UC Riverside UC Riverside Electronic Theses and Dissertations

Title

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

Permalink https://escholarship.org/uc/item/9bs5965k

Author Atamian, Hagop Sarkis

Publication Date 2012

Supplemental Material https://escholarship.org/uc/item/9bs5965k#supplemental

Peer reviewed|Thesis/dissertation

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

Copyright by Hagop Sarkis Atamian 2012 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

vi

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

vii

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

TABLE OF CONTENTS

Acknowledgement	iv
Dedications	v
Abstract of Dissertation	vi
Table of Contents	ix
List of Tables	xii
List of Figures	XV
General Introduction	1
References	43
Chapter One: SlWRKY70 is required for Mi-1-mediated resistance to	
aphids and nematodes in tomato	71
Abstract	72
Introduction	73
Materials and Methods	79
Results	86
Discussion	91
References	97

Chapter Two: A Systemic Phylogenetics-based Nomenclature for

WRKYTranscription Factors	108
Abstract	109
Introduction	111
Materials and Methods	117
Results and discussion	119
References	135
Chapter Three: Sequencing and Comparative Analysis of the Potato	
Aphid Macrosiphum euphorbiae Transcriptome	155
Abstract	156
Introduction	157
Materials and Methods	160
Results	167
Discussion	174
References	181
Chapter Four: In Planta Expression or Delivery of Potato Aphid	
Macrosiphum euphorbiae Effectors Me10 and Me23 Enhances Aphid	
Fecundity	195
Abstract	196
Introduction	197
Materials and Methods	202
Results and discussion	208

References	219
General Conclusions	230
Reference	237
Appendix	239

List of Tables

Table	Page
Table 2.1 Conserved motifs defining all WRKY subgroups and types	
proposed in this study.	150
Table 3.1 Gene Onthology (GO) assignment for the clusters absent in aphids.	193
Table 3.2 Primers designed for sequencing Me_WB29764.	194
Table 4.1 Gateway primers for cloning in the expression vectors	
pEarleyGate100 and pVSP_PsSPdes.	228
Table 4.2 Primers used for gene expression analysis.	229
Table S1 Conserved motif sequences and revised putative full-length	
WRKY sequences of group I members.	

Table S2 Conserved motif sequences and revised putative full-lengthWRKY sequences of group IIA members.

Table S3 Conserved motif sequences and revised putative full-lengthWRKY sequences of group IIB members.

Table S4 Conserved motif sequences and revised putative full-lengthWRKY sequences of group IIC members.

Table S5 Conserved motif sequences and revised putative full-lengthWRKY sequences of group IID members.

Table S6 Conserved motif sequences and revised putative full-lengthWRKY sequences of group IIE members.

Table S7 Conserved motif sequences and revised putative full-lengthWRKY sequences of group III members.

Table S8 The sequences of conserved proteins missing in aphids.

 Table S9 The sequences of putative aphid-specific clusters.

 Table S10 The sequences of putative potato aphid-specific clusters.

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

Table S12 Annotation of *Macrosiphum euphorbiae* salivary gland

 contigs.

List of Figures

Figure	Page
Fig. 1.1 Phylogenetic tree of WRKY group III and amino acid sequence of <i>SI</i> WRKY70.	102
Fig. 1.2 TRV-based virus-induced gene silencing of tomato <i>WRKY70</i> and assessment of its role in <i>Mi-1</i> -mediated resistance against potato aphid and RKN	103
Fig. 1.3 Effect of TRV-VIGS on transcript level of <i>SlWRKY70</i> in control and silenced tomato cv. Motelle (<i>Mi-1/Mi-1</i>) leaves.	104
Fig. 1.4 <i>SIWRKY70</i> temporal expression in tomato leaves and roots after potato aphid infestation or RKN inoculation.	105
Fig. 1.5 Semi-quantitative RT-PCR analysis of <i>SlWRKY70</i> transcript levels after salicylic acid or methyl jasmonate hormone treatments	106
Fig. 1.6 <i>SlWRKY70</i> basal expression analysis in wild type cv. Castelmart (CM) and <i>jai1-1</i> mutant tomato plants.	107

Fig 2.1 Assignment of WRKY sequences from 15 plant species to	
groups/subgroups based on the Arabidopsis WRKY family	
classification.	140
Fig. 2.2 A phylogenetic tree constructed for group I from direct	
alignment of full length sequences	141
Fig. 2.3 Evolutionary relationships of putative group I WRKY	
transcription factors from 15 sequenced plant genomes inferred from	
Bayesian phylogenetic analysis.	142
Fig. 2.4 A phylogenetic tree of group I including secondarily assigned	
members.	143
Fig. 2.5 Evolutionary relationships of putative subgroup IIA WRKY	
transcription factors from 11 sequenced plant genomes inferred from	
Bayesian phylogenetic analysis.	144
Fig. 2.6 Evolutionary relationships of putative subgroup IIB WRKY	
transcription factors from 11 sequenced plant genomes inferred from	
Bayesian phylogenetic analysis	145

Fig. 2.7 Evolutionary relationships of putative subgroup IIC WRKY	
transcription factors from 11 sequenced plant genomes inferred from	
Bayesian phylogenetic analysis.	146
Fig. 2.8 Evolutionary relationships of putative subgroup IID WRKY	
transcription factors from 11 sequenced plant genomes inferred from	
Bayesian phylogenetic analysis	147
Fig. 2.9 Evolutionary relationships of putative subgroup IIE WRKY	
transcription factors from 11 sequenced plant genomes inferred from	
Bayesian phylogenetic analysis	148
Fig. 2.10 Evolutionary relationships of putative group III WRKY	
transcription factors from 11 sequenced plant genomes inferred from	
Bayesian phylogenetic analysis.	149
Fig. 3.1 Overview of <i>Macrosiphum euphorbiae</i> transcriptome assembly.	186
Fig. 3.2 Insect GO-slim terms associated with <i>Macrosiphum euphorbiae</i> .	187
Fig. 3.3 Schematics of the comparative transcriptome analysis.	188

Fig. 3.4 RNA blot analysis of the putative <i>Macrosiphum</i>	
euphorbiae-specific transcripts.	189
Fig. 3.5 Tissue-specific expression of putative Macrosiphum	
euphorbiae-specific transcripts.	190
Fig. 3.6 Schematic of <i>Me</i> _WB29764 transcript.	191
Fig. 3.7 DNA blot analysis of <i>Me</i> _WB29764.	192
Fig. 4.1 Classification of the Macrosiphum euphorbiae contigs using	
the Acyrthosiuphon pisum Gene Ontology (GO) terms.	223
Fig. 4.2 Tissue-specific expression analysis of the eight Macrosiphum	
euphorbiae candidate effectors.	224
Fig. 4.3 Myzus persicae performance on Nicotiana benthamiana plants	
expressing Macrosiphum euphorbiae candidate effectors.	225
Fig. 4.4 Macrosiphum euphorbiae performance on tomato plants	
expressing M. euphorbiae candidate effectors.	226

Fig. 4.5 Alignment of deduced amino acid sequences of Macrosiphumeuphorbiae (Me) effectors Me10 and Me23, with putative orthologsfrom Acyrthosiphon pisum (ACYP) and Myzus persicae (Mp).227

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 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 nearisogenic 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 coiledcoil (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 agedependent 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*-genemediated 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 helixloop-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 VpWRKY1 and V_p 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 AaWRKY1 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) GhWRKY3 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^{COII} 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

Acyrthosiphon 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-offunction 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 fastneutron 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 Ntransgenic N. benthamiana plants. The plants were assayed for attenuation of Nresistance 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 as well (Fradin et al. 2003), the latter being involved in *Ve*-mediated resistance as well (Fradin et al. 2003). 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 Vemediated 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 (*SISERK1*), 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 Gramnegative 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 TFs during *Mi-1.2*-mediated resistance against RKN and potato aphid. The WRKY TFs during *Mi-1.2*-mediated resistance against RKN and potato aphid. The WRKY TFs during *Mi-1.2*-mediated resistance against RKN and potato aphid. The WRKY TFs during *Mi-1.2*-mediated resistance against RKN and potato aphid. The WRKY TFs 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. Hypersensitivelike 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 sapsucking 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 pathogenassociated 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 Ptomediated 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*) FaWRKY1
 and Arabidopsis AtWRKY75 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. Virusinduced 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 leucinerich 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 siteleucine-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 mitogenactivated protein kinases *LeMPK1*, *LeMPK2*, and *LeMPK3* 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 *CrWRKY1* 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 *COII*-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-loophelix 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 genomewide 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 phloemfeeding 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 SIWRKY70 attenuated Mi-1-mediated resistance against both potato aphid and RKN showing that *SlWRKY70* is required for *Mi-1* function. Furthermore, we found SIWRKY70 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 *SIWRKY70* in tomato to assess its role in *Mi-1*-mediated resistance. Moreover, we profile the expression of *SIWRKY70* after potato aphid infestation, RKN inoculation, and treatments with SA or methyl jasmonate (MeJA) hormones. Our data implicate a role for *SIWRKY70* 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 *SIWRKY70* 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 19°C in growth chambers with a 16-h-light and 8-h-dark photoperiod and 200 μ mol 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 22-26°C and nematodes in a greenhouse at 23-30°C. 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 OD₆₀₀ 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 19°C 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 24°C in a growth chamber for three weeks and then transferred to a greenhouse at 23-28°C 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 *SIWRKY70* 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 -80°C 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 -80°C 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 μ g of RNA was treated with DNase I (New England Biolabs) followed by phenol-chloroform extraction and RNA precipitation with isopropanol. cDNA was synthesized from 5μ g 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'-

GTGTGGGGCTCACCTACGTTT-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 μ l with 1 μ l of template cDNA, 2.5 μ l of 10X PCR buffer, 2.5 μ l of 25 mM MgCl₂, 0.4 μ l of 25 mM dNTPs, 0.5 μ l of *Taq* DNA polymerase, and 10 μ M 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 μ l using 2.5 μ l 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, $\Delta\Delta$ Ct was calculated by subtracting Δ Ct (Ct _{target} – Ct _{Ub-3}) and the $\Delta\Delta$ Ct 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 SIWRKY70 is shown in figure 1b. As highlighted in this

figure *SI*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 *SI*WRKY70 suggest that both of these proteins act as nuclear localized transcriptional activators that bind to promoters via their WRKY domains.

SIWRKY70 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 TRV4 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 *SIWRKY70* also contributes to *Mi-1*-mediated resistance against RKN. In *SIWRKY70*-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 *SIWRKY70* 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 *SIWRKY70* transcript levels are clearly reduced in silenced leaflets compared to TRV2 empty vector agroinfiltrated control leaflets though at variable levels (Fig. 3).

SIWRKY70 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
SIWRKY70 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, *SIWRKY70* 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 *SIWRKY70* 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.

SIWRKY70 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 *SIWRKY70* 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 *SIWRKY70* 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, *SIWRKY70* transcript modulation in response to *Mi-1*-mediated recognition of aphids and RKN is enhanced and accelerated, respectively.

In Arabidopsis AtWRKY70 is required for full immunity mediated by the Rgene 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 Rmediated 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 SIWRKY70 transcript levels after potato aphid infestation in resistant plants and accelerated induction of SIWRKY70 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 *SI*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 *SIWRKY70* 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 JAdependent 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 *SlCOII*-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 SAdependent (Li et al. 2006b). Thus, *Mi-1*-mediated aphid recognition may activate WRKY70 function by triggering SA signaling. Besides having a direct role in Rmediated 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 defenserelated 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 *SIWRKY70* 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 *SIWRKY70* 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 an an an an area of the signal. Nevertheless 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), WRKY72type 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. *Coil*-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 Ptomediated 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 *S*/WRKY70. Features conserved between SIWRKY70 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 against provide 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 *SIWRKY70* 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 *SIWRKY70* 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.



Fig. 1.6 *SIWRKY70* 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 *SIWRKY70* 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} . ${}_{5}C_{x22-23}H_{x}H$ pattern of zinc ligands, while those of group III members contain the zincfinger motif $C_{x7}C_{x23}H_{x}C$. 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 Nterminal WRKY DNA-binding domain of group I members serve as specific DNAbinding 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 Nterminal 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 heterodimerization (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 AtWRKY1 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 AtWRKY1 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 AtWRKY1 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 typespecific 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 Ctermini whenever possible. This approach allowed for better-resolved trees and proper inference of orthology among each of the group 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 DNAbinding 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 DNAbinding 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 DNAbinding 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 DNAbinding 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 subclades 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. patens*. 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 DNAbinding domain (Appendix 5 indicated by an arrow). All the members of the other IIDtypes 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

130

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 IIE1type 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 prolinealanine-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.

132

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-glycineglycine-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 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 monocotdicot 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*) FaWRKY1
 and Arabidopsis AtWRKY75 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 VvWRKY11 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 moti	s defining all WRKV	subgroups and types	nronosed in this study
	is ucriming an WKKT	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][L F]	IC
charged-1	GD[ED]x[DE][DE]D[ED]P[ED]SKR[RW]KK[ED]xx	IA, IB1, IC1, IC3,
C	NE	IC4, ID, & IE
YKPxAK	[IV]YKP[QT]AKLVSK[TA]TVSLLANM[GL]N[CF][IA
	NS]T[NS][QS]QQ[PT][QL]Q[PS][AV]E[AT][RN]PQ	
	HS[NS][QH]DK[HF][RN]	
VASR	VASRP[TS]CS[ST]F[KR]SFSELLAGAIN[ATV][ST]	IA
	P[APS][IT]P[SC]	
Q-rich-1	[QM][GS][GPS]FGMSHQQALAQVTAQA[VA][HL	IB1 & IC1
	Q][SA][QPN][SL][RHN]M[FH]D[QH]	
P-rich-1	[PT]PPPPRP[TR][LI][ATS]LPPR[SP][ASF]AE[SA][F	IB1 & IC1
	L]F[TS][GAS][AG][GA][DGS]ASP	
FxxLLxG	[VF][AG]S[RS][PS][ST][AC][AS]D[FC][KR]SF[ST][IB1 & IC1
	QE]LLAGA[IML][AN]SP[PA][AP]x[PS]S	
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]	IC2
	S	
ENSS	[DT][SP][VA]ATPE[ND]SS[AIV][ST]FG[DE]D[ED][IC2
	FA][ED]	
charged-2	G[DG][ED][DF]D[EDG]DEP[DE][AS]KRW[KR][GK	IC2
		102
DxSS	F[ED]SP[DE][AG][VP][DE][AV][ST]S[TA][FLV]S[IC3
		102
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]	IC4
	[AD][SN][AM]P[AV][ED]V[DG][TLS][DS]E[LM][N	
	H]Q[IM][GN][SN]S[EA][NT][GA]LQ[AE]SQ[SV][E	
	D]	
ARYK	[AGS][GL][AGQ][GN][ASG]GARYK[LA]MSPA[KR	IC4
	JLPISR	
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][V	IIA (majority)
	A][VML]C[EA][NK]YNAL[QR][KS][HQ]Lx[ED][L	
	M][VM]x[KA][NA]	
alanin-rich-1	SPEI[RQ][QK][NV]L[VA][EQ][QE]MA[ST][ST]LT[IIA (majority)
	RK]DPNF[TK]AAL[AV][AT]A[IVL]SGR[IL]L	
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	IIA2
]KKP	
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[IIB (majority)
	RK]L[KR]xML[SD][QR]	
alanin-rich-2	DTV[ED][TA][AV]T[AK]AI[TA][AS]DP[NS]FT[AS]	IIB (majority)
	ALAAAI[TS]SIIG	
LDL-2	[MN]A[TS][IL]S[AST]SA[PS][FH]PT[IV]TLDLT	IIB (majority)
	[SQ]	
QSKF	LYNOSKFSGLO	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][S N]	IIB3
FLAR	LMAGSNFLARAV	IIB4
MMPLP	MMPLP[AH]F[ED][HL][GH][HN][HQ][QH][HP][LP Q][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][I LV]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]NSVCTSSA ATSFFTSI[SG]SOLISMSDAAT[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]S	IID4
	D[PQ][NAV][NS][AQ][QH][QH]	
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]V	IIE (all)
	VC[VIH][PV][AT][AP][ADE][ANV]	
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][IIE1
	ES][CD][PSA][MV][VS][FL][RI]	
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
DPFxxxxDP	[FG][GA]D[PA]F[SA]G[LM][VPR]DP[FL][ALS][SH	IIE3
F	T][DE]	
acidic-5	[GH][HQ][PSA][DE]DFF[AS]DLAELESDP[ML]SL[I	IIE3
	L]	
hydrophobic-	SS[ED]LAExL[VA]xK[IV]L[RS][SC]FE[KR][AS][LI	III
1]	
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[V	IIIA3
	A][SA][SA]VQD	
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]QL	IIID1
	RA[LI][VL][LF]	
LETP	[KQH][RKH][RK]R[RK][RLN][DF][KDG][RDE]S[R	IIID1
	V]SL[EV]T[PHN]VP	
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	IIID2
][AT]P	
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

158

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 24°C 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 22-26°C. 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 μ g 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 μ g of the DNase-treated total RNA using Seramag 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 1 μ l on 2% agaorse gel. To assess the diversity of the library, 1 μ l 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 transcript ome.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'-CTCACTGGGGGCA CCTAATGT-3') and *Me*_WB26189R1 (5'-GCAACAAGGGAACCAAGAG-3'), *Me*_WB28140F1 (5'-CAACATGCACATGTACACACCTC-3') and *Me*_WB28140F1 (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 MgCl₂,

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 μ g) 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 μ g of genomic DNA was digested with *Hind*III, *EcoRV* 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 ³²P 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'-CTCACTGGGGGCACCTAATGT-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 either approximate the sequence of the sequence 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 fulllength 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 ATPdependent 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. *Coil*-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 hostsymbiont 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 ($7 \mu g$) 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.

