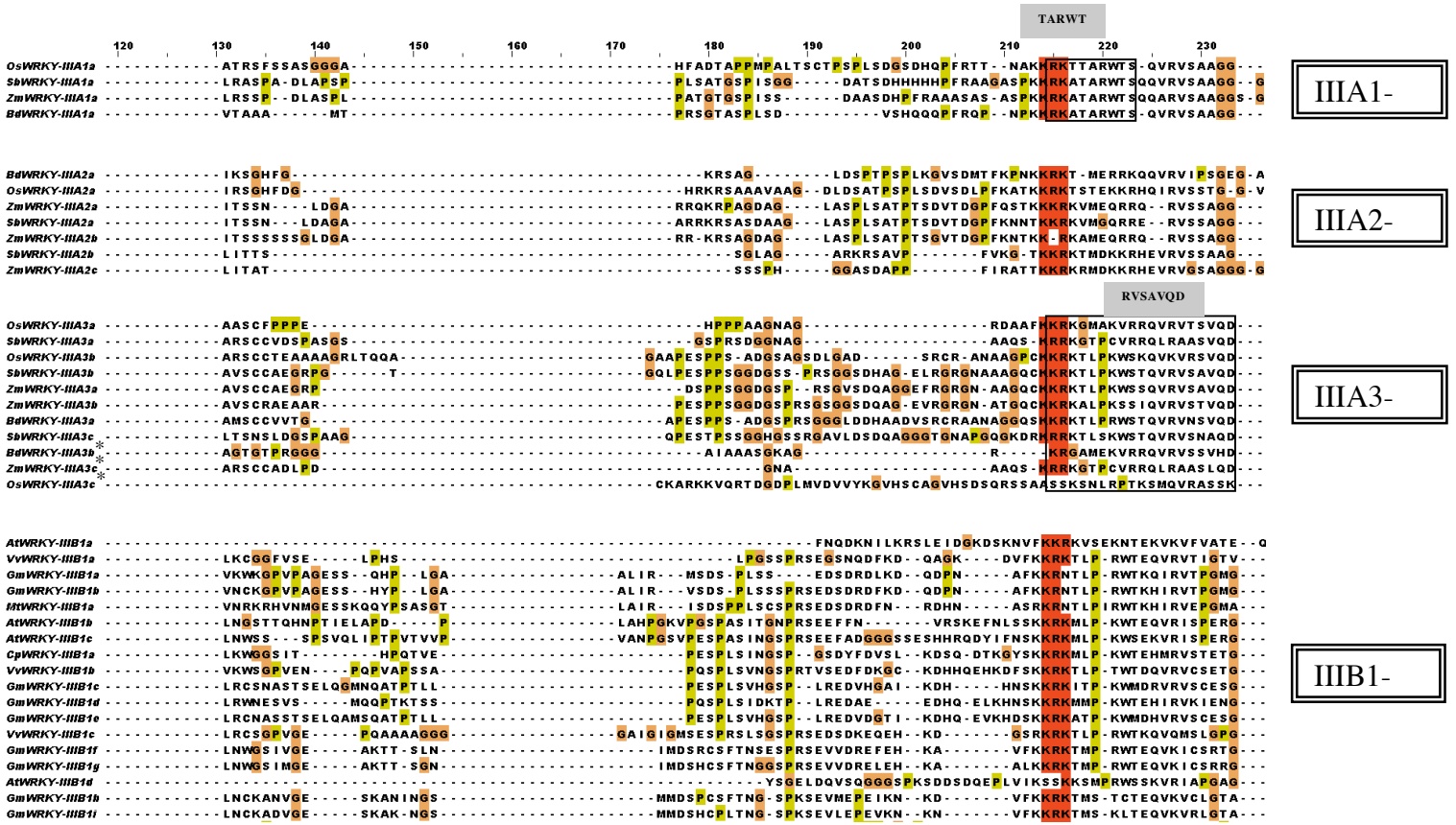
Appendix 7 continued



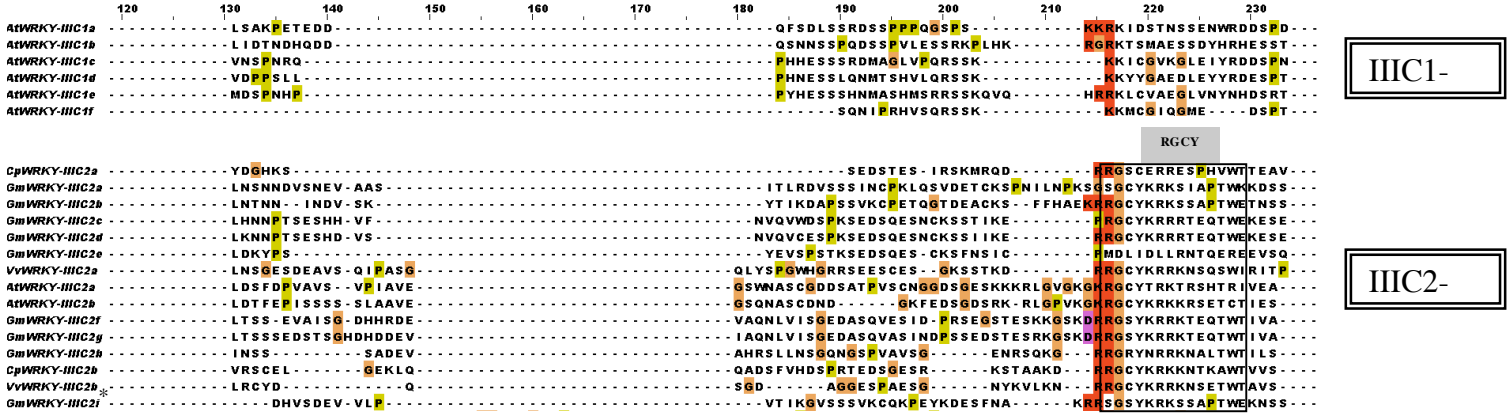