Та	ble 3.1. Gene Onthology (GO	D) assignment for the clusters
ab	ent 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
Me_WB10316R3	AGTAGCCGTTAGCTTCATTG	sequence Me_WB10316
Me_WB26189F3	GGTAATATCTGGTATTGTGGCTTG	
Me_WB26189R3	ACTGTCGTTGTGTTTCACAGTAAC	
Me_WB26189F4	AGGGAGTTGACCCCTCATTTG	
Me_WB26189R4	CGGCAAGAACCATCAATGAC	Primers to amplify and
Me_WB26189F5	GCATATATACGTCTTGTTACCCGATG	sequence Me_WB26189
Me_WB26189R5	ACTCCAGGCCTCTATGAAGCAG	
Me_WB26189F6	CGTTAATCTCACGGTGATAGGT	
Me_WB26189R6	CCGGCTGACAATTAGACCAC	
Me_WB28140F3	AGTCTATCTTTGAAGGGCTCACAC	
Me_WB28140R3	CAACCTGCTGACATTCCTAACG	
Me_WB28140F4	CTATACTTGCCGCACCTACTTG	Primers to amplify and
Me_WB28140R4	GCTGGTCGCCATCGTTAAG	sequence Me_WB28140
Me_WB28140F5	GAGTCGGACACGTGGTTACCT	
Me_WB28140R5	TGAAGAAGCCGGGAAGGTAG	
MeCpwF1	AACAAAGACTGCAAGAACAAATG	
MeCpwF2	GATAACACGTCCGAAAGTGAAG	
MeCpwF3	GAAGGCTGCTAAGCTCAACCT	Oligos used to sequence
MeCpwF4	AAAGATCAACTCTGGTGGTACTTG	the gaps between the
MeCpwF5	AAGTACTGCAACCACATTGCTG	contigs.
MeCpwF6	ATCAAAACCGTTGATGTGTGGT	
MeCpwF7	TCTTATTTTCTGCTGGTATCATTG	

CHAPTER FOUR

In planta expression or delivery of potato aphid Macrosiphum euphorbiae

effectors Me10 and Me23 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, Me10 and Me23, increased aphid fecundity suggesting their ability to suppress N. benthamiana defenses. Five of these candidate effectors, including Me10 and Me23, were also analyzed in tomato by delivering them through the Pseudomonas syringae type three secretion system. In tomato, only Me10 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 Mp10 and Mp42 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 MpC002 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 homologybased 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, Me10 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 μ mol 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 22-26°C 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 75nucleotide 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/ $\Delta avrPto/\Delta avrPtoB$ (*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 μ g/ml), gentamycin (14 μ g/ml) and kanamycin (50 μ g/ml) at 28°C. The cultures were diluted to a final OD₆₀₀ of 0.3 in an induction buffer (10 mM MES, 10 mM MgCl₂, 150 μ M 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 μ g/ml), carbenicillin (100 μ g/ml) and kanamycin (50 μ g/ml) at 30°C for two days. A single colony was inoculated into 200 μ l 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 MgCl₂. Whole plants were vacuum infiltrated with *Pst* (1 x 10³ CFU/ml) containing a candidate effector or GUS in 1 mM MgCl₂ 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 MgCl₂, 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 nonoverlapping 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 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, *Me13* and *Me17* (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 *Me5*, *Me10* and *Me23* 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 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 Me10 (P = 0.004) and Me23 (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 Me10 or Me23 was modest, this increase in aphid performance was consistent in three independent experiments. This suggests that Me10 and Me23 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, *Me*23, 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 DC3000AavrPto/AavrPtoB* (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 Me10 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 *Me23* compared to the GUS control suggesting that Me23 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.

Me10 and Me23 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 *Me10* 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-pathwayderived 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 PCRbased 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. *COII*: 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 *L*27 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. 4.5A-B Alignment of deduced amino acid sequences of *Macrosiphum euphorbiae* (*Me*) effectors Me10 (**A**) and Me23 (**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	g in the expression	vectors pEarleyGate1	00 and
pVSP_PsSPdes				

Contig	Forwrad primer (5'-3')	Reverse primer (5'-3')
Me5	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTAACAATGAATA	GCTGGGTTCAGTTTCCGGTGCAT
	ATCAGGAATTTGTTCATTT	AGGCGAATAAGACG
Me10	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTAGTGTACTAGA	GCTGGGTTTATGCTCCAACGACT
	AGGAACAATGCAATCAATA	GTTGGTTGGGAC
Me13	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTCGGACAACTAG	GCTGGGTTTATCCGTTGTTGGTC
	TAACAATGAGTTTGTTTCC	GAAGATACATTTTT
Me14	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTTGTTCCAAACA	GCTGGGTATATTACATAGGTAGT
	ATGCAGTTTCTAATAT	AATATATAAGCTTA
Me17	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTACCAGCAGGTG	GCTGGGTGAGTTGGTGTATAACG
	TGCAACAATGGAACCGAC	TCAACG
Me20	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTATAACAATGTG	GCTGGGTATCGGTTCGTACTTCC
	GCCGGGATTCACCGTGGTC	TGGGATTTTGAGTG
Me23	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTAACAATGAGGG	GCTGGGTGTTTACGTAACATAAT
	TTTCCCAAGGATTTCCCTA	TAGCAACATTGGTCC
Me25	GGGGACAAGTTTGTACAAAA	GGGGACCACTTTGTACAAGAAA
	AAGCAGGCTGCAACAATGTC	GCTGGGTTCAGAATAGAGCTAG
	ACCAAGTCACGGAGAAAACG	AGTAGCGAGTTTG
GUS	GGGGACAAGTTTGTACAAAA	CACCACTTTGTACAAGAAAGCTG
	AAGCAGGCTCCATGGTCCGT	GGTCTTATTGTTTGCCTCCCTGCT
	CCTGTAGAAAC	

Contig	Forwrad primer (5'-3')	Reverse primer (5'-3')
Me5	CATTCAAGGCCTAGACAGG	TCTTCTCAAGTCGTTTGTCCC
	AC	
Me10	CAAGACCAAAAATAAGTT	GACATTGCGTTAAGAAGATCG
	GTTCCT	
Me13	GAAGTCAATTGGCCATTAA	CCATTTCTGCATTATTTCTGG
	ACAA	
Me14	GTCATTCGAGAGCGTACCA	CAAAAGGACTCGATCGGTTG
	Т	
Me17	GTTTGACGAGATGGACTTG	GTATCTCAACCTTCTTTTCAGCC
	CTTA	
Me20	GAACCGAAGACGACCTCA	CAAGTCTCCATGGGCATCG
	TTCT	
Me23	GGAAATTTGTGTAAATGGC	CATTGGTCCTTTAACAGTTCTTG
	AGA	
Me25	CATGGACAACTAACACAC	GAGCTAGAGTAGCGAGTTTGCA
	AATTC	AC
L27	CCGAAAAGCTGTCATAATG	GGTGAAACCTTGTCTACTGTTAC
	AAGAC	ATCTT
MeC002	GAGCAGGAAGAAGCGTCT	CTTGGTGGGAGCATTGGTTA
	GT	
MeSuc	GAGATCGATCCTATTTATG	CATTCCATTCCCACGGAGATC
	GC	

Table 4.2. Primers used for gene expression analysis.

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 *SIWRKY70* 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 *SIWRKY70* 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 aphidspecific 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 phloemfeeding 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

VYWRKY-IC15





LxER

	10	20	30	40	50	60	70	80	90	100	110
AtWRKY-IC32	EEKSSKRVLER				ELSL	NH <mark>G</mark> QVI <mark>G</mark> LEE	DTSSNHN	· · · · KDSSQS	NVFR <mark>GGL</mark> SER	IAAR#	GFN
AtWRKY-IC3b	HRKRSKAEL <mark>G</mark> R · · · · ·				••••EFSL	NY I K • • • • NE	DSLQTTF···	QESS	· · · RGALRER	I AARS	G F N
GmWRKY-IC3a	EDMVELGIPFR · · · · ·					D	IASLTEQ · · ·	к <mark>р</mark> ss	RGGLVER	MAAR#	GFN
GmWRKY-IC3b	EDSVTRS I SEHS <mark>G</mark> SGRT	GDLFS <mark>G</mark> HRE	PSET <mark>G</mark> KE -		NMNDRAQ	D <mark>G</mark> DS <mark>G</mark> AQLTD	VSFQTEQ · · ·	···· кsss··	R <mark>GGL</mark> VER	MAAR#	G F N
WRKY-IC3a	DDVDSRATSE <mark>P</mark> TNENRN	I E <mark>G</mark> F F L <mark>G</mark> S <mark>P</mark> E	QIMS <mark>G</mark>		DT DKKT L	A <mark>G</mark> DL <mark>G</mark> DQLTE	F <mark>G</mark> AFSDQ	KLSS	RGGLVER	IAARA	GFN
HWRKY-IC3a	DD INSSTITE PPGENRT	K <mark>GLFLGQ</mark> PE	QMTT <mark>G</mark>		NAEKKD <mark>g</mark>	ART S G AQLTE	L <mark>G</mark> SFSEQ · · ·	KSNS	R <mark>GGL</mark> VER	MAARA	GFN
CpWRKY-IC3a	EDMSSKFMSEAL <mark>P</mark> DNKT	EGLFLGSRE	QMRS <mark>G</mark>		· · · · ETHKKS ·	L R E M <mark>G</mark> D Q S A E	LGLISEQ	···· кsss	RGGLGER	IAARA	GFN
BdWRKY-IC3a	DQFPPDPLAEDGS - GNE	AGFEKHGLS	VAVGSP		Q D	E G K P A Q L T P H	FGGRMS	<mark>G</mark> SSS		M QARN	GFS
DsWRKY-IC3a	QNFLADAF PPPELLEGE	GGFEKHGLS	VAVGSPP-		<mark>р</mark> т <mark>рррр</mark> ед	GCSPLPLTPQ	FGQKFG····	sgggg-	GG S L AD R	RARG	GFSN
WRKY-IC3a	RGVRPDTVGGHSNEDAM	PGFEKHGF S	VDISSP		QE	EGRSLPLTPO	FG-9KT		· · · · · · SLAER	M QARA	GFK
mWRKY-IC3a	RAVRPDTVEGHSNEDVE	AGFEKHOLS	AALNSP			EGRSLPLTPQ	FG . QNS	s <mark>pg</mark> s		M QARA	GFK
dWRKY-IC3b	EDFSSGPFSNLFSENGS	NKPHDHSEP	RG-EFVD-		L R D Q V P A	OSAEATLOKD	ISLEPNL	FNANGKPNP		MASRA	GFS····
SWRKY-JC3b	EEFSSGPFSDIFCDNGS	NKHODGLG	SK-AFID-		· · · · SSREETA	QLAKK	· · FESNL · · ·	FGANOKSSS		MAART	GFG····
dWRKY-IC3c	EDLSSGPCSDVFGDN	- KPODGIDO	AKPSLVV-				· · FEPNL · · ·	FGAKEKPIS	GSSLAER	MAPGN	GLC
WRKY-IC38*	FFFSSFPFSGFFSFHGT	NKPHDOSER	SR-EVVN-		SSEEVPA	HAVNDPFOKG	FSLK <mark>P</mark> NL · · ·	FSANHKSNS		RAARA	GFS
WRKY-IC3b							S	<mark>.</mark>	· · · · · S LAE 8	RAAK F	GFDAS
mWRKY.IC3e	MEODNINNN							NEVSDS	RGRKRSLAFE		GENS
WRKY-IC3d	· · MEODNINNNN · · · ·						• NCDFF • • • •	TEVEDS	RETKRSLAFE		GENS
NWRKY.IC1a							. DAKDP	KEVDES	PKPSIAFP		GEN
WRKY.IC3b	NKNGGGEINNYS						. YSKE	NEEDES	PKRSLAFE		GEN
MARKY-IC26	MGINBGAAAS						VKPESPN		APGIVETVER		GENNKK
WRICH-ICSN	MONNIPAARVET					1.1.140/0.11	LETKEEN	COVCC	NNDCICICE	BAAKO	CENNK .
	CONNEAAETRI	NW <mark>R</mark> oporre				or a p <mark>p</mark> a copp	LEIKSSN		NNKOTSTALK		G FRMR

AtWRKY-IC4a		MN <mark>P</mark> QANDRKEFQ <mark>G</mark> DCS	· · · · · · · · AT GDLT AKHDSAGGNGG · · · · · · · · · · · · · · · · ·
PtWRKY-IC4a		MDN NT SR	· · · · · · · · · · · · · · · · · · ·
PtWRKY-IC4b		MDN I N S R	· · · · · · · · · · · · · · · · · · ·
GmWRKY-IC4a		•••• MDAA•••••• ATNS <mark>GEPRP</mark> SSE••••••	· · · · · · · · LOTT GESEDPNRSGSGO · · · · · · · · · · · · · · · · · · ·
GmWRKY-IC4b		MDGTTNSGFPRSSSV	
MWRKY-IC4a	·		GEMENHSEDPNRTGSSSRPDSESDGPT
GmWRKY.IC4c			
GmMRKY-ICAd			Provence of the second se
0111111111111111111			
ADDIN 10 AL			FOTBTBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
MWRKY-IC4b		•••• MEVT•••••• TTNSG••••••	••••••••••••••••••••••••••••••••••••••
MWRKY-IC4b VvWRKY-IC4a		MEVTTTNS <mark>G</mark> MEES-SFQ <mark>P</mark> HNSSKE	E G T P T D P N R N G S V Q
MtWRKY-IC4b VvWRKY-IC4a BdWRKY-IC4a	MSI		E GT PT DPNRNGSVQ AT SL PPLHSSNVASGSD SS PEKPCPADRRVAALTG
MtWRKY-IC4a VvWRKY-IC4a BdWRKY-IC4a OsWRKY-IC4a	MSI MAI	MEVTTTNSG MEES-SFQPHNSSKE DSPNPGSCDLPGGAG DSPNPSSCDHPAGVG	E G T P T D P NRNG S VQ
MtWRKY-IC4b VvWRKY-IC4a BdWRKY-IC4a OsWRKY-IC4a SbWRKY-IC4a		. ME VT	E GT PT DE NENGSVO A TSI PPE INS NVAS QS D SS PE KEC P ADRIVAALT Q GS PE KOP VDRIVAALA QAA QS TE KEV VARVAALA QAA
IRTWRKY-IC4b VVWRKY-IC4a BdWRKY-IC4a OsWRKY-IC4a SbWRKY-IC4a ZmWRKY-IC4a	. MS 	MEVT	E GT PT T DPMRNG SV0. AT SL PP LHS SNVAS GS D. SS PE KC CP A DR VVA LT G. GS PE KO P V DR RVAA LT G. GS PE KO P V DR RVAA LAGGAA. GS TE K PV LADR RVAA LAGGAA.
httwRkY-1C4b VvWRkY-1C4a BdWRkY-1C4a OsWRkY-1C4a SbWRkY-1C4a ZmWRkY-1C4a ZmWRkY-1C4a	- MS - MA - MA - MA	. ME VT	E GT # T DE NRWG SVO. A T S L PE LISS NVA SG S D. S S # EKD C P ADRR VAAL T G. G S E KO P V DRR VAAL A G. G S E K V LADRR VAAL A G. G S # EK # V LADRR VAAL A G.

GmWRKY-ID1a	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 - MV	S S I	KSA	D	 	 	
GmWRKY-ID1b	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 - MV	S T I	OKSAI	D	 	 	
MIWRKY-ID1a	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 - MV	LSO	ESA	D	 	 	
MWRKY-ID16*	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 - MV	LSO	ONVI	D	 	 	
VyWRKY-ID1a*	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 MG	T S			 	 	
PfWRKY-ID1a	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 . MV	5 5 1	EVP		 	 	
PHWRKY-ID46	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 . MV				 	 	
ANNOLO IN		 	 			 		 	 	 		 	 	 			 				 	 	
Alward 1-1019	 	 	 • •	 	 	 • •	 	 - MAI			A	 	 	 									
CRWKKY-ID12	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 	 - M V 🛛		1 S Y S	1	 	 	

BdWRKY-IE1a	 																						 	 	 	 	 							·E	PK	EM	ME	тD	AS	ΕV	GK	AE	PK	EM	MET	A					 		 	-
AtWRKY-IE1a	 • •		• •	 ·										• •	-	• •							 • •	 • •	 	 	 				• •		• •	• •	• •						• •		· M	EЕ	DT	3 I D)		• •	• •	 		 	
CpWRKY-IE1a	 			 ·		• •		• • •	• • •	• •	• •					• •			• •				 	 • •	 	 	 						- 1	LF	AR	RK	LV	MA	ΕA	ET	CK	AV	QV	ER	REI	RHE	(- -	• • •			 • •	• •	 	-
GmWRKY-IE1a	 • •	• •	• •	 ·					• • •					• •		• •							 	 • •	 	 	 							• •								- E	PD	тκ	GNI	DSH	4		• •		 		 	-
GmWRKY-IE1b	 			 -					• • •							• •							 	 • •	 	 	 							• •	• •							- E	S D	тι	GNI	DSH	4	• • •			 • •		 	-
VvWRKY-IE1b	 • •	• •	• •	 •		• •		• • •	• • •	• •				• •		• •			• •				 	 • •	 	 	 	• •			• •	• •		• •	• •						• •		• •		• • •	• • •		• • •	• •		 • •	• •	 • •	-
VvWRKY-IE1a	 • •	• •	• •	 •		• •		• • •	• • •	• •	• •		• •	• •		• •	• •	• •	• •	• •	• •		 	 • •	 	 	 	• •			• •	• •		• •	• •			• •			• •	• •		• •	• • •			• • •	• •		 • •	• •	 • •	-
MtWRKY-IE1a*	 	• •		 •	• • •	• •	• • •			• •	• •	• •	• •				• •	• •	• •	• •	• •	• •	 	 	 • •	 	 		• •	• •		• •			• •		• •	• •			• •	- V	GP	тα	LPI	EC	8				 • •	• •	 • •	-
GmWRKY-IE1c*	 	• •	• •	 •		• •				• •	• •			• •					• •				 	 	 	 	 				• •				• •						• •	• •	• •		• • •	ME			• •		 • •	• •	 • •	-
GIIIWRKY-JE18*	 			 •		• •				• •	• •								• •				 	 	 	 	 																			· MP					 		 	-
PTWRR Y-IE12"	 			 •		• •				• •						• •	• •	• •	• •				 	 	 	 	 							• •						- 0	GD		GE	EE	LES	3 U Y	/				 		 	•



ID1-

IC3-

IC4-





















IC2-

SPTTG

 256
 376
 326
 326
 460
 410
 420
 420
 440
 450
 460
 470

 SPTTGKFFFLFVNGN
 ALSSE
 KAKADE
 FFDDIGAS
 FSFHVSRSSSFOTTERMS-VDVG
 RSSSHQAEEVKGSENI

 SPTTGKFFFLFVNGNMRHSELSSDAFE
 KAKADE
 FFDDIGAS
 FSFHVSRSSSFOTTERMS-VDVG
 RSSSHQAEEVKGSENI

 SPTTGKFFFNVGGNMRHSELSSDAFE
 KAKAD
 FDDITTSSFAKRATOSGSSFV
 AGGRKM
 INT
 LPOGSLFQEVSAGSENSFQC

 SPTTGKFFFNVGSNNRSSALLFAAPD
 KKCHN
 FFDDITSSFAKRATOSGS FVHAGRKM
 INT
 LPOGSLFQEVSAGSENSFQC

 SPTTGKFFFNGSSNRSTASSSHD
 FFDDITSSFAKRATOSGS
 FVHAGRKM
 INT
 LPOGSLFQEVSAGSENSFQC

 SPTTGKFFFNGSSNRSTASSKL
 KAKDE
 FSKED
 FFDDITSSFAKRATOSGS
 FVHAGRKM
 INT

 SPTTGKFFFNGSSNRSTASSKL
 NSKED
 FFDDITSSFAKRATOSGS
 SSTSTGKFFFLGTNVGASKLSA
 A
 IPOGSTGFFGUSSKSSFFGS

 SPTTGKFFFNGSSNRSTUSSEVD
 NSKED
 SSTGKFFFGUSSKNSTASSKLFSA
 A
 IPOGSTFSHOVSWKETNE
 A

 SPTTGKFFFUSSSKNSTUSSEVD
 NSKED
 ISSTAFFAROSSSFFLGGTNK
 OSSTSFFLGGTNK
 IPOGSTFSHEVSVOSENSLGS

AtWRKY-IC32 AtWRKY-IC36 GmWRKY-IC3 GmWRKY-IC3 VvWRKY-IC3a PtWRKY-IC3a C#WRKY-IC3# . BdWRKY-IC3a OSWRKY-IC3a ShWRKY-IC3a ZmWRKY-IC3a BdWRKY-IC3b OSWRKY-IC3b BdWRKY-IC3c STITGTLEFLMANANALAL EVINKU. EDLS. KUU UT LÄNKET I INKOLOSSATUR I SVUNKU DESLUSSSIAIK. STITGTLEFLMANNOKSI IPÄÄKIT – EDSÄ. FONDESSON HUSESATURS. TAEKOYGAYQQ. OSISI HOQE SLUSSSIAIK. STITGTEVISI IHEGSVNSVAT VD. GOQAN. NFS. ASGKKK HAN. FISLA (FSMGFEFAT. EFSTEASKKNYVFDSATD. VG. SFFTLALI SHEGSNLTTEQGN. VDVST. AFDV. AASKK KROHDINELPPEASLNQVSSH. HSVKGENRESHLLAQVQPLD. S. SFFTLALF, SHEGSNLTTEQGN. VDVST. ASDV. DASKK KROHDINELPPEASLNQVSSN. HSVKGENRESHLLAQVQPLD. SHWRKY.ICTH VVWRKY-IC3b GmWRKY-IC3 GmWRKY-IC3d LEMINGG HRDGGMLVOG GG VUVUT A SIN KONNEN DVKMVDO LAGVGKUVUT A SIN KONNEN DVKMVDO LINI VUVUT A MIWRKY-IC32 MIWRKY-IC3b PtWRKY-IC3b PtWRKY-IC3c PTTGTLPFLMASNAKSTIPSATKMD.....EDCT.......FGNDTFSFQPHVGSRRPNFS.AAEKGPNACHQ.......QSLSNIHQRESSLQSSFTAVK...... OSWRKY-IC3 AtWRKY-IC4a PtWRKY-IC4a PIWRKY-IC46 GmWRKY-IC4 TGSLRKLQQTVHGSMASAASAT - KV<mark>P</mark>ADFNN<mark>H</mark>VSKQSTQVE<mark>G</mark> - - - - - - - - -GmWRKY-IC4 MWRKY-IC4 GmWRKY-IC4 GmWRKY-IC4 MWRKY-IC46 VvWRKY-IC4a RAMPKY.ICA TTGSLGMAAIMHKSAHPDIL · · PTTGTLSMAAIMNKSANPDIL--OsWRKY-IC4 SE LE SERAA MENSAN DE LE CONTRACTOR DE L ShWRKY.ICA ZmWRKY-IC4 ZmWRKY-IC4 VVWRKY-IC48

GmWRKY-ID1b			·····
MANUTACY ID.4.	NAK		NEWEDED AMAGDI VE
INCOMERCY - ID 12	••••••••••••••••••••••••••••••••••••••		NEVSDSDAMACDLVS
MtWRKY-ID16*	••••••••••••••••••••••••••••••••••••••	·····	••••••••••••••••••••••••••••••••••••••
VvWRKY-ID1a'	• • • • • • • • • • • • • • • • • • • •	••••••••••••••••••••••••••••••••••••••	
PtWRKY-ID1a	ASQ	QEGIPAS	YAPGKS-LENSGGRIS
PtWRKY-ID1b		QEGIPPS	
AtWRKY-ID1a	••••••••••••••••••••••••••••••••••••••		·····
CpWRKY-ID1a	••••••ENH•••••••	••••••••••••••••••••••••••••••••••••••	·····

-	•		:	;	:	1		-	•	1	E	T	P	D	ĸ			. E / E	P	D	Q	P G	н •	т ч			-	-	-	-	;	-			
															-	sv	x	xS	5																
-					-		-		s	P	т	s	s	A	Y	S E	. 0	L	P	s	P	F	P	H											
•	-	•	-	•						P	s	L	A	A	s	s c) s	L	т	v	т	P	c	L٤	;						-				
•	-	•	-	-	-	-	-	•	-	ı	C	т	т	s	L	SE	1	s	P	т	s	v	т	QS	ş.	-	•	•	•	•	-	•			_

MVAVPCFIAPPGFAGQFAMSHQAALASVTAQAHMQLQ	· · · · · · · · · · · · · · · · · · ·
NVEDVLEVNETDSVKETVVSAIV <mark>P</mark> VDEVEENRQVE <mark>T</mark> S	
SLRDEERAREIHQEDLAGATPQ	
KAESKETVGSPEKEIIDRGAVQPLQTQTEDRLQVS	ACST <mark>P</mark> LSDLS <mark>P</mark> TSVTQS
KAESKETVESPEKETIERGAVQPLQTQTENRLQVS	
MSHQEVLASVT AKAAQAPVQMDQP	
	LNST SQS LAS - VE FKS
KSE I KET VGPP ENEKE VAKSEAAQ · · · · · · · · · · · · · · · · EP SAQAE YQ I QVT · ·	· · · · · · · · · · · · · · · · · · ·
KAESKVSAQPPDEETIEKDAVEAPQKQTENQLQS	VCSTSLSELSPTSVSHS
KAESKVSAQPPDEETVVKDAVEAPKKQTENRLQS	FCSTSLSGVSPTSVSHS
EAEYKEQVGLSHQEVLGSVADKNSGAQTQNQLQPS	
	NVAVE CFIAPEGFAGGAANSHQAA LASVTAQAHBQLQ NVEDVLEVNETDSVKETVVSAIVE VDEVEENROVETS SINDEERRAFIH GELAGATFQ VDEVEENROVETS KAESKETVGSE HLIDRAVQ HLIDRAVQ KAESKETVGSE KILDRAVQ HLIDRAVQ KAESKETVGSE KILDRAVQ HLIDRAVQ KAESKETVGSE KILDRAVQ HLIDRAVQ KAESKETVGSPE KILDRAVQ HLIDRAVQ KAESKETVGSPE KILDRAVQ HLIDRAVQ KAESKETVGSPE KILDRAVQ HLIDRAVQ KAESKEVGOP KILDRAVQ KAAQAPQOPO KAESKYSQOP ELIKETVGPEKEVAKSEAAQ SAQAEVQOVT KAESKYSQOP ELIKAVSE ARKATENKQOV KAESKYSQOP ELIKAVSVSAQAP ARKATENKQOV KAESKYSQOP ELIVKDAVE ARKATENKQOV

IE1-

IC3-

IC4-

ID1-

C,

	480	490	500	510	520	530	540	550	560	570	580	
		1	· · · ·	- <u>-</u>		1	· · · · ·	I I	· · · · ·	1 1		
AtWRKY-IA12			QSEAVSY	<mark>6</mark> KSVS0	G		· · · · · · · THRA <mark>G</mark>	PNLVQK	••••••••••••••••••••••••••••••••••••••	ES	· · · · · · · · · · · · · · · · · · ·	•
PtWRKY-IA1a			LQ <mark>PV</mark> ETR	· · PQLSKO	DKH		••••NFSSQLT	SNLHQN · · ·	· · · · · · I PSP/	EA-DHTTE	<mark>, LRLTSLNQEED</mark> KTI	L
PtWRKY-IA1b*			• • • LQS <mark>V</mark> EAR • • •	<mark>Р</mark> ас <mark>р</mark> ка	NKQ		•••••NFS <mark>S</mark> QLT	SNLHQN	• • • • • • • • • • • • • • • • • • •	· · · · · · H/	VILTSQNQE - DPKAI	· [
VvWRKY-IA1a			• • • LAQVEAR • • •	VQ <mark>PP</mark> NQ	DRQ		H S R <mark>P</mark> H L S	SNLHQT	F <mark>P</mark> SQE	ET - DRTSE	SKTASQNLEEDQK <mark>P</mark> I	1
GmWRKY-IA1a			QQ <mark>P</mark> MEAN	FQHSIH	E • • • • • • • • • •		KFRTNTS	SNLDQS	••••••••••••••••••••••••••••••••••••••	EINYQSSE	SKMVQQNIEEDQKVI	L .
GmWRKY-IA1b			L V <mark>M</mark> N <mark>P</mark> I	FNQ <mark>P</mark> QD	s		•••••ALLT••R	ARISIR···	VLPR-		QKQ I I RAVNL DQKAL	1
MWRKY-IA1a			QKSRETN	•• <mark>P</mark> QHLNH	D · · · · · · · · · · ·		· · · · · NFRANMS	SKLHQN · · ·	••••••••••••••••••••••••••••••••••••••	ET - YQATES	CMMA <mark>p</mark> qn i eedqkal	1 I
CpWRKY-IA1a	• • • • • • • • • • • • • • • • • • • •		•••MQL <mark>A</mark> HAR•••	AQNSNO	DKH • • • • • • • •		•••••NFMSRFC	SNMHQC · · ·	••••• <mark>Р</mark> SHT	EA-ELTIGN	ISKTAAQNLEEDQKTS	;



<i>GmWRKY-IB1a*</i> · · · · · · · · ANS · · · · · NMH I QAEHS · · · T QASAAT FNT T QQL · · · · · · · I <mark>PP</mark> L NADSWAT MT E S · · · · · · · ·	ADHS-HSEQRLQSS
<i>GmWRKY-IB1b</i> * ANS NMH I QAEHS T QA <mark>P</mark> AAASNT T QQL M <mark>PP</mark> L T S D SWAAMT E S	IDHS-HSEQRLQSS
<i>PtWRKY-IB1</i> 2* · · · · · · · · · LAQS · · · · QMHM <mark>H</mark> AQYQ · · · <mark>P</mark> SSLTA <mark>P</mark> TELLTRH <mark>P</mark> SF · · · · · · · NPGEALQQQQ · MPHSTS · · · · · · D	T QNSVVELTEFSHSE RKYQ P
<i>MtWRKY-IB1a*</i>	TDHA <mark>P</mark> LSEQRLQ <mark>P</mark> S··········
GmWRKY-IB1¢* • • • • • • • • • • • • • • • • • • •	L DYS - HSEQKLQSS
<i>OSWRKY-IB1a</i> [★] ····································	<mark>T</mark> L <mark>P</mark> SHT DN <u>SN</u> I E ST E V SH <mark>G</mark> FQT

IB2-

IC1-

IB1-

AtWRKY-182a BdWRKY-182a SbWRKY-182a ZmWRKY-182a	N D - D	SQPQPQRPDTF PYQPPSTTYFNHADP PYPPYLVADQ PYAPYLHVDQ	PHHMPPSTSVAVHGR PPQWPPRAPSSSLPPP PPQWPPRADSSM PPQWPPPRATVVGSSL	QSLDVSQVD RGNNFALQHEQ PHSNFTVFPR PHSNFTVLFPS	ORARNHYNN <mark>FG</mark> NNNNNRSYNV QOQOSSMOMOLLRALGR <mark>P</mark> HQALASA N <mark>F</mark> YDHDMOLRATALFGGSSGLAHALPP N <mark>F</mark> YEHDMOLRATALFGGGGGSYSYTLPP
AtWRKY-IC1a		GNN VHMQQ SQQSE	Y <mark>P</mark> SSTQQQQQQQQASLTE	· · · I <mark>P</mark> SFSSA <mark>P</mark> RSQIRAS · · · · VQE	r s q <mark>g</mark> - q r e t s e i s v f e h r s q
AtWRKY-IC1b		AN ANMQ P QTE	Y <mark>PPP</mark> SQ	· · · VQSFSS <mark>G</mark> · QAQ PTS · · · · AP	- · A · QRETSDVT I I EHRSQQ · · · · · · · · · · · · · · · · · ·
PtWRKY-IC1a		ANS NMHVQQEY	· · · ST SAMSSTQFLTSINNS · · · · ·	AAQQQQQQQMA <mark>G</mark> SVT	• • • • • • D SRVTVQELS <mark>G</mark> <mark>P</mark> HADR • • • • • • • I RSES
PtWRKY-IC1b		ANS IMHVQ PEY	• • • ST <mark>P</mark> AMSST • • FTSTQ <mark>G</mark> • • • • • • •	AHQQQQKVRSVA	· · · · · D <mark>S</mark> RVK I QE LSDFSRSD · · · · · · · QRSES
PtWRKY-IC1c		FAQS QMHMQAQYQ	• • • <mark>P</mark> SSVTAAKELLTQY <mark>P</mark> SF • • • • • •	· · · N <mark>PG</mark> EALQQQQL · M <mark>PP</mark> STS · · · ·	D <mark>A</mark> QNSMVE <mark>P</mark> AEFSHSE
PtWRKY-IC1d*		LAQSQMHMHAQYQ	· · · <mark>P</mark> SSLTA <mark>P</mark> TELLTRH <mark>P</mark> SF · · · · · ·	· · · N <mark>PG</mark> EALQQQQQ · M <mark>P</mark> HSTS · · · ·	D <mark>T</mark> QNSVVELTEFSHSE
GmWRKY-IC1a		AQS · · · · · · HMHMQADYQ	M <mark>P</mark> SVTA <mark>P</mark> TE <mark>PP</mark> VQQLSF	ALNEASEQQVVSCVS	· · · · · E <mark>P</mark> RNAQ LEA <mark>P</mark> ELSQAD · · · · · · · · KKYQ <mark>P</mark>
GmWRKY-IC1b		AQS HMHMQADYQ	• • • M <mark>P</mark> AVTA <mark>P</mark> TE <mark>PP</mark> VRQLSF • • • • • •	ALNEASEQQVVSCVSSVS	· · · · · E <mark>P</mark> RNAQLEA <mark>P</mark> ELSQAD · · · · · · · · KKYQ <mark>P</mark>
MWRKY-IC1a		AQSQ NMHMQPEYQ	· · · LVSYEA <mark>P</mark> TERLAEQ <mark>P</mark> SY · · · · ·	· · · TRNEA <mark>P</mark> EQQVTA <mark>P</mark> VS · · · · · · ·	· · · · · E <mark>P</mark> RNAQMETSEITHSD · · · · · · · · KKYQ <mark>P</mark>
VvWRKY-IC1a		S-QSHMFIQAEYQ	PSSLEAPVESLAQDPSF	· · · · ITDATTHQQ · · V <mark>PP</mark> L <mark>P</mark> S · · · ·	D <mark>P</mark> KSSMTESSEVSHSD
BdWRKY-IC1a		ASHS PLR MFDHIEQ	· · · · · <mark>P</mark> SFSAAASSSEAVQH · · · · · ·	· · · · · · · · · · · · · · · · · · ·	· · · EMA <mark>T</mark> · I SNNDNAAFHSAE · · · · · · ASQRYQV
ShWRKY-IC1a		ATHS PLR MFDHLEQ	· · · · · <mark>P</mark> SFSTAATTS <mark>G</mark> ALQH · · · · · ·	· · · · · · · INSAASMA <mark>G</mark> I S · · · · · · ·	DMT <mark>M</mark> ATANNENASFQSAE ASQRYQV
ZmWRKY-IC1a		ATHS PLR MFDHLEQ	· · · · · <mark>P</mark> SFSTAATMS <mark>G</mark> ALQH · · · · · ·	· · · · · · · MNSAASMA <mark>G</mark> I S · · · · · · ·	DMT <mark>MATANNENTSFQSAE ASQRYQV</mark>
OsWRKY-IC1a		ASHS PLR MFDHTEQ	· · · · · <mark>P</mark> SFSAA <mark>P</mark> TSSEAMQH · · · · · ·	·····MNAAVNMT <mark>G</mark> IS······	DMV <mark>MGP</mark> TNNENVAFQ <mark>P</mark> AE ASQRYQV
ShWRKY-IC1b		AVHS QYNMINHADYA	· · · · I <mark>P</mark> FSSTTT <mark>P</mark> ALITAQH · · · · · ·	ANSSANVTSAQ	EK <mark>PAL</mark> PSHT <mark>G</mark> NSK I ESNE VSQ <mark>G</mark> LK <mark>P</mark>
ZmWRKY-IC1c		AVHS QYNMINHTDYS	· · · · I P F S S T T A P AL I T A Q H · · · · · ·	ANSSANVASAQ	· · · EK <mark>PALP</mark> SHA <mark>G</mark> NSNIESNE · · · · · VSQ <mark>G</mark> LKT
ZmWRKY-IC1d	• • • • • • • • • • • • • • • • •	ATHS PLR MFDHLEQ -	· · · · · <mark>P</mark> SFSTAATTS <mark>G</mark> ALHH · · · · · ·	••••••••••••••••••••••••••••••••••••••	DMT <mark>M</mark> ATANNEN <mark>P</mark> SFQSAE ASQRYQV
BdWRKY-IC1b		AVHS QYT I VNQADYP	L P F S S - T T S A F T S - Q H	VNSSANITSTE	· · · ET <mark>PT PP</mark> SLT <mark>G</mark> NSNFK <mark>P</mark> NE · · · · · · VSQ <mark>G</mark> FQT
VvWRKY-IC1b*	• • • • • • • • • • • • • • • • • •	AHS · · · · · · · HMQ LQAE F P	• • • S S L S V S <mark>P</mark> A A S L T Q F <mark>P</mark> S F • • • • • •	ASNTKAHEQM <mark>PP</mark> LVS	D <mark>A</mark> RTAVKESS <mark>G</mark> LSQSD QRSQ <mark>P</mark>