Appendix 7 continued



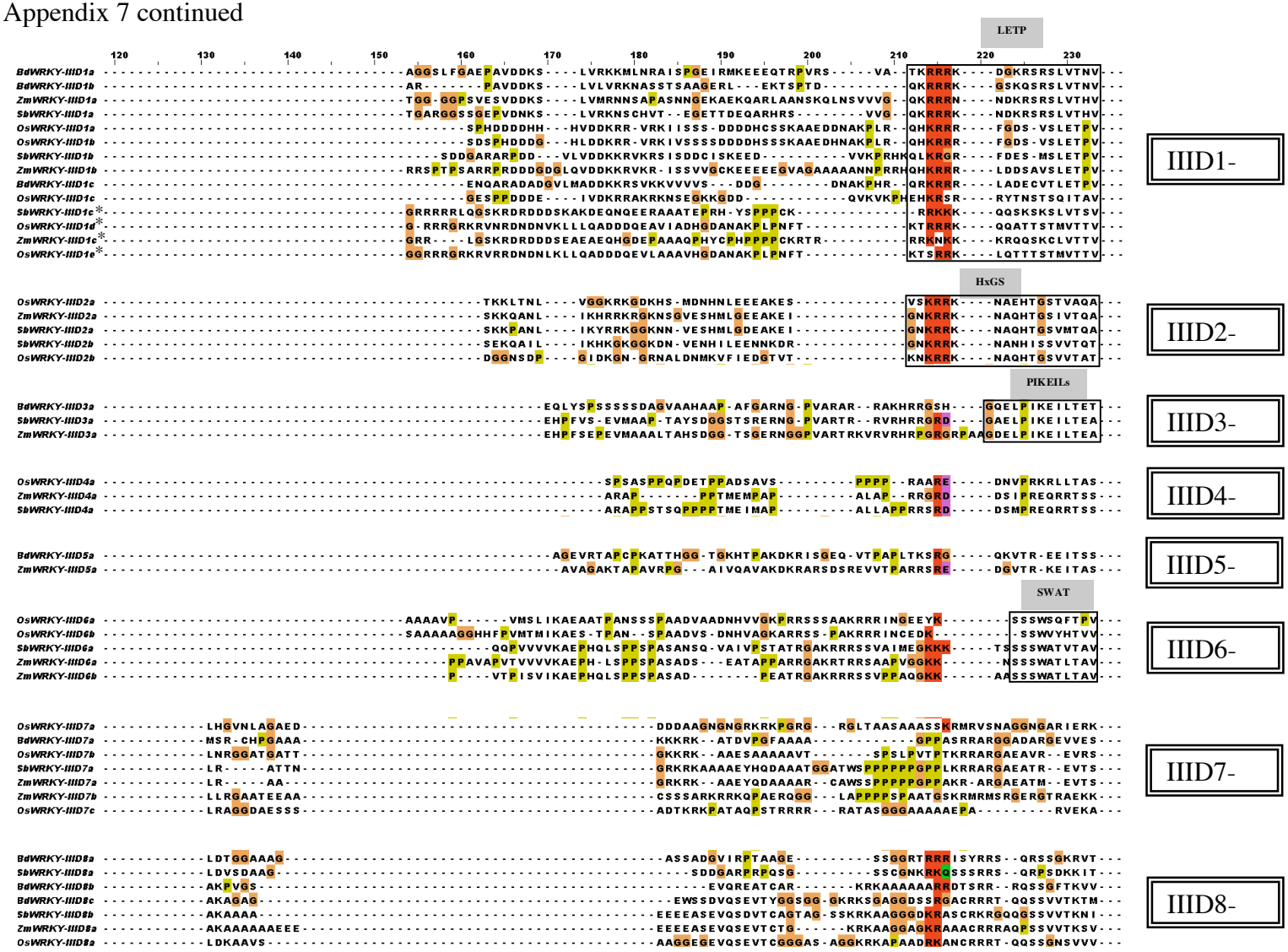
Appendix 7 continued