	480	490	500	510	520	530	540	550	560	570	580	
PpWRKY-IC2a	MSQQQAQ	VMAHMQSR	- 000400000	00100000	o		LSNOS	SORAAV	EAENS	DSEQDNOP	PPPMAAP	· · · - · ·
AtWRKY-IC2a			· · · · · Y <mark>G V</mark> QDHE ·	KKQEMI <mark>P</mark> N			· · · · E AT	QNNN	••••••••••••••••••••••••••••••••••••••	ERQIKI <mark>P</mark> AY	MVS	
AtWRKY-IC2b		<u>.</u>	• • • Q L D L <mark>F</mark> K • • • •	· · · SEIMS · · ·			<u>.</u>				<u>.</u>	
PtWRKY-IC2a	FQE <mark>P</mark> AI	KQDEFVS <mark>G</mark> ··	•••KSNM <mark>v</mark> kmey•	- NSNSMKSFS -			• • • • • P EIAA	IQAN	PQSNN	3 F Q S D H <mark>G</mark> N Q	Ρααγ	
GmWRKY-IC2a	FQEATH	KQDNFSS	··· <mark>G</mark> K <mark>G</mark> MMKTE··	- NSSSMQSFS -			PEIAS	VQT	NHSN	3 F Q S D Y <mark>G</mark> N Y	<mark>РР</mark> Q	
MtWRKY-IC2a	YIENTI	NQNAFSSE	KNMIQTTE-	- NNSSMQSFS -			• • • • • PEIAS	νατη	· · · · · NTNN	FQSDYSNY	aa <mark>p</mark> aaa <mark>p</mark>	
GmWRKY-IC2b	••••••••••••••••••••••••••••••••••••••	KQD <mark>G</mark> FASR	- ISMSM <mark>V</mark> KTE	- TTSAMQSLT -			• • • • <mark>P</mark> ENN •		•••••NHRN	FQSDHKNY	a <mark>p</mark> aa	
VvWRKY-IC2a	· · · · · · · · · · · · · · · · · · ·	NQ	· · · · L <mark>G</mark> V <mark>F</mark> KNQ · ·	PSKIVFHQ-			••••KKLR•		•••••ATE	FATMDSNRI	MAVT	
BdWRKY-IC2a			· · · S F <mark>P</mark> S I KEQQ -	VAQVS			<u>.</u>			· · · · · · · NN	KSN	
BdWRKY-IC2b			SF <mark>P</mark> S <mark>F</mark> KEQQ -	QQQVSKS -			vv <mark>p</mark>		ASN:	5 · · · · · NN	K S N	
ShWRKY-IC2a			· · · S F <mark>P</mark> S <mark>F</mark> K Q E Q -	QQQVEAA -			· · · · AT T NK	Q S A V	V <mark>A</mark> SSNI	NKQ <mark>G</mark> SS <mark>G</mark>	66 <mark>NSS</mark>	
ZmWRKY-IC2a			· · · SL <mark>P</mark> S <mark>F</mark> EEEQ ·	QQQVEKA-			• • • • • • • • • • • •		· · · · · · · · S S N I	R AS <mark>G</mark>	<mark>GGNG</mark> · · · · ·	
OsWRKY-IC2a			· · · S F <mark>P</mark> S <mark>F</mark> K E Q E ·	QQVVESS -			••••KN <mark>G</mark> ••	· · AA · · · · · ·	•••••AASSNI	< - <u>-</u> s <mark>G</mark>	GGG <mark>N</mark>	
AtWRKY-IC2¢			SQEQQKK	NQ SEQWS -			<u>.</u>		•••••QTE	rr <mark>p</mark> n NQ	AVS	
CpWRKY-IC22	FDKSTKQ	QQHDFS	- L <mark>g</mark> kt <mark>ga</mark> ksdf -	APLECFS -			TEIN <mark>P</mark>	FQS	SN	- A <mark>G</mark> AQSSFS	QYNSQ <mark>P</mark> C	
PtWRKY-IC2b	••••FNQS•••F	KQTDFSL	- DQKT <mark>GV</mark> KSVS -	A <mark>P</mark> VQSFS -			SELV <mark>P</mark>	LQAN	<mark>M</mark> QNS	- TA <mark>P</mark> Q <mark>P</mark> SYN	QYN-QA <mark>G</mark>	
PtWRKY-IC2¢	••••FDQL•••F	KQTDFSS	- DQKT <mark>GV</mark> KSEF -	APEQSFS -			S E L V <mark>P</mark>	LQAN	<mark>M</mark> QSVI	NT AAQ <mark>P</mark> S FN	QYN-QSA	
GmWRKY-IC2¢	····WKFNE <mark>P</mark> TH	KQSDFS <u>5</u>	· · · ERTA <mark>T</mark> KSEY ·	<mark>P</mark> SIQKFS-			••••SEMAA	GK <mark>P</mark> E	IQSN:	5 V <mark>P G</mark> S <mark>G</mark> Y F D	YT S	• • • • • A
GmWRKY-IC2d	····WKFNE <mark>P</mark> TH	KQTDFS <mark>P</mark> ···	· · · ERTA <mark>T</mark> KSEF ·	<mark>P</mark> SIQSFS -			••••SEMAE	GK <mark>P</mark> E	IQSS:	5 V <mark>P G</mark> S <mark>G</mark> Y F D	YT S	• • • • • A
MtWRKY-IC2b*	····WKFNE <mark>P</mark> K	KQTDFSS	· · · ERTASKSEF -	QSTQSYS-	· <u>-</u>		SE I V <mark>P</mark>	I К <mark>Р</mark> Е	IHSN:	SVT <mark>G</mark> S <mark>G</mark> YYN	Y <u>NNN</u>	· · · · A
BdWRKY-IC2c	Q <mark>G</mark> EEAYR	<mark>6</mark> 99	- QQ <mark>PWGG</mark> YQQ <mark>P</mark> -	- AAA <mark>G</mark> MEASAN	N <mark>P</mark> A		S F T A	А <mark>РР</mark> L	• Q A T S S	···EMA <mark>P</mark> HA	Q <mark>GGG</mark> AYRQT	нs
OSWRKY-IC2b	····L <mark>G</mark> DEAYRS	sqqq	- QQ <mark>PWGY</mark> QQQ <mark>P</mark> -	AGMDAGAN	A		ASFGA	A <mark>P</mark> - F	QATSS	···EMA <mark>P</mark> QV	Q <mark>GGGG</mark> YSQ <mark>P</mark>	Q S
ShWRKY-IC2b	····L <mark>G</mark> EEAYRS	50000	- QQ <mark>PWGY</mark> QQQQ -	- Q <mark>P</mark> AGMDAGSS	Q		AAYG	G A F • • • • • • • •	Q AG S S I	DA <mark>G</mark> AMA <mark>P</mark> HV	<mark>P</mark> AS <mark>GG</mark> YSHQ	AQ
ZmWRKY-IC2b	····LGEEAYRS	saa	- QQ <mark>PWAY</mark> QQQQ -	<mark>PG</mark> MDAGSS	QQA		APYG	E <mark>P</mark> F	QAASSI	DAATMA <mark>P</mark> HV	<mark>P</mark> AS <mark>GG</mark> YSHQ	AQQS
ZmWRKY-IC2c	<mark>P</mark> VEEAYR	<u> 6</u> 00 - <u>-</u>	· · · Q PWAYQQQQ -	AAGMDAGSS	Q · · · · · · ·		AAYG	<mark>GР</mark> FH	QAASSI	DAAAMA <mark>P</mark> HV	<mark>P</mark> AS <mark>gg</mark> yshq	AQQS
BdWRKY-IC2d	••••• <mark>G</mark> DSYNRI	EQQQ <mark>P</mark> QQQ	- QQ <mark>P</mark> WS <mark>Y</mark> QSDT -	TR <mark>P</mark> ADFTT <mark>P</mark> FD			FEAA <mark>P</mark>	DN	MLGNN	VAS <mark>GG</mark> YSTA	<mark>P</mark> AGTGVRAQ	SG
OsWRKY-IC2c	····· <mark>GD</mark> PYNGI	EQQ	· · · Q PWS YQE PT ·	TRPAEFTS			· · · · · SAAA	G D	VAGNG	SYSQVAA <mark>P</mark> A	AA <mark>G - G</mark> FRQQ	5
ShWRKY-IC2c	V <mark>G</mark> DSYN <mark>G</mark> (GELQQQQQ	- Q P P WAYQE P C -	· · · TQFEA <mark>P</mark> · ·			AAAQ <mark>P</mark>	DNS	MLGNG	3 - Y <mark>ggapg</mark> p	A V S <mark>G</mark> C F R E Q	SQS
ZmWRKY-IC2d*			S F <mark>P</mark> S <mark>F</mark> K E Q Q -	QQQVEAA -			···ткз··	· · <u>·</u> A · · · · · ·	V <mark>P</mark> SSNI	< AS <mark>G</mark>	<mark>6666</mark>	
GmWRKY-IC2e'	· · · · · MISSEAA	KQTDFSS	· · · ERTE <mark>T</mark> K <mark>P</mark> EY ·	<mark>P</mark> STQ <mark>G</mark> FS-			· · · · AALAS	K <mark>P</mark> E	IQSN:	S A <mark>P G</mark> S VH F N	STY	· · · · A
GmWRKY-IC2f*	MIFNEAA	KQTDF <u>S</u> S	· · · ERTE <mark>T</mark> KSEY ·	<mark>P</mark> S T Q <mark>G</mark> F S -			•••• <u>T</u> ALAS	I KHE	IQSN:	5 A <mark>P G</mark> S	<u>.</u>	
MtWRKY-IC2¢*	SDK <mark>P</mark> EH	КQТНF <mark>P</mark> ····	- QEKT <mark>GV</mark> KSEF -	ASLQSLS -			PEIAS	IQTN	<mark>M</mark> Q S N	- N I <mark>P</mark> QS <mark>G</mark> RS	HHA - Q <mark>P</mark> S	
ZmWRKY-IC2e*		- MLQQQ	PWT Y QE P T -	AQ FEAP			AAQA	GN	MFGTG	3GYGAAPG-	FREQ	RQ 5





-----TIQIQNMSP

IE1-PKVNSECMPEGNGKNSSDLKT I NVSNT PVAEVNKKNT SGGRAL

PKANSVOVSKODKOTPSDOTTI

- PKANSVQVPKVDKGTPSDGTTL

- PEVKNACISEAGNOSSAELKAL

VvWRKY-JE1a

MWRKY-IE1a*

-----FSSVPSP-

	WRKY DNA-binding domain		
600 610 AttWRKY-IA1a ETST 600 ST 900 KWKKY-000 CV KO PHWRKY-IA1a SHA SNO DR SV DO YNWEKY-000 CV KO 900 KWKY-000 CV KO PHWRKY-IA1a SHA SNO DR SV DO YNWEKY-000 CV KO 900 KWKY-000 CV KO VWRKY-IA1a TS SV NO DR SV DO YNWEKY-000 CV KO 900 KWKY-000 CV KO GmWRKY-IA1a TS SV NO DR SV DO YNWEKY-000 CV KO 900 KWKKY-000 CV KO GmWRKY-IA1a TS SV NO DR SV DO YNWEKY 000 CV KO 900 KWKKY-000 CV KO GWWRKY-IA1a FS TN NA DR SV DO YNWEKY 000 CV KO 900 KWKKY CO KV KO GWWRKY-IA1a SV VV SS DR SV DO YNWEKY 000 CV KO 900 KWKKY CV KO	C20 C30 C44 E CFR SYN CTH PK CP VKKY ERS C44 C44 S CFR SYN CTY NCP VKKY ERS C44 C44 S VFR SYN CTY NCP VKKY ERS C44 C44 S VFR SYN CTY NCP VKKY ERS C44 C44 S VFR SYN CTH NCP VKKY ERS C44 C44 S VFR SYN CTH NCP VKKY ERS C44 C44 S VFR SYN CTH NCP VKKY ERS C44 C44 S VFR SYN CTH NCP VKKY ERS C44 C44	CONTRACTOR	IA1-
GmWRKY-181a* - LLNVDK - PADDQYNWRKYODKQVKQ GmWRKY-181a* - LLNVDK - PADDQYNWRKYODKQVKQ PWRRY-181a* - PADDQYNWRKYODKQVKQ FWRRY-181a* - SLNVDK - PADDQYNWRKYODKQVKQ GmWRKY-181a* - JNVDNK PNDQYNWRKYODKQVKQ OSWRKY-181a* TALTEDK - PADDQYNWRKYOQKAVKQ	SE FPRSYYKCTHPNCSVKKKVERS- SE FPRSYYKCTHPNCPVKKVERS- SE VPRSYYKCTHINCPVKKVERS- SE PPRSYYKCTHINCPVKKVERS- SE FPRSYYKCTHPNCPVKKKERS- Ge VPRSYYKCTHSCPVKKKVERS-	LEGH VTATI IY KOCHMHOR PHP NKRSKDT-MTSNANSNIOGS	IB1-
A GWRKY-1829 VN - • VDK - PADDOYNWRKYOOK PIKO Bawrky-1829 PAAAAIEOPGKDOYNWRKYOOK PIKO SWRKY-1829 PPP-AIEOPAKDOYSWRKYOOKOLK ZmWRKY-1829 PPPLAIEOPAKDOYSWRKYOOKOLK	CEY <mark>PRGYYKCTHVNCPVKKKVERS</mark> - AE <mark>SPRSYYKCTREACPVKKIVERS</mark> - AES <mark>PRSYYKCTRGCPVKKVVERS</mark> - AES <mark>PRSYYKCT</mark> RDGCPVKKVVERS -	EDGE I TO I I YKGAHDHE <mark>R</mark> PAN RAGGOGR - DS - D	IB2-
ABWRKY-1619 FONADK-FADDGYNWRKYGGKUVK ABWRKY-1619 FUNDK-FADDGYNWRKYGGKUVK PWRKY-1619 SLAUDK-FADDGYNWRKYGGKUVK PWRKY-1619 SLAUDK-FADDGYNWRKYGGKUVK PWRKY-1619 SLAUDK-FADDGYNWRKYGGKUVK PWRKY-1619 SLAUDK-FADDGYNWRKYGKUVK PWRKY-1619 SAJDGYNWRKYGKUVK GRWRKY-1619 FAAU-FN-FADDGYNWRKYGKUVK GRWRKY-1619 SCAIDK-FADDGYNWRKYGKUVK GRWRKY-1619 SCAIDK-FADDGYNWRKYGKUVK BWRKY-1619 SAJDGYNWRKYGKUVK BWRKY-1619 SAJFK AD-VDK-FADDGYNWRKYGKUVK GSWRKY-1619 SAJTFK FADDGYNWRKYGKYUK GSWRKY-1619 SAJFFK SAJTFK FADDGYNWRKYGKYUK GSWRKY-1619 SAJFFK SAJFFK FADDGYNWRKYGKYUK GSWRKY-1619 SAFFFK SATFFK FADDGYNWRKYGKYUK GSWRKY-1619 SAFFFK SATFFK FADDGYNWRKYGKYUK ZIWWRKY-1619 SAFFFK SUBGYNWRKY-1619 SAFFFK SUBGYNWRKY-1619 SAFFFK GS	S G F P RS YY K CTH P A G P V KKK VE R3 S F F P RS YY K CTH P A G P V KKK VE R3 S F F P RS YY K CTH P G C V KKK VE R3 S F Y RS YY K CTH N C V KKK VE R3 S F Y RS YY K CTH N C V KKK VE R3 S F Y RS YK CTH N C V KKK VE R3 S F Y RS YK CTH N C V KKK VE R3 S F Y RS YK CTH N C V KKK VE R3 S F Y RS YK CTH N C V KKK VE R3 S C P RS YK CTH N C V KKK VE R3 S C P RS YK CTH N C V KKK VE R3 S C P RS YK CTH N C V KKK VE R3 S C P RS YK CTH N C V KKK VE R4 S C P RS YK CTH N C V KKK VE R4 S C P RS YK CTH N C V KKK VE R4 S C P RS YK CTH S C V KKK VE R4 S C P R S YK C TH S C V KKK VE R4 S C P R S YK C TH S C V KKK VE R4 S C P R S YK C TH S C V KKK VE R4 S C P R S YK C TH S C V KKK VE R4 S C P R S YK C TH S C V KKK VE R4 S C P R S YK C TH S C P KKKK VE R4 S C P R S YK C TH S C P KKKK VE R4 S C P R S YK C TH S C P KKKK VE R4 S C P R S YK C TH S C P KKKK VE R4 S C P P S S YK C TH S C P KKKK VE R4 S C P P S S YK C P R S C P KKKK VE R4 S C P P	LDGQVTEIIVKGGHNHELFGG, KRGN, NNGSCKSSDIAN	IC1-

	WRKY DNA-binding doma	nain
		••••••
600 610	620 630	640 6 50 660 670 680 690 700
PpWRKY-IC22 TIMDRPSEDGYNWRKYGOKOVKGS	SEYPRSYYKCTOTNCPMKKKVI	VERS - HDGQVTEIVYKGDHNHPKP - QPTRRMALSGAH SLSDGLSRDGDGNDSRPDS
AtWRKY-IC22RNSNDGYGWRKYGQKQVKKS	SENPRSYFKCTYPDCV <mark>SKK</mark> IVI	VETAS - DGQ ITEIIYKGGHNHPKP EFTKRPSQSSLPSSVNGRRLFNPASVVS
AtWRKY-IC2b · · · · · SNKT <mark>SDDGYN</mark> W <mark>RKY</mark> GQKQVKGS	SENP <mark>RSYFKCTYPN</mark> CL <mark>TKKK</mark> VI	VET <mark>S</mark> LVK <mark>GQMIEIVYK</mark> GS <mark>HNHPKP</mark> ··QSTKRS <mark>P</mark> ·····STAIAAHQNS······
<i>PtWRKY-IC2a</i> - QSVREQKR <mark>SDDGYNWRKYGQKQ</mark> V <mark>KG</mark> S	SENP <mark>RSYY</mark> KCTFPNCPTKKKVI	VE <mark>RS</mark> - LDGQ I TE I VYKGSHNHPKP - · QSTRRSSST <mark>Q</mark> S - · · · · · · · N <mark>P</mark> AMIPAPNSNSNE I QDQSYVTH
GmWRKY-IC22 - SQTLSR - R <mark>SDDGYNWRKYGQKQ</mark> V <mark>KG</mark> S	SENP <mark>RSYY</mark> KCTYPNCPT <mark>KKK</mark> VI	<mark>VERS</mark> - LDGQ I TE I VYKGTHNHPKP - · QNTRRNS - SNS - · · · · · · · SSLA I <mark>P</mark> HSNS I RTE I <mark>P</mark> DQSYATH
MIWRKY-IC22 • TQTLSR • RSDDGYNWRKYGQKQVKGS	SENP <mark>RSYY</mark> KCTYPNCPTKKKVI	VERS · IEGOVTEIVYKGTHNHPKP · · QCTRRNSSSSS · · · · · · · · NALVVVPVN <mark>P</mark> · INEIHDQSYASH
GmWRKY-IC2b · VQTLSR · RSDDGYNWRKYGQKQVKGS	SENPRSYYKCTYPNCPTKKKVI	VEKS · LDGQ I TE I VYKGTHNHPKP · · QAAKRNSLSAS · · · · · · · SSLA I PHSNHGSNELPHH · · · · ·
VVWRKY-IC2a - THHLKQRR <mark>S</mark> EDGYNWRKYGQKQVKGS	SENPRSYYKCTYPNCPTKKKVI	VERS LEGQVTEIVYKGTHNHPKP · QSTRRSS · LS · · · · · · SSQTIQASNPPNNEVPDQ · · · ·
BawkKy-JC22KQ-LEDGYKWKKYGOKOVKGS	SENPRSYYRCTYSNCSMKKKVI	VERSLADGRITOTVYKGAHHHPKP-LSTKKHNISPP-
	SEDPRSYYKCTHAGCSMKKKV	
	SENPRSTICTINS COMPREND	VERALADARITATVIKAANNERE - LSIRENSSSGOVAADEEQQAAANSLSAAAAAAGUU
	SENDRSYYKCTYNGCSMKKKV	
	SENPRSYYKCTEPNCPTKKKV	
C#WRKY-IC2# - QYVREQRRAEDGYNWRKYGQKQVKGS	SENPRSYYKCTFPNCPTKKKV	VERS - LDGQ ITE IVYKGSHNHPKP - · QA · RRSSSH · · · · · · · · · SICNNSE ISDQSGSLLDT · · · ·
PtWRKY-IC2b HYMRENGRSEDGYKWRKYGQKQVKGS	SENP <mark>RSYY</mark> KCTYPNCTTKKKVI	VERS - LDGQ ITE IVYKGSHNHSKP - QSTRRSSSQSV YPSGGANSE ISYQSGAP
PtWRKY-IC2¢ • HYMRENKR <mark>SDDGYNWRKYGQKQVKG</mark> S	SENP <mark>RSYY</mark> KCTYPNCPTKKKVI	VERS - LDGQ ITE IVYKGSHNHPKL - QSSRRSSSQLV QPSGGASSE ISDQS IAP
GmWRKY-IC2¢ SLSVREQKRAEDGFNWRKYGQKQVKGS	SENP <mark>RSYY</mark> KCTHPNCSVKKKVI	VE <mark>KT</mark> - LEG <mark>Q</mark> I <mark>T</mark> E I VY <mark>K</mark> GQ <mark>HNHPK</mark> L - · QSTRRTNSQS I N - · · · · · · Q <mark>P</mark> SSSCTNS <mark>G</mark> I SDQSVVTL <mark>G</mark> N - · · · ·
GmWRKY-IC2d SQSVREQKR <mark>TEDGFN</mark> WI <mark>KYGQKQ</mark> V <mark>KGS</mark>	SENP <mark>RSYY</mark> KCTH <mark>PNCSVKKK</mark> VI	VE <mark>KS</mark> - LDGH I <mark>T</mark> E I VYKGQHSHPKP - · QSTRRTNSQS I H - · · · · · · QPSSSCTNS <mark>G</mark> I TDHSVVT L <mark>G</mark> N - · · · ·
MtWRKY-IC2b* SQFVREQKR <mark>SEDGYNWRKYGQKQ</mark> V <mark>KG</mark> S	SENP <mark>RSYY</mark> KCTNPNCSM <mark>KKK</mark> VI	VE <mark>R</mark> D·LDGQITEIVYKGTHNHPKP··QSNRRTNS······QPTSSCTNS <mark>G</mark> ISDQS······
B#WRKY-IC2¢ •••••QRRS <mark>SDDGYNWRKYGQKQ</mark> VKGS	SENP <mark>RSYYKCTFPSCPTKKK</mark> VI	<mark>VETS</mark> L-EG <mark>QITEIVYK</mark> GT <mark>HNHAKP</mark> LNTRRSS <mark>GAG</mark> <mark>G</mark> AAAQVLQS <mark>G</mark> A <mark>GG</mark> DTSEHSFGGG
OsWRKY-IC2b QRRSSDDGYNWRKYGQKQVKGS	SENPRSYYKCT FPNCPTKKKVI	VERSL-DGQITEIVYKGTHNHAKPQNTRRNS <mark>G</mark> SSAAQVLQS <mark>GG</mark> DMSEHSFGG-
SbWRKY-IC2b · · · · · · RRSSDDGYNWRKYGQKQVKGS	SENPRSYYKCT FPSCPTKKKVI	VERSL-DGQITEIVYKGTHNHAKP-QNTRRNS
ZmWRKY-IC2b · · · · · QRQSSDDGYNWRKYGQKQMKGS	SENPRSYYKCTFPGCPTKKKVI	VERSL-DGOITEIVYRGTHNHAKPQNTRRNSSAAAQLLQSGGGDASEHSFGG-
	SENPRSYYKCTFPSCPTKKKVI	VERSL DGGITETVYKGTHNHAKP - ONTRKNSSS- AAAAQLLQ - GGDASEHSFGG
	SENPRSTIKCSAPGCPTKKKV	
	SENPRSTICT FOCFICKEN	VEDEP DOVIEIVINGANSPREF, UNKRAGGOGO - TALDO - TALDO - AASDATSAD
ZmWRKY-IC2d*TKI EDGYNWRKYGOKOVKGS	SENPRSYXKCTYHSCSMKKKV	
GmWRKY-JC2e* PKS I REQKRSEDGYNWRKYGEKQVKGS	SENPRSYYKCTHPSCPTKKKV	VERS-LEGHITEIVYKGSHNHPKP-LG-RKNGSGSIHOTSSSCTNSGISDO
GmWRKY-JC2f* SIREQRR <mark>SEDGYNWRKYGEKQVKGS</mark>	SENPRSYYKCTHPSCPTKKKV	VERS - LEGHITEIVYKGSHNHPKP - HG - RKNGSQSIH
MWRKY-IC2¢* · ESYREQRR <mark>SDDGYNWRKYGQKQ</mark> VKGS	SENPRSYYKCT FPNCPTKKKVI	VERS - LEGHITEIVYKGTHSHPKP - OP - KRSSSQ - S
ZmWRKY-IC2e* · · · · · HR <mark>P</mark> S <mark>SDDGYNWRKYGQKQMKGS</mark>	SENPRSYYKCSFPGCP <mark>T</mark> KKKVI	VEQSP - DGQVTE I VYKGTHNHPKP - · QSTRRGASSAP - · · · · · · · · ASYVVQSA - · SDAVMPEHSWG - ·

IC2-



Appendix 1 contin

	710	720	730	740	750	760	770	780	790	800	810	820	
AtWRKY-IA1a						PL	WNNOSNDSTONR	FEKMSE <mark>g</mark> evit <mark>f</mark>	FEFAVPRST	NSN <mark>PG</mark> TSD		s <mark>g</mark> ckssoc	
PtWRKY-IA1a						E	GREENQNEIGLP	HS I YQ <mark>g</mark> ka <mark>pp</mark> s	Y DP AGT GT I	NAGTGTSDN -		S C G V S G E C	
PtWRKY-IA1b*											
VvWRKY-IA1a						E	GRIENQNEV <mark>G</mark> LS	FHSTY <mark>pg</mark> ka <mark>p</mark> ln	IY <mark>D</mark>			S <mark>GRWCG</mark> DC	
GmWRKY-IA1a					EVRIENT <mark>G</mark> LSM	HSDYYVK	VPRPNDSALN '	GATNAGGVSTE	N			S C <mark>G</mark> L S <mark>G</mark> E C	
GmWRKY-IA1b							I PQPNDSSLN	I <mark>g</mark> atna <mark>ggg</mark> sme	N			S C <mark>G</mark> L S <mark>G</mark> E Y	
MWRKY-IA1a						N	QNERNE <mark>G</mark> RIENQ	/KASL <mark>P</mark> DDSALE	т	<u>.</u> . <u>.</u>		S C <mark>G</mark> L S <mark>G</mark> E C	
CpWRKY-IA1a						· · · · E	GRVENQNEM <mark>G</mark>	· L S C F Q <mark>G</mark> R Y <mark>P P</mark> A	YEPTAAGT I	HVGTGNSEN - ·		S L <mark>G</mark> L S <mark>G</mark> E C	
								-					
GmWRKY-IB1a*						M	SKMDPESSOAT -	ADHLS	TSESEEVG-				
GmWRKY-IB16*						M	SKMDPESSOAT	ADRI S	TSDSEEVA.				
PtWRKY-IB1a*						AHSV	GRDGESTOAD	SE · · · · · · · PPC	PSDSEEAG-			DAAV	
MWRKY-IB1a*						M	SKMEPESSOAT -	/EHLS	TSDSEDVG -			DRET	
GmWRKY-IB1c*						v	SKMDLGSSQAT -	3EHGS	TSDSEEVD-			DHET	
OsWRKY-IB1a*						A	STRSEQGSQDYS	SKFKA · · · · · ·	SNDGG PS	SSRR <mark>G</mark> DR		<mark>G</mark> EQI	
								-					
AtWRKY-IB2a						т	E V <mark>GG</mark> A <mark>G</mark> QMME S S I	DD 5 <mark>6</mark> Y	RKDHD				1
BdWRKY-IB2a						AD	A <mark>GSG</mark> ADAEEEEI	EH <mark>G</mark> · · · · · · · · · ·				DQL	
SbWRKY-IB22						· · · · · L	AGGGNDA - LAAAI	EED				VD <mark>G</mark>	ID <i>2</i> -
ZmWRKY-182a						· · · · · L	A <mark>GGG</mark> NDA <mark>G</mark> LAAA	EED				AE <mark>G</mark>	
AtWRKY-IC1a						K	SKRDOFTSOVIT	FOMS F	ASDSEEVG-				
AtWRKY-IC1b							REQHEAVSOATT	EHLS	ASDGEEVG-				
PtWRKY-IC1a							- RKDOESSOAT P	H · · · · · · vs	MSDSEEVS -			DTET	
PIWRKY-IC1b							- RKDOESSOAT P	H · · · · · · · I S	MSDSEEVG-			· · · · · · · · DTEA	
PtWRKY-IC1c							DP	SE ·····PPO	SSDNEEAG-			NAAV	
PtWRKY-IC1d*						AHSV	IGRDQESTQAD	SE	PSDSEEAG-				
GmWRKY-IC1a						- NSS V	PESDOTSNOGAP	RQL LPG	SNESEEVG -				
GmWRKY-IC16						- DSS V	AKSDOTSNOGAP	- RQL LPC	SSESEEVG.				
MWRKY-IC1a						- DSSPPE	PESDLTSNOGAL	- RP RP 0	SSESEEVG-			NAEN	IC1_
VyWRKY-IC1a						ANSV	PGMDQETTQAMP	Q VNG	SSDSEEVG-			DAET	
BdWRKY-IC1a						A	AEHNEQSNDTAS	LSGV	RRDQEAVYA	MS		EQL	
ShWRKY-IC1a						A	ADQNEQSNDTTS	LSGA	KRDODNIYG	MS		EQA	
ZmWRKY-IC1a						A	ADHNEQSNDTAS	LSAA	KRDODNIYG	MS		EQA	
OsWRKY-IC1a						A	ADQNEQSNDTVS	LSG	KRDGEALYG	MS		EQL	
ShWRKY-IC1b						· · · · · T	STRSEPGSQDHS	КНЕ <mark>G</mark> · · · · · ·	SNDGIAGPS	VSRRGEG		HEQL	
ZmWRKY-IC1c						· · · · · i	STRSEPGSQEHS	<mark>акне</mark> д	SNDGILGPS	VSRR <mark>GGG</mark> ····		DEQL	
ZmWRKY-IC1d*						A	FDONEOSNDTTS	L S G A	KROODNIYG	MS		EQA	
BdWRKY-IC1b						T	LNRSEQGSODHS	AKFEV	SNDGITVPS				
VvWRKY-IC1b*						AYSL	SKKDQQSSQAIP	EH • • • • • • • L <mark>P (</mark>	SSDSEEMD -			DAET	





IC2-

ppendi		mucu						DxSS				
	710 7	20	730	740	750	760	770	780	790	800	810	820
AtWRKY-IC32		VSCNNTQQQ	GGSNENNVE	E <mark>G</mark> STR	FEY <mark>G</mark> NQ - S <mark>G</mark> S	I O A O T G -	GQYESGD <mark>P</mark> V -	- VVDASSTFS	NDEDEDDRGTH	3 s		VSL <mark>G</mark> Y
AtWRKY-IC3b	· · · · · · · · · · · · · · ·	IDATEYEGF	A GTNENIEW	T S <mark>P</mark> V S A E • • •	LEY <mark>G</mark> SH - S <mark>G</mark> S	MQVQN <mark>G</mark> T	HQF <mark>GYG</mark> D···	· · · AAADALY	RDENEDDRTSH	M S · · · · · · · · · ·		VSLTY
GmWRKY-IC3a	•••• DL GWAN \	QK <mark>GN I AG</mark> AA	NWKHENIEA	T S S A S V <mark>G</mark>	PEYCNQ - S <mark>P</mark> N	LQAQN <mark>G</mark> -	THLDS <mark>G</mark> E···	• AVDASSTFS	NEED DQVTH	3 <mark>5</mark>		VSL <mark>G</mark> Y
GmWRKY-IC3b	•••• DL GWAN \	QK <mark>GN I AG</mark> AA	SWICHDNLEA	ASSASV <mark>G</mark>	PEYCNQQ <mark>PP</mark> N	LQTQN <mark>G</mark> -	THFDS <mark>G</mark> E···	- AVDASSTFS	NE EDE DDQGTH	3 <mark>8</mark>		VSL <mark>G</mark> Y
VvWRKY-IC3a		<mark>P</mark> L T <mark>P</mark>	DWRHDNLEV	T S - S S L <mark>G</mark>	PEFCNT - STT	LQ <mark>G</mark> QN <mark>G</mark> -	A <mark>P</mark> FESSD · · ·	- AVDASSTFS	NDEDDDDRVTH	3 <mark>8</mark>		VSL <mark>G</mark> Y
PtWRKY-IC3a	· · · · · DSAWAGT	QKG I AAGT S	DWRRDNVEV	T S S A S <mark>G G P G</mark> -	PEFGNP-SSS	VQAQS <mark>G</mark> -	TPFESAD	- AIDASSTFS	NDEDDD-RATH	3		s v <mark>g</mark> y
CpWRKY-IC3a	DSIWANT	QK <mark>G</mark> TAS <mark>G</mark> TP	DWRHDNVEV	T	PEYGNQ - SAA	LHAQN <mark>G</mark> -	AQLES <mark>G</mark> D · · ·	- AIDASSTFS	NDEDEEDRATH	3 <mark>8</mark>		VSL <mark>G</mark> Y
BdWRKY-IC3a	NAGGATO	QPNAEAR - P	LWHGGGGMG	VQDWR <mark>GG</mark> D- <mark>O</mark>	LEATSS PGEL	CDSSASM	QVHDGTATE -	- GVDVTSAVS	DE V <mark>DG</mark> D <mark>D</mark> R VRAI	H <mark>G</mark> SMSQ		- GHNQGAA
OsWRKY-IC3a	DGGGGSTF	VGAGQAG - A	EWHNGG	V V <mark>G G E G</mark> L V D A	AT S S <mark>P</mark> S V <mark>P G</mark> E L	CESTASM	QVHEGAAAE -	- GVDVTSAVS	DEVORDOKATH	V L <mark>P</mark>		· · LAAAAA
ShWRKY-IC3a		NAAAEANHQ	PWRAG	VQDGMDA A	AT S S <mark>P</mark> S V <mark>P G</mark> E L	CDSSASM	QQVEYAPRE -	- GADVTSAPS	DEV <mark>DGGD</mark> RVTL	3		- SMSHA <mark>g</mark> a
ZmWRKY-IC3a		AAE AN - Q	AWRAG	VQD <mark>G</mark> VD A	AT S <mark>PP</mark> SV <mark>PG</mark> EL	CDSAASM	Q-VDCAARE -	- GADVTSAVS	DEV <mark>DG</mark> DDRVTL			TH <mark>GG</mark> A
BdWRKY-IC3b		- LDSASLWE	NGRSGCIQD	VQSE <mark>G</mark> VDAR ·	· · · · · PGTRLP	VSAY <mark>g</mark> dt	SIVESQD	· AVDVSSTLS	NEEID··RATH	3 т		VSLDC
OsWRKY-IC3b		- QATATSWE	NAANGHLOD	VGSEVLTK · ·	LSASLT	TTEHAEK	SVMDKQE	- AVDISSTLS	NEEDD RVTH	RA <mark>P</mark>		LSL <mark>G</mark> F
BdWRKY-IC3c			· · · · HHFQD							.		
ShWRKY-IC3b*			· · · · · · · · IK	VQGKGVEAR -	SAAFL <mark>P</mark>	VSAHSDA	SLLESQD	· AVDVSSTLS	NEEE···RATH	эт		VSIEC
VVWRKY-IC3b		GSKNDRAGS		TSSTSAVS	· · · · ALSNSLS	NT <mark>GG</mark> ISM	GIFESAG	- TPDLSLTVA	SODDGEDGATO	s		
GmWRKY-IC3c	· · · · · · · RN 101	GLKDTKOSF		TSSTSAVT · ·	• • • • ELSD <mark>P</mark> IS	TNKAKSL	RIFELED	• TPELSSTLA	SHODDEDGTAH	ALV		
GmWRKY-IC3d		GLRETKOSF		TPTTSAVT · ·	· · · · ELSD <mark>P</mark> IS	TNNAKSL	CMLESED	• TPELSSTLA	SHOGDEDGTAQ	ALV		SAED
MIWRKY-IC3a		WGKDAKHNP	ERKPDGOER	TSPPSGVT	· · · · ELSD <mark>P</mark> MK	R ARSO		- APEHSSALG	NHDGDKDGATO	AVL		
MWRKY-IC3b		RCKDAKHNP		TSOPSOVT		R ARSO	GMEESDD	AOFHSSALD	NHDGDKDGAT .			
PHWRKY-IC3b			DOKYDOLER	TSSTSVVT		AAOVKSI	GTSESTE	TPEISSILA	SHDD	s		
PIWRKY-IC3c							. TYESTK	TPELSSTLA	SHDDD OVTO	s		
OsWRKY-IC3c*		HNSASIW	NAKNDCLOD	VOSEVIETR		VSTNCDT	SIMESOD	AVDVSSTIS	NEEDD. PATH	т		
000000000000				- 40 I E I K	. AAUL					••••••••		

AMARKY ICAA	EKOSOVY, NI SNONFOTONDEVED I SASTADOGE ABASNOND
AUVAA 1-1042	
PTWRK Y-1042	DRSPGATGQVPHATEPNGALELSIG-ANDUIGEGAE
PtWRKY-IC4b	DKS <mark>PG</mark> AYGQVSHAIE <mark>PDG</mark> APELS <mark>PG</mark> TTN <mark>DDTGEG</mark> AE
GmWRKY-IC4a	DKASTMY <mark>G</mark> QVSHAAE PNST PESS PVATNDDG LEGAG FVSNE
GmWRKY-IC4b	DKAT AMYGQVSHAAE PNST PESS PVAT NDDG LEG V
MtWRKY-IC4a	DKDFNNY <mark>G</mark> QMSHAAERDST <mark>P</mark> ELS <mark>P</mark> IAANDGS <mark>PEGAG</mark> FLSNE
GmWRKY-IC4¢	DK <mark>Q</mark> SNMC <mark>G</mark> QQSHLAE <mark>P</mark> DQK <mark>P</mark> ELL <mark>P</mark> VATNDGDLDGLGVLSNRNND
GmWRKY-IC4d*	* • • • • • • • • • • • • • • • • • • •
MWRKY-IC4b	DKASNS
VvWRKY-IC4a	DKS <mark>P</mark> SIY <mark>G</mark> QMAHNID <mark>P</mark> NGTPELS <mark>P</mark> VAANDDVVEGAILD_
BdWRKY-IC4a	DKSSNVLSIL <mark>GNP</mark> VHST <mark>G</mark> MAE <mark>P</mark> VPGSASDDDIDAGAGR <mark>P</mark> YD
OsWRKY-IC4a	DKS SNVLS I L <mark>G</mark> NAVHT A <mark>G</mark> M I E <mark>P</mark> V P G SAS D D D N D A GG G R P Y D
ShWRKY-IC4a	DK <mark>P</mark> SNIYSNLCNQVHSA <mark>G</mark> MIDTV <mark>PGP</mark> ASDDDVDAGGGR <mark>P</mark> YD
ZmWRKY-IC4a	DK <mark>P</mark> SNIYSNIYSNIYSNIYSNIYSNIYSNIYSNIYSNIYSNIY
ZmWRKY-IC4b	DK <mark>P</mark> SNIYSNIYSNIXSNIYSNIXSNIYSNIYSNIXSUDDVDAGGGCRSYD
VvWRKY-IC4b	HKT SHAH <mark>G</mark> QT SYH <mark>G</mark> E LD SV <mark>P</mark> E V <mark>PP</mark> FT AS DD E QE AD E