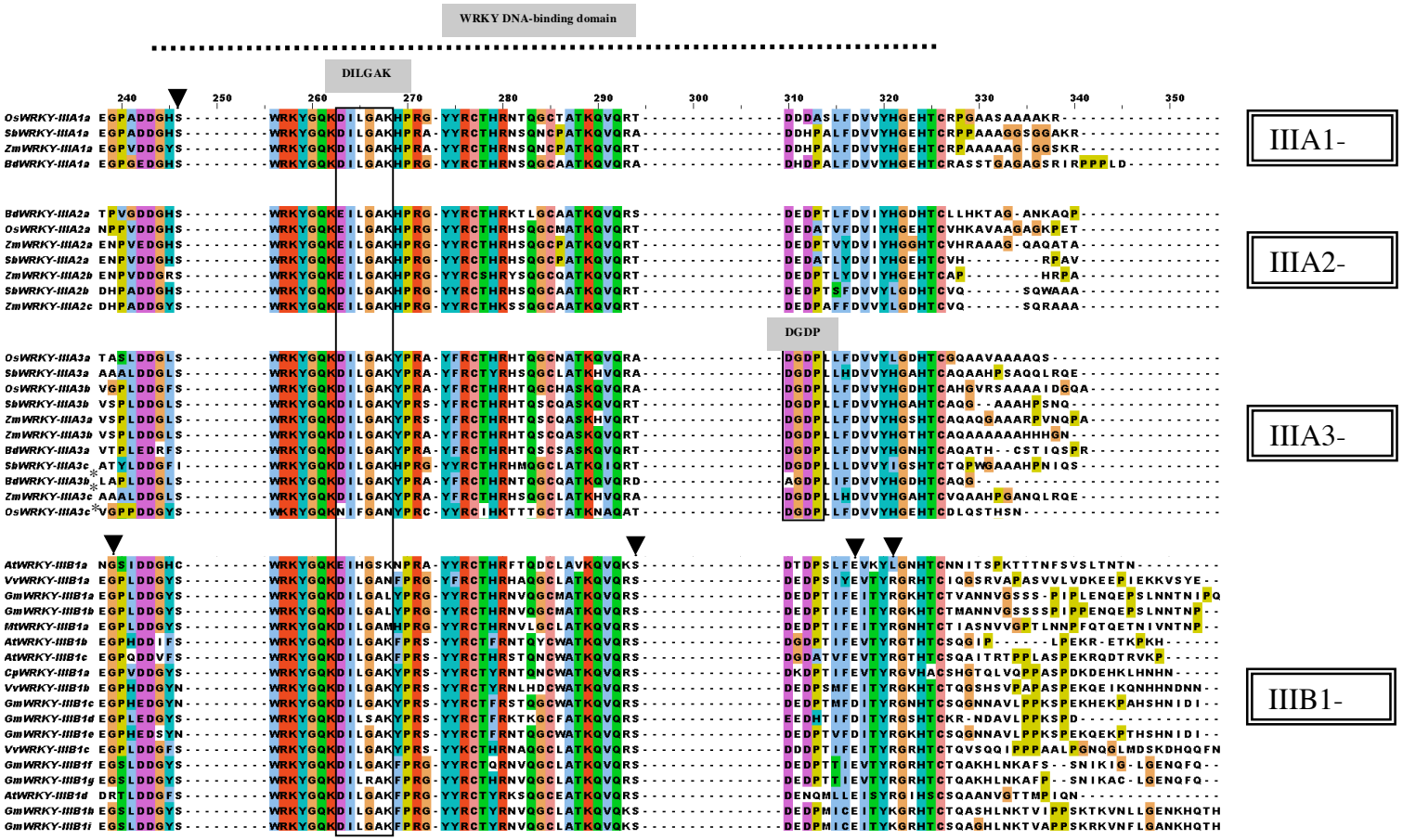
Appendix 7 continued