GmWRKY-ID1a	KASVEHGCMPQQ QP QS F PP AKVS P VNKLNASHL S L TKN
GmWRKY-ID1b	V F F D I I HE DKAS VEH <mark>G</mark> CM <mark>P</mark> QQ I K <mark>P</mark> LQS F P P AKVS <mark>P</mark> VNE LKAAHL QL T KN · · · · · · · · · · · · · · · · · ·
MWRKY-ID1a	QENS S VE YE SM <mark>P</mark> RQVT <mark>P</mark> L R F H <mark>P P</mark> S K V S R T
MWRKY-ID16*	QDN
VvWRKY-ID1a*	
PtWRKY-ID1a	ESHE <mark>P</mark> H · · · · PIESTNT PQISSVTSSEDVKRVLSE PKRD · · · · · · · · · · · · · · · · · · ·
PtWRKY-ID1b	ON SE
AtWRKY-ID1a	EPPKIHGGLHVSVIPPADDVKTDISQSSRITG
CpWRKY-ID1a	

BdWRKY-IE1a	T <mark>P</mark> T SNKLKKSVVENSEQQL F CSSDC <mark>EGD</mark> A <mark>G</mark> I KS · · · · · · · · · · · · · · · · · ·
AtWRKY-IE12	SD <mark>P</mark> SASTKEYICESQTLV <mark>D</mark> RKRHC
CpWRKY-IE1a	
GmWRKY-IE1a	D <mark>P</mark> SISSKESLQEA <mark>P</mark> CSTDKKRQNTSNIS <mark>G</mark> NGKVIL
GmWRKY-IE1b	D P S I S S K E S L Q E A <mark>P</mark> C S N D K K R Q N T S N I S <mark>G</mark> N D K V I L
VyWRKY-IE1b	DLSNYKSE <mark>PG</mark> KASVAM <mark>P</mark> ELERQNSSNS <mark>D</mark> SNT <mark>G</mark> IKA
VyWRKY-IE1a	SDAS DE NPE IKV
MtWRKY-IE1a*	D <mark>P</mark> STSSK-AQEET <mark>P</mark> CSSDKKLQNSSDIN <mark>G</mark> NG <mark>KIVL</mark>
GmWRKY-IE1c*	DPSSS <mark>P</mark> KE <mark>P</mark> LQEA <mark>PCNG</mark> DKNLENSSNVE-NGKIIL
GmWRKY-IE1d*	DPSSS <mark>P</mark> KELLQET <mark>P</mark> CSGDKNLENSSNVE-NGKIIL
PtWRKY-IE1a*	E <mark>P</mark> AT LS I E L VQET SA I SERKRQSSSSS <mark>DE</mark> NK <mark>E</mark> T Q I · · · · · · · · · · · · · · · · · ·



IC3-

IC4-

ID1-





														WRKY	DNA-bind	ling domain		
			Charg	red.2														
				,cu 2														•••
L30 L40 L50 L70 L											-1-		X 7					
		830	840	850	860	870		880	890	900	1	910	Ŧ	920	930	940	950	960
	PpWRKY-IC2a	NED	GDDDEQDS		ISKDATAAP		TIREPR		vvvqt t	SDVDILDD		YGOKVV	KGNPH	PRSYYKCT	NVGCPVF	KHVERASTD	IKAVITTYEGH	HNHDVPAAR
	AtWRKY-IC2a	EID	EEEE <mark>Q</mark> PEMK	RMKREG - ·	EDEGMSIE	· · · · · · s	KG V KEPR			SDIDVLID	G F RWRK	<mark>γοακ</mark> νν	KGNTN	I P R S Y Y K C T	FQGCGVK	KQVE RSAAD	E RAVL TTYEGE	HNHD I PTAL
	AtWRKY-IC2b	· · · · · · · · · <mark>G</mark> KD	IGEDETEAK	RWKRE	EN		· · V <mark>KEPR</mark>		vvv<mark>qt</mark>t	SDIDILDO	G <mark>Y</mark> RWRK	<mark>γοακ</mark> νν	KGNPN	I P R S Y Y <mark>k</mark> c t	FTGCFVF	RKHVE RAFQD	P KSVITTYEGH	HKHQ I PTPR
	PtWRKY-IC2a	К · · · · · · · GDD	LDEDEPEAK	RWKRER - ·	DNE <mark>G</mark> ISA <mark>P</mark> -	<mark>G</mark> S	RT V <mark>REP</mark> R		<mark>v v v q т</mark> т	SDIDILDO	IG <mark>Y</mark> RWRK	<mark>Υσακ</mark> νν	KGNPN	I P R S Y Y <mark>k</mark> c t	FQGCPVF	RKHVE RASHD	L RAVITTYEGH	HNHD VP AAR
	GmWRKY-IC2a	KC · · · · · · GDE	Y D E D E P D A <mark>k</mark>	(RWKIEG - ·	ENE <mark>G</mark> MSA <mark>P</mark>	<mark>G</mark> S	RT V RE P R		<mark>vvvqt</mark> t	SDIDILDE	IG Y RWRK	<mark>Υσακ</mark> νν	KGNPN	I P R S Y Y K C T	HPGCPVF	RKHVE RASHD	LRAVITTYEGH	HNHD VP AAR
	MtWRKY-IC2a	RS · · · · · · · GDE	FDEEEPEAK	(RWKNEG - ·	ENEGISAQ	AS	RT V RE P R		vvvqt t		IG Y RWRK	(YG <mark>QK</mark> VV	KGNPN	I P R S Y Y K C T	HPNCPVR	REN A CARACINA C	LRAVITTYEGH	(HNHDVPAAR
	GmWRKY-IC2b	· · · · · · · · · · · · · · · · · · ·	FDNDEPDAK	(RWR IEG - ·	ENEGISAV	•••• <mark>6</mark> 8	RTVREPR		VVVQT T	SDIDILDO	GYRWRK	(YG <mark>Q</mark> KVV	KGNPN	IP RS YYKCT	FPGCPVF	RENVERASQU	LRAVITTYEGH	HNHDVPAAR
	VvWRKY-IC2a	KS	FDEDEPEAK	RWKKES	ENEGISAP	· · · · · · · G S	RTVREPR		VVVQT	SDIDILDO	GYRWRK	GOKVV	KGNPN	PRSYYKCT	NPGCPVF	RKHVERASHD	LRAVITIYEGH	HNHDVPAAR
	BdWRKY-IC2a		SOGAEPQAK	RWKEDA - L	0N - EGSS		VREPR		···· LVVQTL	SDIDILDO	GFRWRK	YGOKVV	KGNPN	PRSYYKCT	TVACPVE	REHVERASHD	NRAVITIYEGH	HNHDVPLGR
	BOWRKY-JC2D	·····	SDGAEPEPK	RWREDA-L		GOGG				COLOILOD		YCOKW	KGNPK	PRSYYCI	THEODY	CRHVERASHD	TRAVITTYECH	
	2000RR 1-1620		SDODEPDAN	PWW OFD	ENEGREA	GAGG			LYVOT			VCONVV	KONPA	PREVYKCT	TACCAN		RAVITTYEOR	HNHDVPVOR
	OcWRKV-1023		SEGDEPEAK	PWK.ED.	DNEGSSOG	M. 000 A00	PVPEPP			SDIDILDO	GERWRK	VOOKVV	KONPA	PRSYXKCT	TVGCPVB	KHVERASHD	TRAVITTYEGH	HNHDVPVGR
GWMNY//G2 SN GG DD TEFE ARD WG DN E NE 1 1 1 1 1 1 1 1 1	AtWRKY-IC2c	VS DE E	DCGSEPEAK		ETNGGNGG		TVREPR		IVV0TT	SDIDILDO	GYRWRK	YGOKVV	KGNPN	PRSYYKCT	TIGCEVE	KHVERASHD	MRAVITTYEGH	HNHDVPAAR
	CuWRKY-IC2a	SN · · · · · · · · · · · · · · · · · · ·	DDDTEPEAK		ENEGLIGT	G S	TVREPR		· · · · · · · · · · · · · · · · · · ·	SDIDILDO	GYRWRK	YGOKVV	KGNPN	PRSYYKCT	TVGCPVB	KHVERASOD	VRAVITTYEGH	HNHDVPAAR
PWWRY/-C22 N GEDNANCEPDARWEDON - ENERGY STUEPA IVVDITE DIDIDOGYMWRDYOK VON PARSYNCTSVCTVCTVCTVLEPAADDLAVITTYCCTMHUPDAAR GWWRY/-C22 SY GDDILDOGYMWRDYOK STUEPA IVVDITE DIDIDOGYMWRDYOK VON PARSYNCTACSVCTVCTVLEPAADDLAVITTYCCTMHUPDAAR GWWRY/-C22 SY GDDILDOGYMWRDYOK VOUTTE DIDIDOGYMWRDYOK VON PARSYNCTACSVCTVCTVCTVLEPAADDLAVITTYCCTMHUPDAAR GWWRY/-C22 SY GDDILDOGYMWRDYOK VON PARSYNCTACSVCTVCAADDLAVITTYCCTMHUPDAAR BWWR/Y-C22 - NDLUDDEPDSTWRDYOK VON PARSYNCTACCSVCRHUE RAADDLAVITTYCCTMHUPDAAR BWWR/Y-C22 - NDLUDEPDSTWRDYOK VON PARSYNCTACCSVCRHUE RAADDLAVITTYCCTMHUPDAAR BWWR/Y-C22 - NDLUDEPDSTWRDYOK VON PARSYNCTACCTVCCTVCCTVCCTACCVKHUE RAADDLAVITTYCCTMHUPDAAR SWWRY-C23 - NDLUDEPDSTWRDYOK VON PARSYNCTACCTVCCTVCCTACCVKHUE RAADDLAVITTYCCTMHUPDAAR SWWRY-C23 - NDLUDEPDSTWRDYOK VON PARSYNCTACCTVCCTVCCTVCCTVCCTACCVKHUE RAADDLAVITTYCCTMHUPDAAR SWWRY-C23 - - NDLUDEPDSTWRDYOK VON PARSYNCTACCTVCCTVCCTVCCTVCCTACCVKHUE RAADDLAVITTYCCMHUPDAAR SWWRY-C23 - - NDLUDEPDSTWRDYOK VON PARSYNCTACCTVCCTVCCTVCCTVCCTACCVKHUE RAADDLAVITTYCCMHUPDAAR SWWRY-C23 - - NDLUDEPDSTWRDYOK VON PARSYNCTACCTVCCTACCVKHUE RAADDLAVITYCCMHUPDAAR SWWRY-C23	PtWRKY-IC2b	SN	DNENEPDAK		ENESILGA	<mark>G</mark> S			IVVQT	SDIDILDO	GYRWRK	YGOKVV	RGNPN	PRSYYKCT	SAGCPVE	RENVERASHD	LRSVITTYEG	HNHD VP AAR
GIM WRY//C22 SY GO E DILG - FDARTWAGEN - NUD YS VS DS T VIE PR VVVDT T E DILDOG FINARENYOK VKON PARSYNK C A PGG SV KHVER AAND I KAVIT TY EOKHHUND AAR GIM WRY//C22 SY GO EDILG - FDARTWAGEN - NUD YS VS AS ST VKE PR VVVDT T E DILDOG FINARYOK VKON PARSYNK C A PGG SV KHVER AAND I KAVIT TY EOKHHUND AAR GIM WRY//C23 SY GO EDILDOG FINAR SYNK C A PGG SV KHVER AAND I KAVIT TY EOKHHUND AAR GIM WRY//C23 SY GO EDILDOG FINAR SYNK C A PGG SV KHVER AAND I KAVIT TY EOKHHUND AAR GIM WRY//C23 SY GO EDDL G SA ST VKE PR VVVDT T E DILDOG FINARYOK VKON PARSYNK C A PGG SV KHVER AAND I KAVIT TY EOKHHUND AAR GIM WRY//C23 SY GO EDDL G SA ST VKE PR VVVDT HEDILDOG FINARYOK VKON PARSYNK C T AGG V KHVER AAND I KAVIT TY EOKHHUND AAR GIM WRY//C23 SY GO EDDL G SA ST VKE PR VVVDT HEDID I LDOG FINARYOK VKON PARSYNK C T AGG V KHVER AAND I KAVIT TY EOKHHUND AAR GIM WRY//C23 GO EDDL G SKIWK KDG - DGED I S HAG NIT V KEPR VVVDT HEDID I LDOG Y WKEY GOK VKON PARSYNK C T AGG C V KKHVE RA HDL RAVIT TY EOKHHUND V AAR ZIM KRY/C23 GO EDDL D SKIWK KDG - DGED I S HAG NIT V KEPR VVVDT HEDID I LDOG Y WKEY GOK VKON PARSYNK C T AGG C V KKHVE RA HDL RAVIT TY EOKHHUND V AAR ZIM KRY/C23 GO EDDL D SKIWK KDG - DGED I S HAG NIT V KEPR VVVDT HEDID I LDOG Y WKEY GOK VKON PARSYNK C T AGG C V KKHVE RA SHD L RAVIT TY EOKHHUND V AAR ZIM KRY/C23 GO E F D DE EP SKIWK GG E A MAG NIT V KEPR VVVDT HEDID I LDDG Y KWEY GOK VKON PARSYNK C T AGG C V KKHVE RA SHD L RAVIT TY EOKHHUND V AAR ZIM KRY/C23 GO E F D DE EP SKIWK GO E E A MAG NIT V KEPR VVVDT HEDID I LDDG Y KWEY GOK VKON PARSYNK T HOFO Y KKHVE RA SHD L RAVIT TY EOKHHUND VAAR C SKIWK Y GOK V KON PARSYNK Y GOK V KON PARSYNK Y GOK VKON	PtWRKY-IC2¢	N	DNANEPDAK		ENESILGA	<mark>G</mark> S	TVREPR		TVVQT	SDIDILDO		YGOKVV	KGNPN	PRSYYKCT	SVGCPVF	RENVERASOD	L RAVITTYEGH	HNHD VP AAR
Gummery/JC24 SY Gummery/JC24 GUMery/JC24 GUMery/JC24 <td>GmWRKY-IC2c</td> <td>SY GODED</td> <td>N L G P D AK</td> <td>RWKEDN</td> <td>KNDGYSVS</td> <td> <mark>G</mark> S</td> <td>RTVREPR</td> <td></td> <td><mark>vvvqt</mark>t</td> <td>SEIDILDO</td> <td>G F RWRK</td> <td><mark>γοακ</mark>νν</td> <td>KGNPN</td> <td>ARSYYKCT</td> <td>APGCSVR</td> <td>RENVERAAND</td> <td>I KAVITTYEGH</td> <td>HNHD VP AAR</td>	GmWRKY-IC2c	SY GODED	N L G P D AK	RWKEDN	KNDGYSVS	<mark>G</mark> S	RTVREPR		<mark>vvvqt</mark> t	SEIDILDO	G F RWRK	<mark>γοακ</mark> νν	KGNPN	ARSYYKCT	APGCSVR	RENVERAAND	I KAVITTYEGH	HNHD VP AAR
MINURY/S/22* SYSGONDRALV-FEAKUWKGON - ENERYKGASASTYKEFR VVVDTTE ID ILDOGFWYRKYGKYVKOHPARSYKCTAPGONVRKHVERAAHD I KAVITYEGKHHHUPDAAR BWWRY/S/22* GD ED DD PD SKIWK KDG - DO ED IS MG AN TYREFR VVVDTME DI ILDOGYWRKYGKYVKOHPARSYKCTAGCONVRKHVERAAHD I KAVITYEGKHHHUPDAAR GWWRY/S/22* GD ED DD PD SKIWK KDG - DO ED IS MG AN TYREFR VVVDTME DI ILDOGYWRKYGKYVKOHPARSYKCTAGCONVRKHVERASHD LAVITYEGKHHUPDAAR GWWRY/S/22* GD ED DD ED SKIWK KDG - DO ED IS MAAN TYREFR VVVDTME DI ILDOGYWRKYGKYVKOHPARSYKCTAGCOVRKHVERASHD LAVITYEGKHHUPDAAR ZWWRY/S/22* GG ED ED ED SKIWK KDG - DO ED IS MAAN TYREFR VVVDTME DI ILDOGYWRKYGKYVKOHPARSYKCT TAGCOVRKHVERASHD LAVITYEGKHHUPDAAR ZWWRY/S/22* GG ED ED ED SKIWK KDG - DO ED IS MAAN TYREFR VVVDTME DI DI LDOGYWRKYGKYVKOHPARSYKCT TAGCOVRKHVERASHD LAVITYEGKHHUPDAAR ZWWRY/S/22* GG EF DE ED ED SKIWK KDG - DO ED IS MAAN TYREFR VVVDTME DI DI LDOGYWRKYGKYVKOHPARSYKCT TAGCOVRKHVERASHD LAVITYEGKHHUPDAAR ZWWRY/S/22* GG EF DE ED ED SKIWK KDG - DO ED IS MAAN TYREFR VVVDTME DI DI LDOGYWRKYGKYVKOHPARSYKCT TAGCOVRKHVERASHD LAVITYEGKHHUPDAAR ZWWRY/S/22* GG EF UD ED ES KIWK KDG - DO ED IS MAAN TYREFR VVVDTME DI DI LDOGYWRKYGKYVKOHPARSYKCT TAGCOVRKHVERASHD LAVITYEGKHHUPDAAR ZWWRY/S/22* GG EF UD ED ES KIWK KDG - DO ED IS MAAN TYREFR VVVDTME DI DI LDO GYWRKYGKYV	GmWRKY-IC2d	SYGGDGD	N L G P D AK	RWKGDN - ·	ENDGYSVS-	AS	R S V <mark>R E P R</mark>		VVVET	SEIDILDO	G F <mark>RW</mark> RK	<mark>γοακ</mark> νν	KGNSN	IA <mark>RSYYK</mark> CT	APGCSVR	RKHVE RAAHD	I KAVITTYEGH	HNHD VP AAR
BAUMRY/S/22	MWRKY-IC2b*	SYS <mark>GG</mark> NDN	ALV - · PEAK	RWKGDN - ·	ENE <mark>G</mark> YCAS	••••AS	RT V <mark>KEP</mark> R		<mark>v v v q t</mark> t	SEIDILDO	G F <mark>RW</mark> RK	<mark>γοακ</mark> νν	KGNPN	IA <mark>RSYYK</mark> CT	APGCNVF	RKHVE RAAHD	I KAVITTYEGH	(HNHD VP AAR
Osumary:/cize	BdWRKY-IC2¢	N D	LDDDEPDSK	RWRKDG - ·	DGEGIG	••••V <mark>GG</mark> N	RT V REPR		<mark>vvvqт</mark> м	S D I D I L D D	G <mark>Y</mark> R W R K	<mark>Υσακ</mark> νν	KGNPN	I P <mark>R S Y Y K</mark> C T	TVGCPVF	RKHVE RA <mark>S</mark> QD	L RAVITTYEGH	(HNHDVPAAR
Sawarky-/c2a GD EFD EFD SKIWERKDG - DG EFD SKIWERKDG - DG EFD S MAAN TV REFR VVVDTM BD D ILD DG YWERKYG KVV KOH PV RS YK CT TAGC VKH VERA #HD L AVI TV EGK HHUP DA AR ZwWRRY/C2b GG FD EFD SKIWERKDG G DG ED S MAAN TV REFR VVVDTM BD D ILD DG YWERYG KVV KOH PV RS YK CT TG CC VKH VERA #HD L AVI TV EGK HHUP DA AR ZwWRRY/C2b GG FD EFD SKIWERKDG G DG ED S MAAN TV REFR VVVDTM BD D ILD DG YWERYG KVV KOH PV RS YK CT TG CC VKH VERA #HD L AVI TV EGK HHUP DA AR SwWR/Y/C2b GD EFD EFD SKIWER DG GG EA MA L G N TV REFR VVVDTM BD D ILD DG YWERYG KVV KOH PV RS YK CT TG CC VKH VERA #HD L AVI TV EGK HHUP DA AR SwWR/Y/C2b GD EFD EFD SKIWER DG GG EA MA L G N TV REFR VVVDTM BD D ILD DG YWERYG KVV KOH PV RS YK CT TG CC VKH VERA #DL AVI TV EGK HHUP DA AR SwWR/Y/C2b GG EFD V DE FD SKIWER DG GG EA MA L G N TV REFR VVVDTM BD D ILD DG YWERYG KVV KOH PV RS YK CT TG CC VKH VERA #DL AVI TV EGK HHUP DA AR SwWR/Y/C2b CA GE DL DD EFD SKIWER DG GO GG EA SL SG NN TV REFR VVVDTM BD D ILD DG YWER YG KVV KOH PV RS YK CT TG CC VKH VERA #DL AVI TV EGK HHUP DA AR SwWR/Y/C2b VA GE DL DD EFD SKIWER DG GO GG EA SL SG NN TV REFR VVVDTM D D I D UD DG YWER YG KVV KOH PV RS YK CT TG CC VKH VERA #DL AVI TV EGK HHUP DA AR SwWR/Y/C2b VA GE DL DD EFD SKIWER DG GO GG EA SL SG NN	OsWRKY-IC2b	<mark>G</mark> D E	FDDDEPDSK	(RWRKDG - ·	D G E G I S	MAGN	RT V <mark>REP</mark> R		<mark>VVV</mark> QТМ	<mark>S D I D I L D D</mark>	IG <mark>Y</mark> RWRK	<mark>Υσακ</mark> νν	KGNPN	I P R S Y Y K C T	TAGCPVR	RKHVE RASHD	L RAVITTYEGH	(HNHDVPAAR
Zum RYS // C2:	SBWRKY-IC2b	• • • • • • • • • • • • • • • • • • •	FDEDEPDSK	(RWRKDG - ·	DGEG S	····MAGN	RT V RE P R		<mark>VVVQT</mark> M	SDIDILDO	I G Y RWRK	(YG <mark>QK</mark> VV	KGNPN	I P R S Y Y K C T	TAGCPVR	RKHVE RA <mark>S</mark> HD	L RAVITTYEGH	HNHD VP AAR
Zawarky/cize GO E FO DE PD SKIWK ROG 0 DE ROG 0 DE PD SKIWK ROG 0 DE RO	ZmWRKY-IC2b	• • • • • • • • <mark>66</mark> •	- DEDEPDSK	RWRKDGDC	SVGEGIS	•••• MAAN	RT V RE P R		<mark>VVVQT</mark> M	<mark>5 D I D I L D D</mark>	IG Y RWRK	(YG <mark>QK</mark> VV	KGNPN	I P R S Y Y K C T	TPGCPVF	RKHVE RACHD	L RAVITTYEGH	(HNHD VP AAR
BAWRY/1/22 VAGE FEVDEPESKIWEN.GGGEBABAIGENTYKEFEVVVDTIBDIDIDOGYMWENYGKVVKON PNRSYYKCTAGGCVKHVERABOLLAVVITYEGKHHUDVPAR SOWRY/1/22 VAGE DIDDEPDSKIWENGG GOEGEVSIVAGINTYKEFEVVVDTMBDIDIDDOGYMWENYGKVVKON PNRSYYKCTAGGCVKHVERABOLLAVVITYEGKHHUDVPAR SOWRY/1/22 LAADDLUDDEPDSKIWENGG GOEGEVSIVAGINTYKEFEVVVDTMBDIDIDDOGYMWENYGKVVKON PNRSYYKCTAGGCVKHVERACHDTRAVITIYEGKHHUDVPAR ZWWENY/1/22 SVGGODDLD - ALKIWKGEN - ENDERSSGAGGFVKEFEVVDTMBDIDIDDOGYMWENYGKVVKON PNRSYYKCTAGCVKKHVERACHDARAVITIYEGKHHUDVPAR GOMENY/1/22 SVGGODDLG - NEAKIWKGEN - ENDERSSGAGGFVKEFEVVDTMBDIDIDDOGYMWENYGKVVKON PNRSYYKCTAGCVKHVERACHDARAVITIYEGKHUDVPAR GOMENY/1/22 SVGGODDLG - NEAKIWKGEN - ENDERSSGAGGFVKEFEVVDTTBEIDIDDOGYKWENYGKVVKON PNRSYYKCVAPGCVKKHVERACHDARAVITIYEGKHUDVPAR GOMENY/1/22 SVGGODDLG - NEAKIWKGEN - ENDERSSGAGGFVKEFEVVDTTBEIDIDDOGYKWENYGKVVKON PNRSYYKCVAPGCVKKHVERACHDARAVITIYEGKHUDVPAR GOMENY/1/22 SVGGODDLG - NEAKIWKGEN - ENDERSSG	ZmWRKY-IC2c	····GDE	FDEDEPDSK	RWRKDGG	DGEG S	· · · · MAGN	RTVREPR		VVVQT M		IG Y RWRK	(YG <mark>Q</mark> KVV	KGNPN	IP RS YYKCT	TVGCPVF	REN PERASHD	LRAVITTYEGH	HNHDVPAAR
0 3/WR/Y/12/2 VA	BdWRKY-IC2d	GEE	FEVDEPESK	RWR	GGEGA····	MAICGN	RTVREPR		· · · · · · · · · · · · · · · · · · ·	SDIDILDO	GYRWRK	(YGQKVV	KGNPN	PRSYYKCT	MAGCPVF	RENVERASQU	LRAVVTTYEGH	HNHDVPAAR
SAMARY-1922 LA A DUELDUEFD SKAWARUE OUSDOUL V.S. SONAFI KEFR LVVDT LA DUG YKWARVGGK VVKAH PRSYKCT A GCFVKH VERACHD TA AV IT YEGA HAHD V FOR SAMARY-1922 SY OG GDD LG NEAK WKGE O. ENDERSS A GS TVKEPR LVVDT LD I LD OG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR GMWRKY-1922 SY OG GDD LG NEAK WKGE N.ENDEYSYS S A GS TVKEPR VVVDT TE LD I LD DG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR GMWRKY-1922 SY OG GDD LG NEAK WKGE N.ENDEYSYS S A GS TVKEPR VVVDT TE LD I LD DG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR MWRKY-1922 SY GG GDD LS NEAK WKGE N.ENDEYSYS S A GS TVKEPR VVVDT TE LD I LD DG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR MWRKY-1922 SY GG GDD LS NEAK WKGE N.ENDEYSYS S A GS TVKEPR VVVDT TE LD I LD DG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR MWRKY-1922 SY GG GDD LS NEAK WKGE N.ENDEYSYS S A GS TVKEPR VVVDT TE LD I LD DG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR MWRKY-1922 SY GG GDD LS NEAK WKGE N.ENDEYSYS S A GS TVKEPR VVVDT TE LD I LD DG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR MWRKY-1923 SY GG GDD LS NEAK WKGE N.ENDEYSYS S A GS TVKEPR VVVDT TE LD I LD DG YKWARVGGK VVKGN PR SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR MWRKY-1924 SY GG GDD LS NEAK NWKGE Y NEAK INKERSK NYK SYKCVA PG FVKH VERACHD HAV IT YEGA HADVE LOR MWRKY-1925 SY GG GDD LS NEAK NWKGE Y NEAK INKERSK NYKCVA PG FYKK HAV FYGA WKA FYGA HAV FYGA HA	OSWRKY-IC2c	VA·····GED	LDDDEPDSK	RWRRDG - C	DGEGV····		RTVREPR		VVVQTM	SDIDILDO	GYRWRK	YGOKVV	KGNPN	PRSYYKCT	TAGCPVR	RHVERASNO	PAVYTTYEGH	
AMPRINGED EN LE	SRWKKY-IC2¢		SCOPERDAK	RWKKDG - C		V-SLSGNN	OVPEPR			SD I DVL DD	C F RWRK	VCORVV	KGNPK	PRSYYKCT	TAGCPVH	RHVERACHD	ARAVITTYEGH	HNHDVPVGR
GENERATING STRUCK GENER	2mvvrKY-IC28	ev odopp	DIG NEAK	PHAR OF N	ENDOLEVED	AGGG	TVEPR		EVVQTL	S S I D I L D D	O Y DWDW	VOOKVV	KONPA	PREYYKC	AGCPVH	KHVERACHD	MKAVITTYEG	HIHDVPLGR
	GmWRKY.IC28	sy		RWKGEN	ENDGYSYSS		TYKEPK		TT	SELDLIDO	GYRWRK	YGOKYY	KGNP	PRSYXKC	APGCPVE	KHVERASHD	MAVITTYEGH	HIHDVPLGR
	MWRKY-IC2e*	SR	D. ENEREAK	RWKGEY	ENFAISAS		TVKEPR		TT	SPUDUDO	GYRWRK	YGOKYY	KGNP	PRSYXKCT	SIGCEVE	KHVERSSKO	RAVLTTYEGH	HNHDVPAAR
	ZmWRKY-IC2e		LDDDEPDSK		DADAAGCS	V-SVASNN	TVREPR		· · · · · vvvoTM	SDIDVLDD	GYRWRK	YGOKVV	KGNPN	PRSYYKCT	GAGCLVE	KHVERACHD	CAVVTTYEGH	HNHD VPPAR