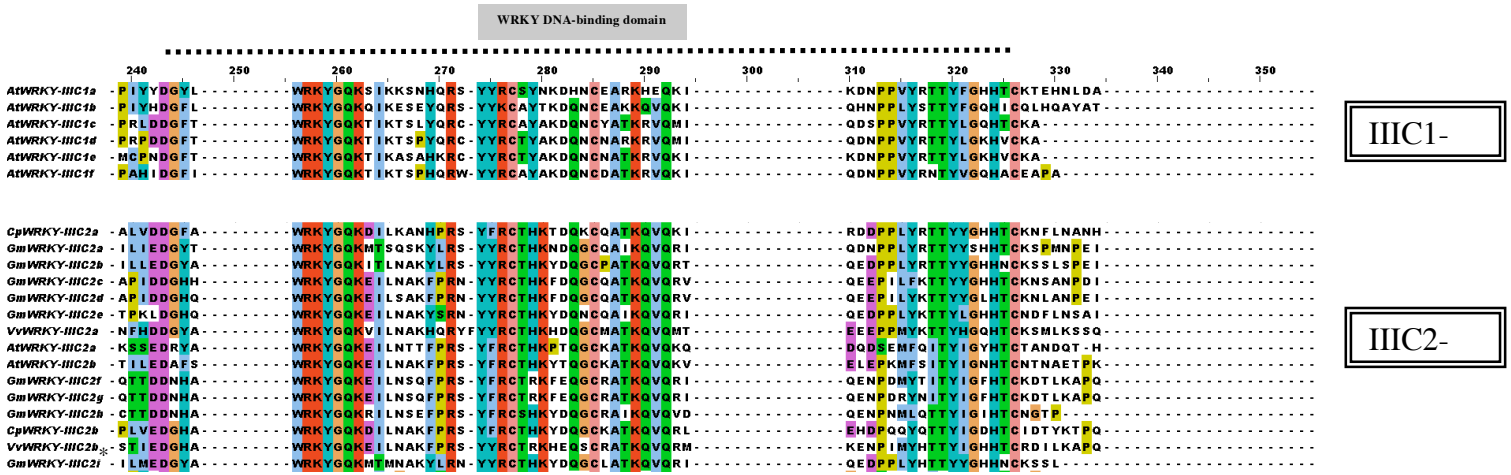
Appendix 7 continued



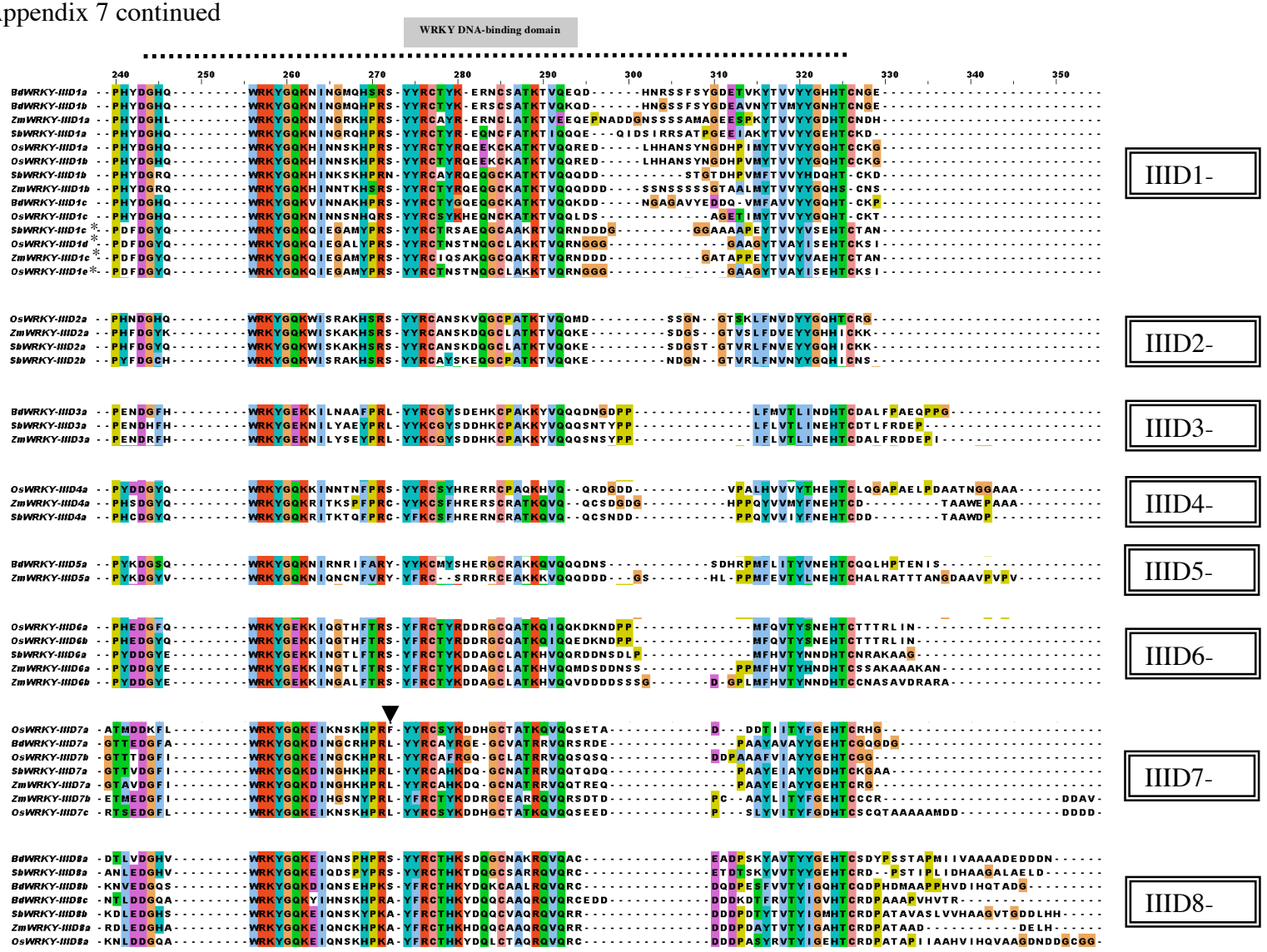
Appendix 7 continued



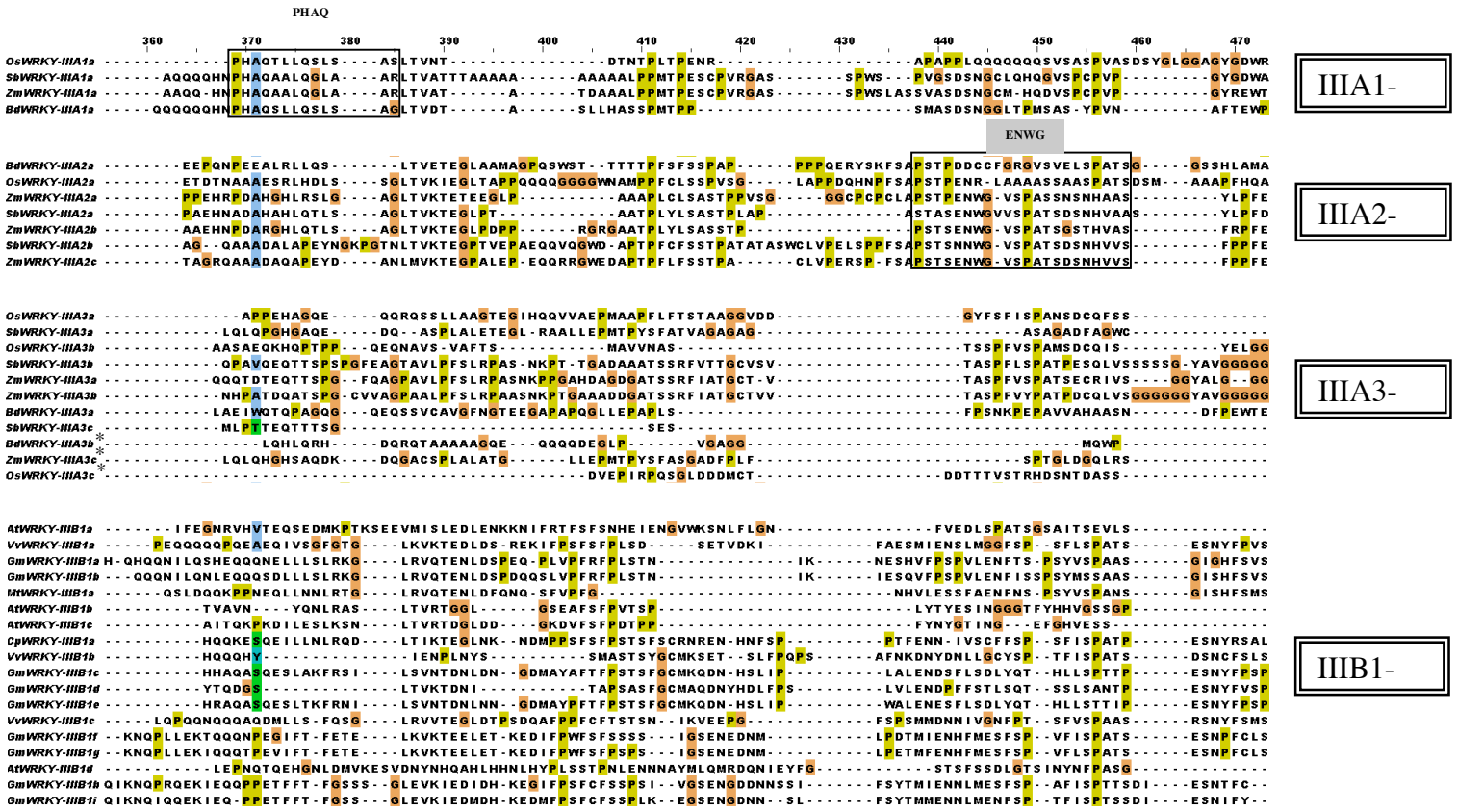
Appendix 7 continued