IC2-

Append	lix 1 co	ontinued			WRKY DNA-binding	domain	
		Charged-1			••••••		
AUWRKY-(C3a AUWRKY-(C3b Gm WRKY-(C3b Gm WRKY-(C3b Gm WRKY-(C3b FWWRKY-(C3a BWWRKY-(C3a SWWRKY-(C3a BWWRKY-(C3a BWWRKY-(C3b Gm WRKY-(C3c Gm WRKY-(C3b MWWRKY-(C3b Gm WRKY-(C3c MWWRKY-(C3c C) WWRKY-(C3c C) WWRKY-(C3c) C) WWRKY-(C3c C) WWRKY-(C3c) C) WWRCY-(C3c) C) W	100 00	E40 E50 000 E 00 E 0 E S KR RK L E A F A A E V E 0 E 0 E S KR RK L E A VA T E 00 E 0 E S KR RK L E S VA - E 00 E 0 E S KR RK L E S VA - E 00 E 0 E S KR RK L E S VA - E 00 E 0 E S KR RK L E S VA - E 00 E 0 E S KR RK L E S VA - E 00 E 0 E S KR RK L E S VA - D E 0 E S KR RK L E S CA - D E L E S KR RK L E S CA - D E L E S KR RK L D S CA T D D E L E S KR RK L D A L A T N D U V E H K R K M D - D E E S KR RK L E D AL A T N D E L E S KR RK L E S YA - D E E S KR RK L E S KR RK L E S KR RK L A A A T N D E L E P K I R K K S YA Y A D E E S KR RK L E S YA - D E E S KR RK L E S KR RK L E S KR RK L A A A T N D E L B P K I R K S YA Y A D E S KR RK K E S Y Y A D E S KR RK K E S Y Y A D E S KR RK K E S Y Y A D E S KR RK K E S Y N Y A D E S KR RK L E S K R K L E S K Y Y A D E S KR RK K E S Y N Y A D E S KR RK K E S Y N Y A D E S KR RK K E S Y N Y A D E S KR RK L E S K R K L L A A L A A D B S KR RK K E S K N Y A D E S KR R K K E S Y	EG0 B70	890 900 1	10 200 300 RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YYK ÖT APÖ T IV RK YGÖK VUK ÖN PPR RS YK ÖT FÖ T IV RK YGÖK VUK ÖN PPR RS YK ÖT FÖ T ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT FÖ T ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT FÖ T ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT FÖ T ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT FÖ T ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT FÖ T ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT F ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN PPR RS YK ÖT S ÖT IV RK YGÖK VUK ÖN		IC3-
AtiwRKY-IC4a PtwRKY-IC4a PtwRKY-IC4a GmWRKY-IC4a GmWRKY-IC4a GmWRKY-IC4a GmWRKY-IC4a BdWRKY-IC4a BdWRKY-IC4a SbWRKY-IC4a ZmWRKY-IC4a ZmWRKY-IC4a		E P D D D D F S KR - K RME G - A D D D D F S KR - K R L D A G V D D D F S KR - K R L D A G V D D D F S KR - K ME L G N E V D D D F S KR - K ME L G N E V D D D F S KR - K ME L G N E V D D D D F S KR - K ME L G N U T A L E F L S KR M K L D G V D D D D F S KR - K ME L G N U T A L E F L S KR M K L G N G D A T D D L E S KR K ME S A D A V D D D L S KR K ME S A D A V D D D L D S KR K ME S A D A V D D D L D S KR K ME S A D A N D D L L D S KR K ME S A D A N D D L J S KR K ME S A C N D D D L S KR K ME S A C N D D D L J S KR K ME S A C N D D D L J S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K ME S A C N D D D L S KR K K K ME S A C N D D D L S KR K K K K K S A C N D D D L S KR K K K K S A C N D D D L S K K S A K K K S A C N D D D L S K S K K K K S A K S A C N D D D L S K S K K K K S A K S A C N D D D L S K S K K K K S A	M		RH YGOR VV KGN PN PRS YYK CTANG CPY RH YGOR VV KGN PNPRS YYK CTAG CPY RH YGOR VV KGN PNPRS YYK CTANG CPY RH YGOR VV KGN PNPRS YYK CTANG CPY RH YGOR VV KGN PNPRS YK CTANG CPY	RKHVE RASHDPKAVITTYEGKHDHDVPT RKHVE RASHDPKAVITTYEGKHHDVPT RKHVE RASHDPKAVITTYEGKHNEVPT RKHVE RASHDPKAVITTYEGKHNEVPT RKHVE RASHDPKAVITTYEGKHNEVPT RKHVE RASHDPKAVITTYEGKHNEVPT	IC4-
GmWRKY-iD1a GmWRKY-iD1b MWWRKY-iD1a MWWRKY-iD1a VWWRKY-iD1a PtWRKY-iD1a PtWRKY-iD1a CpWRKY-iD1a		QVHDNEEPESKELKKDNIN QVHDNKEPESKELKKDNSNA DESKELKKDNSNI ETDNDDDPDSKRQKKDINN EVDVDDDQSKRQKKSSCND NIHKBYNBPTAKRKKQQNI EFDTSDDPGLKERKKDNYN			RK YGDK LVKGNTNPRSYYRCSNPGCPU RYGDK LVKGNTNPRSYYRCSNPGCPU RYGDKUVKGNTNPRNYRCSSPGCPU RKYGRKUVGKTINRNYFRCAYPGCTU KRYGRKLVKGNTNPRSYYRCSSPGCPU RKYGDK LVKGPNPRSYYRCSSPGCPU RKYGDKLVKGPPPRSYYRCSSPGCPU RKYGDKLVKGNPHPRSYYRCSSPGCPU	KKHVERAŠHDŠKUVITYEGOHDHEIPPO KKHVERAŠYDŠKTVITYEGOHDHEIPPO KKHVEKŠŠONITĮVITYEGOHDHAPPTO KKYVEKŠŠONITĮVITYEGOHDHAPPTO KKYVERAŠHDPLVITYEGOHDHDPPS KKHVERAŠHDPLVITŠYEGOHDHDMPPS KKHVERAŠHDTKLLITYEGKHDHDMPPO KKHVERTŠHDTKLLITYEGKHDHDKPPA	ID1-
BdWRKY-IE1a AtWRKY-IE1a CpWRKY-IE1a GmWRKY-IE1b VVWRKY-IE1b VWWRKY-IE1a MWRKY-IE1a GmWRKY-IE1e* GmWRKY-IE1e* PtWRKY-IE1a*		DE HE SAE POPKR ENE AVE FEPKR KEHNNE VPKR KEHNNE PKR KEHNDE PKR KEHNDE PKR KEHNDE PKR KEHNDE PKR KEHNDE PKR KEHNDE PKR KEHNDE PKR KEHNNE KR KEHNNE KR KEHNNE FKR KENSE	EATTENSSPULE SSDSVSPOKK SDS TDMDSVSPOKK TDMDSVSPOKK CSSLFFTINE SDS UPDKP SDS ESSAVL ECSKAVL COM ECSKAVL COM COM COM COM COM COM COM COM	REGK I I V AGKHSDGYRW F VVHAADVG G GDGYRW - AGVG G GDGYRW - VVHAADVG G GDGYRW F VVHAADVG G BDGYRW - FVVHAADVG G BDGYRW - FVVHAADVG G BDGYRW - VVHAADVG G BDGYRW - FVVHATDVG G BDGYRW - FVVHATC G G GGYRW - FVVHADGYG G GDGYRW - FVVHADGYG G G GGYRW	RK YOOK I V KON PMP RE YY R C TH DOC PY RK YOOK MY KON PHP RN YY R C TS AG DY KK YOK MY KON PHP RN YY R C TS AG DY RK YOK MY KON PHP RN YY R C TS AG DY RK YOK MY KON PHP RN YY R C TS AG DY RK YOK MY KON PHP RN YY R C TS AG CY RK YOK MY KON PHP RN YY R C TS AG CY RK YOK MY KON PHP RN YY R C TS AG CY RK YOK MY KON PHP RN YY R C TS AG CY RK YOK NY KON PHP RN YY R C TS AG CY	RKHVE KAADD INNMUVIT VEOKNINOOP FO RKH IE TAVEN KAVIIT VKOVINNOOP FO RKH IE TAVEN SAVIIT VKOVINNOOP FO RK II TAVEN SAVIIT VKOVINNOOP FO RK II TAVEN SAVIIT VKOVINNOOP FO	IE1-