Appendix 7 continued

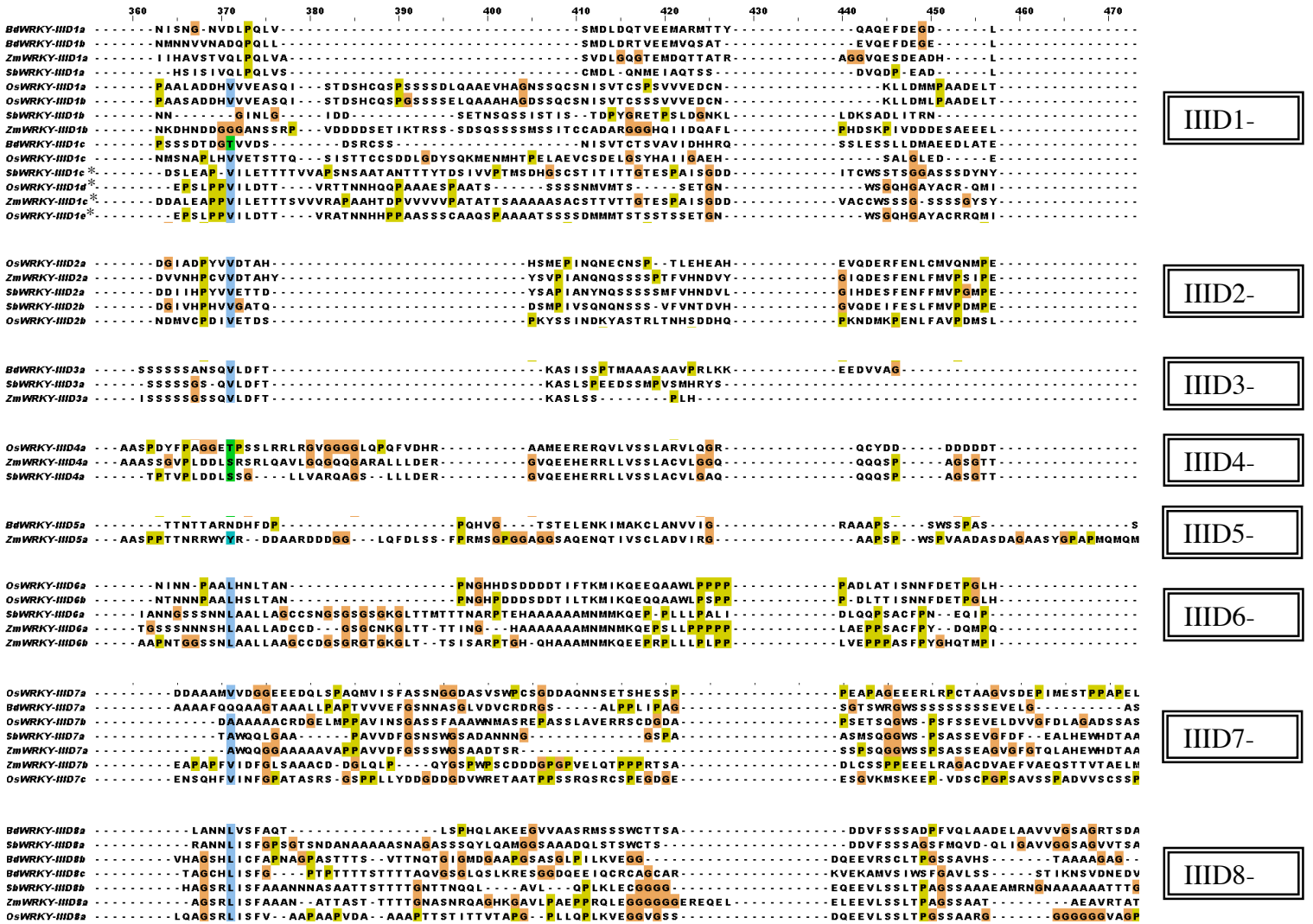


Appendix 7 continued

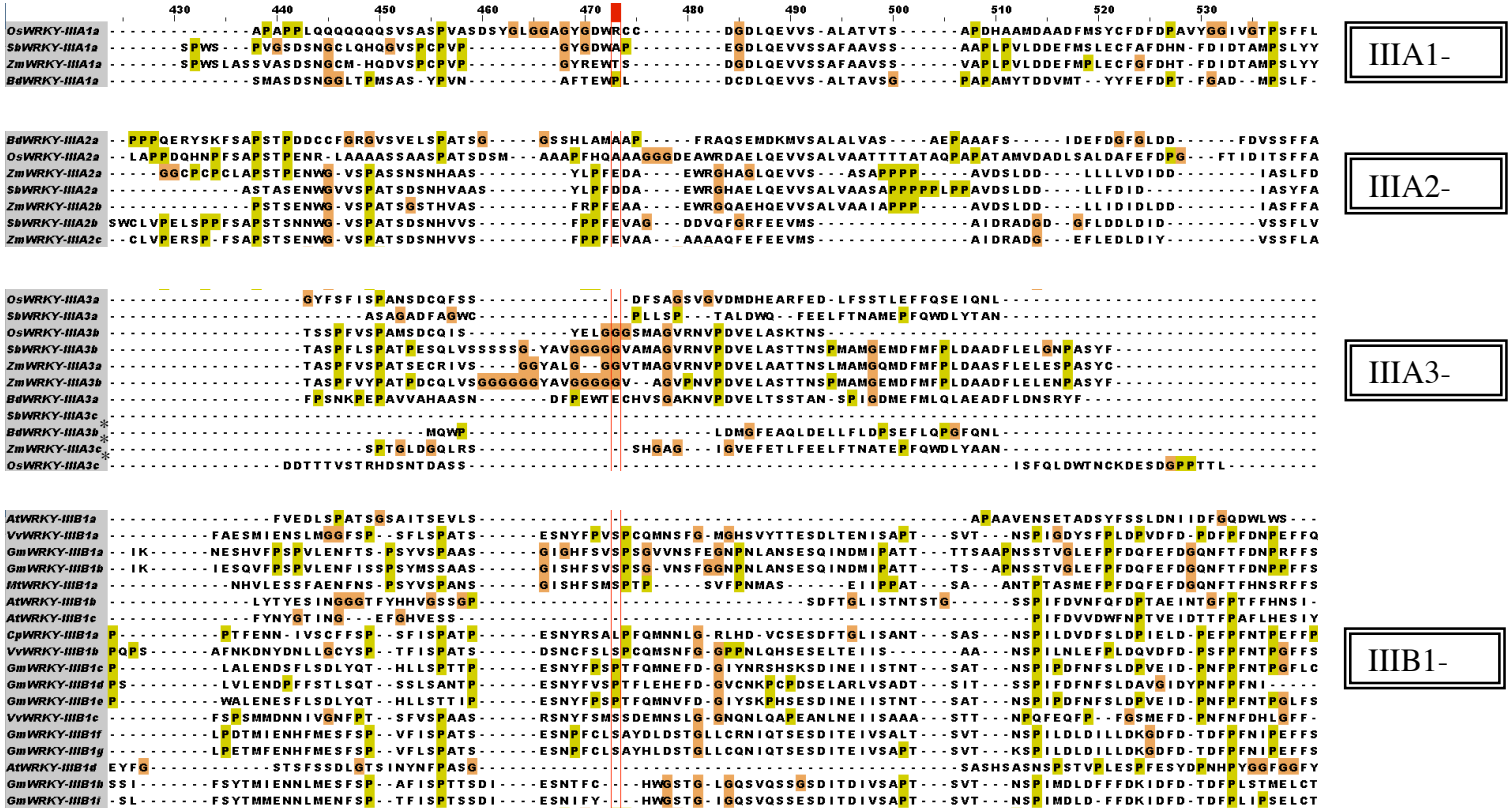


Appendix 7 continued

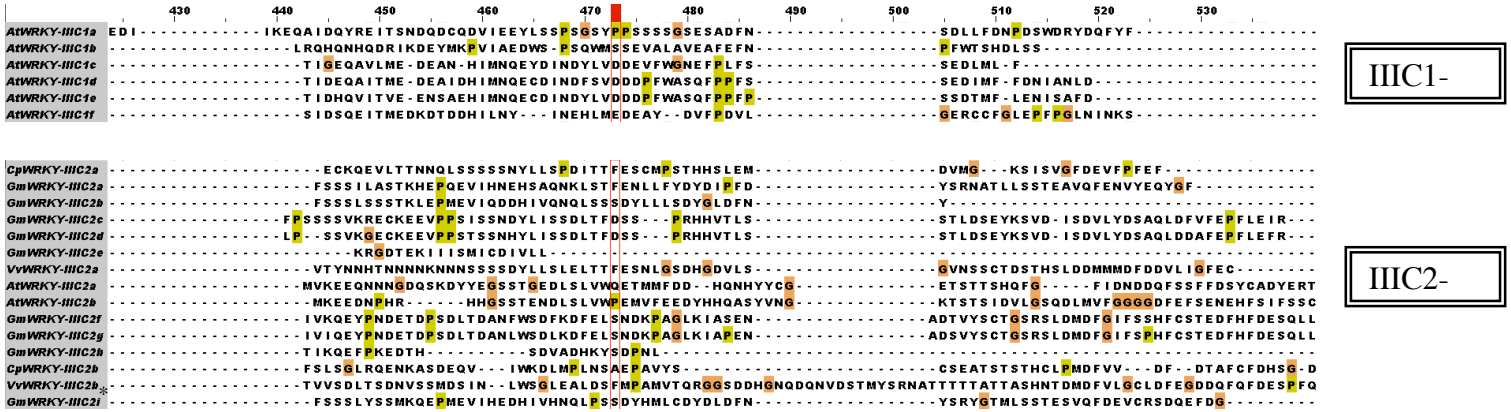
Appendix 7 continued



Appendix 7 continued



Appendix 7 continued



Appendix 7 continued