CC-A	
	100 110
OSWRKY-IIA12 VEAKVTQMSEENRRLTEVIARLYGGQIPRLGLDGSA <mark>SPP</mark> R <mark>P</mark> VS <mark>P</mark> LSGKKRSRES	
SBWRKY-IIA12 LEAKLTQVSEENRRLTEMIAYLYASQVARQSSS <mark>PDTT</mark>	
ZmWRKY-IIA12 LEARLTOVSEENRRLTEIIAYMYASQVAARRS.PD	
BUWRKY JIATA LOAK · · · · · LAGUS DE · · · NRRLTEM I AH LYGNRUARESSPSACTAASPSSNK · · · · · RURALSPPATSTAASPASSK KRGRES	
ZmWRKY-IIAIB LELDRPAAGCRATSSLSDEQ···VTDRPTDRLLAARLPSWLRARVVLVPPIGTGVATHDHDSARRTGGSPRRSPTCTAPRHHHASGRKKSRDS	IIA1-
	I S CAARE DAA
SWWRKT-11A7B LEEK K. KRVSE ENKRLAPALDALLS ADRSNHFKALAIS	
GWWRKY-IIA22 LVEE	E S D Q W C T
GMWRKY-IIA26 LVEE	E SDQCC T
MWWRKY-HA23 LVEE	E SESCIN MF
MtWRKY-IIA20 LVEE	\$ VENCF
GmWRKY-IJA2e LFEELRRLSSENKRLTE.TLNHLCESYVALQKHLSEFSQLRNANFDKENFDKEGTCAVPSLKRKP	E SENCV
GMWRKY-JJA24 LVEE · · · · · · LRRLSN · · · · ENKRLTE · TLKHVCENYVALQKHLNEFSQLRNA · · · · · · · · NFDKE · · · · · AG · TVPSLKRKA · ·	E · · · S VN L F · · · · · · · · · · · · · · · · · ·
<i>Płwr.ky-iiA23</i> Lvee · · · · · · LT <mark>rms</mark> S · · · · Enkklte · MlgvvcDnymalokhladlvsknsv · · · · · · · · · · · · · Ke i S · · · · · · · · · · · · · · · · · ·	•••••EDYSI••••••
<i>PłWRKY-IIA26</i> LVEE······LN <mark>R</mark> VNS·····ENQKLTE·VLGVVCEKYLTLQKHLADLTSKNSE······KELMT······T·PVISMKRKAES·	EDYSN
CPWRKY-IIA22 LVEE · · · · · · · LN <mark>R</mark> ISN · · · · ENKKLTE · MLVAVCONYTOLHNRFVELVNKNAG · · · · · · · · · EKEQ · · · · · · · IATSRKRKADSS	T · · · TEDYCS <mark>G</mark> H I A · · · · · · NN
<i>Atwrky-ija2a</i> REE · · · · · · · LN <mark>R</mark> VNS · · · · ENKKLTE · MLARVCESYNELHNHLEKLQSRQSP · · · · · · · · · E · IEQ · · · · · · TDI P IK · K <mark>r</mark> Kqd <mark>P</mark> ·	
Atwrky-IIA20 LQDE · · · · · · · IN <mark>R</mark> VNS · · · · ENKKLTE · MLARVCEK <mark>Y</mark> YALNNLMEELQSRKS <mark>P</mark> · · · · · · · · · ESVNF · · · · · QNKQLT <mark>GK</mark> RKQEL ·	·····DEFVS······
VYWRKY-11A22 LVEE······LR <mark>R</mark> MST·····ENKKLTE·LLSIMCENYNALHNHLKELMRKNSD·······HQLF······NSRKRKAES·	·····EDNIN·····
VYWRKY-IIA26 LVEKERBYNAERBYNLTE-MLTYMCENYNALRSHYMEYISKNPERESERESP.SSRKRKAESSNN	N · SNNN <mark>G</mark> · · · · · · · · · · · · · · · · · · ·
GmWRKY-HA2e LEEE · · · · · · LKRVSA· · · · ENKKLAE · MLSVVCENYNT LRSHLMEYMRKNG · · · · · · · · · · EKEV · · · · · · SPTSKKRKSESSNN	NN SN LM <mark>G</mark> · · · · · · · · · · · · · · · · · · ·
GWWRRY-IIA21 LEEE	N - SN L MG
GmWRKY-IIA29 LAEE	NST I PMG
GMWRRY-HAZA LAEE	NNSI <mark>PMG</mark>
	N - DNN I A
	N-DNNIV
	A
Charter and a second and a se	
<i>PtWRKY-IIA3</i> 2 LQFE ······LEHRQK····ENESLRF·MLQVMTRRESILQAEIQETEEH	
VVWRKY-IIA32 LKIE ······LERLRK····ENEDLRL·GLEIMGSKYEVLQAHLQKXXQ·······	
GmWRKY-11A3a LKAE · · · · · · LQ <mark>R</mark> VRE · · · · ENNTLRL · MLEVLNSKCTKLEIHLQEIN · K · · · · · · · · · · · · · · · · ·	
GMWRKY-11A3# IKAE · · · · · · LK <mark>R</mark> VRE · · · · ENNTLRM · MLEVL <mark>S</mark> SKYTKLETHLQE IN · K <mark>·</mark> · · · · · · · · · · · · · · · · · ·	••••••••••••••••••••••••••••••••••••••
MWWRKY-NA33 LELELQQVREENNTLRF.MLEVMSIKCTKLESHLQDIK.K	·····
GmWRKY-HAJC LEAKLQRVKEDNRTLRV-MLETLSSKCEKLQSHLQEINNEVGTKSDQSGSVLL	
ECCS	
EUGS	
OSWRKY-11446 LEAE	D. SLDDAAHHRK <mark>P</mark> SPPFV
SBWRKY-HA4B LETE	MDT AP AP P AAAQQQQQHGTHLHQQ
ZWWRKY, HAAA LETE	TDTAPP



	WRKY DNA-bin	ding domain										
				SINS		LDL-	1			alanine-rich-1		
	240	250	260	270	280 290	300	310	320 330	340	350 3	60 370	
OsWRKY-IIA	12 EDSSLLVATY	EGEHNHPHPS	PRAGELPAAAG	A - GGSLPCSI	SIN	· · · SSGPTITLDLTI	NG GAVQVVEAA	H - PPPPPDLKEVCRE	VASPEFRTALVEO	MASALTS DPKFTG	ALAAAILQ <mark>K</mark> LPEF	
Zm MRKY-HA	A EDSSLLVATY	EGEUNUPSPT	-RAGELPSSAS	ATAS <mark>GP</mark> V <mark>P</mark> CST	S I N	SSGPATTIDLT		AAEAPDLKKLCQE	IASPDERTALVEO	MARSLISDSKFTH		
BdWRKY-IIA	12 ENSSVLEATY	EGEHNHPOPT	RAGELTSSCV	R - SGSVPCSI	S I N	· · · SSGPTITLDLT	NG GGVQVVEAG	E - AQOP - DIKKVCRE	VASPEFRRELVEO	MARELTSDENFTG	ALAAAILOKLPEY	
ZmWRKY-IIA	18 [*] EDS <mark>S</mark> VLVATY	EGEHNH <mark>P</mark> C <mark>P</mark> T	- RAGDLPSSAT	A S <mark>g</mark> S V <mark>P</mark> R S I S N	S I N	SS <mark>GP</mark> AITLDLTI	IN <mark>G - A</mark> GGVRVLE	GAEAPDLRKLCQE	IASPDERTALVEO	MAR <mark>sltkd</mark> ckftd	ALAAAILQ <mark>QLPG</mark> Y	A]-
OsWRKY-IIA	10 <mark>ed</mark> r <mark>s</mark> mlva <mark>ty</mark> i	EGEHNHAL <mark>s</mark> t	QTTEFVAS GCT	S QHAG	<mark>g</mark> ssss <mark>p</mark> l <mark>p</mark> cs - is	INSS <mark>GR - TITLDL</mark> TI	IQAGSGSIASCGV -	· · · · EAAAVSGELVT	V L S <mark>P</mark> E L R R <mark>H</mark> L V E E	VVQVLKN <mark>D</mark> AE F VE	AVTNAVAARVVDQ	
ShWRKY-IIA	# DDNLMLVA <mark>TY</mark>	E <mark>GEHNH</mark> EQ <mark>H</mark> A	QSEYSYINDAS	T S Q Q Q Q <mark>P</mark> Q A G	<mark>G</mark> SSSSTL <mark>P</mark> CSIIS	INSL <mark>G</mark> R - <mark>T</mark> ITLGLAI	DQRR <mark>PG</mark> SSS	····NAEAAAVVVGE	I VT <mark>P</mark> ELRKVLVDE	LA <mark>S</mark> LLKN <mark>DP</mark> EF I E	S LATAVADRVMER	•
ZmWRKY-111	Ac DDNLMLVATY	E <mark>gehnh</mark> eqra	QSEY-VVTDAS	SLHQA <mark>G</mark>	S S L <mark>P</mark> R S I I S	IDSSGRKKITIGLA	ADQR <mark>PP</mark> ADS	····NA <mark>G</mark> AAV <mark>G</mark> EI··	IVT PEFRKALVDE		KLASAVADRVMER	
										SGHPSA	SLQNIFIFL	
GmWRKY-IIA	22 EDPTILVTTY	E <mark>gehn</mark> ae i <mark>s</mark> l	VSNQREA <mark>PP</mark> K <mark>G</mark>	5 - 5 <mark>P</mark> V S S <mark>P</mark> T <mark>P</mark> T	I R S · · · · · · · · · · · ·	· · · A A C PTVTFDL VI	SGLVELESAQ · · ·		- KSSIQQFLVQQ	MA <mark>TSLTRDT</mark> NFTT	ALATAI <mark>SGK</mark> IL <mark>E</mark> A	
GmWRKY-IIA	26	• <mark>GP YN</mark> • • • • •		· · · · · · · · <mark>·</mark> · · ·		· · · · · CYNV · · · · ·	•••••					
MWRKY-IIA	22 EDPTILVATY	EGEHNAE I SM	IISSQSEEA <mark>P</mark> LG	- VHVTSPQQI	IQR	TCSTMKLDNV	•••••		- KSSIQQFLVQQ	MATSLTNDPNFTA	ALATAISGRILDH	
MWRKY-IIA	26 EDQNVLVTTY	EGEHNPEMSL	TSSNQSETTPT	NLVPASSSSP	INW	RTAQASKLDLV	PROLVVDDSH		- KSSIQQLLVQQ	MATSLTRDPNFAA	ALATAISGRIIIL	
GmWRK T-IIA	20 EDPSVLVIII	EGENNAETST	N555K5E1 <mark>P</mark> K <mark>P</mark> /	A · V S <mark>P</mark> SAMN I R	55A		SKVVD IN - AQ		- INSSEQUELVIO	MATSI TROPNETA	ALASAISGRILDD	
PIWRKY-IIA2	2 EDPSILVATY	EGEHNHASHS	OHEPSLOSSHO	NEGPIPTOSP	I R S	· · · SAP · TVTLDLI	SGRHGDTATK		MEVPALOKILVOO	MASSLTRDPNFTA	ALAAAI SGRENQT	
PtWRKY-IIA2	• ENPSILVATY	EGEHNHASHS	QPELSLGSSQS	S F G P V P S P S S	I RT	SVP - TVT L D L I (SGMHVDSAKK ····		LQV <mark>P</mark> EVQKVLVQQ	MASSLTRDPNFTA	ALAAAI SGR FNQT	
CpWRKY-IIA	20 EDPSILIA <mark>ty</mark>	EGEHNH <mark>P</mark> I <mark>P</mark> S	RAELSLT PSRG	AAVAGSSA	MR <mark>P</mark> ····	· · · TASGAVT LNL I	SSTLV <mark>PG</mark> EEE···	RQVEE	KKAA <mark>G</mark> IQQVLV <mark>QQ</mark>	MA <mark>SSLTRDP</mark> NFRA	ALAAA I <mark>SGK</mark> FL <mark>G</mark> Q	
AtWRKY-IIA2	2 EDPSLLVA <mark>ty</mark> i	E G T <mark>HNH</mark> L <mark>G P</mark> N	ASE		τ s q	· · · <mark>G</mark> GSS <mark>T</mark> VTLDLVI	I <mark>G</mark> CHRLALEKN	•••••ERD••	NTMQEVLIQQ	MA <mark>SSLT<mark>KDS</mark>KF<mark>T</mark>A</mark>	ALAAA I <mark>SGR</mark> IM <mark>E</mark> Q	∥ IIA2-
AtWRKY-IIA	8 EDPSFLVATY	E G T <mark>HNH</mark> T <mark>G P</mark> H	AS			····VSRTVKLDLV	BOGLE <mark>P</mark> VEEKK	E R	<mark>G</mark> T I QE V L V Q Q	MASSLT <mark>K</mark> DPKFTA	ALATAI SGRLIEH	11112
VVWRKY-IIA	29 EDQSLLIATY	EGEHNHQQPS			MVS	···PRP·TVTLDLIC			NSFLVEQ	MASSLTRDPSFTA	ALAAAISGRILDQ	L
GmWRKY.III	20 EDUSTLVATT	EGEHNHPOAS	SOMEATSGSGR	SVTI GSVPCSA	SIS		TKSKGSNDSK			MATSI TT DPN FRA	ALVAALSGRII.	
GmWRKY-IIA	21 DDQSVLVATY	EGEHNHPOFS	SQMEATSGSGR	SVT LGS VPCTA	S L S	· · · TSTPTL · VTLD	TKSQGSNDSK		GDSPKVPQVLVEQ	MATSLTTDPNFRA	ALVAAI SGRLLHN	
GmWRKY-II/	29 DDQ <mark>S</mark> VLVA <mark>TY</mark>	EGEHNHTH <mark>P</mark> -	SQMEVTTGSNR		S L S	S S A <mark>P T</mark> V T L D I	VTKSKSSSESK	NVN <mark>P</mark> K	TES <mark>PEVP</mark> QVLV <mark>EQ</mark>	MA <mark>TSLT<mark>K</mark>DP</mark> NFRA	ALVAAI <mark>SGK</mark> MLHN	
GmWRKY-II/	28 DDQ <mark>S</mark> VLVA <mark>TY</mark>	EGE <mark>HNH</mark> PH <mark>P</mark> -	SQME VTT GSNR	MT L <mark>g</mark> sv <mark>p</mark> csa	S L S	SS <mark>PPT</mark> AT L DI	VTKSKSSSESK	····NTS <mark>P</mark> K	TES <mark>P</mark> EV <mark>P</mark> QVLV <mark>EQ</mark>	MA <mark>TSLTKDP</mark> NFRA	ALVAAI <mark>SGK</mark> MLHN	
PtWRKY-IIA2	e, <mark>DDQS</mark> VLVA <mark>TY</mark> I	EGEHNH <mark>P</mark> H <mark>P</mark> -	- SMEATSGSSH	BLT L <mark>GS V</mark> PCSA	SLA	· · · SS <mark>GKT</mark> N · ITLDI	TKSKSSNDAK	· · · · · · · · · · · ssk <mark>p</mark> k	T D A <mark>P</mark> E V R Q F L V <mark>E Q</mark>	MA <mark>SSLTKDP</mark> NFTA	ALAAAI <mark>SGR</mark> MLQQ	
PtWRKY-IIA2	d* DDQSVLVATY	EGEHNHPHP -	- SMEATSGSNR	S L T L <mark>G P</mark> A <mark>P</mark> C I A	SLA	SS <mark>GPT</mark> ITLD	TKSKSRNDDR	SSKPR	ADT PEVRKFLVEQ	MASSLTKDPNFTA	ALASAI SGRMFQQ	
AIWRKY-IIA	C EDQSVLVATY	EGENTHPMP -	SQIDSNNGLNRI	VRECEACEAC	VAA	OPTMTMDLTPS		SPTSR	THFPQVQKLLVEQ	MASSLTKUPNFTA		
CPMAA I-IM	20 EDGS VIVAIII	EGENTHE	AI <mark>F</mark> SSN	· VK303A03A0			I SKEN KOME ····					
										7		
PtWRKY-IIA	Ba EDDSVLVASY	DGEHNHE PNG	SHGQYL CS	HTSSKISITN	HVL	EIPPLOPSIALDLT	LS NQQKEN	PSKRSMEDCG	K INNNCNKN <mark>Y I E</mark> E	YVASLTKDPTFSV	ALAAAVASSM <mark>SD</mark> L	
VVWRKY-IIA	32 EDS <mark>S</mark> ILVA <mark>TY</mark>	EGA <mark>HNH</mark> E <mark>PP</mark> -	- HDA <mark>P</mark> AG GS	SYS <mark>P</mark> S <mark>P</mark> IK <mark>G</mark> LV	AN F <mark>P P</mark> T	T V D P F Q P T V T L D L T I	L S <mark>G T G</mark> QEKRR		KSN CGRR - VEE	YVASL <mark>TK</mark> DTNFTL	ALAAAVARS I TOQ	ſ
GmWRKY-II	132 H <mark>dks</mark> ivva <mark>ty</mark>	DGEHNHAAI -	- HES S	S S T S <mark>G</mark> S S <mark>P</mark> V V N	N - L <mark>P</mark> MT	SILNDKE <mark>PMNIGL</mark> AI	LS <mark>G</mark> -QTDHRR	HCEDAMQQNN	NN <mark>GG</mark> SNIR <mark>IE</mark> E	Y	IS LAEAVVRT I TGQ	
GmWRKY-II	130 HDKSILVAIY	DGEHSHGPL -	- DES S	S S T <mark>P G</mark> S S <mark>P</mark> V A N	NKL · · · · · · · PMT	SV <mark>P</mark> NYIEAMNIDLA	LSG-QTD-RR	HCEDVMQQNN	NN <mark>GG</mark> SNIK IEE	YVSSLIKDPDFTM	SLAEAVART I TOQ	IIA3-
MWRKY-IIA	32 HDRSILVATY	DGEHSHGVP -	- NESFK PS	SST <mark>PG</mark> SS - ISN	K L P	TRUSDKE	IS	ICENVMQQFG	VER HIK IEE	YASSLIKDPDFTA	ALAEAVARTITOQ	
GIII WKK 1-III	ASE QUESTLUAT	EGRANNOVE -	- HULLK <mark>-</mark> 3	551 <mark>F</mark> ISIMINN	L <mark>F</mark> · · · · · · · · · · · · · · · · · · ·			·····LCDEIK····		TASPLUNOF	- CHERVVHJERJU	
OsWRKY-IIA	42 EDNTILVATY	EGEHNHGQ P P	PPLQ	S A A Q N S D <mark>G</mark> S <mark>G</mark> K	SAGKPPHA	<mark>Р</mark> ААА <mark>РРАР</mark>	VV <mark>P</mark> HRQHE <mark>P</mark> V···	• VVNGEQ - QAA	AASEMIRRNLAEO	MAM <mark>T L T R D P</mark> S F K A	ALVTALSGRILEL	
BdWRKY-IIA	42 EDRTVLVATY	EGEH <mark>NH</mark> GOPP	PAQP	QQQHDGKNGA	АК <mark>Р Р</mark> Q А	P VAAAAALHQQI		VVNGEAAA	AASEL I RRNLAEO	MAM <mark>TLTMDP</mark> SFKA	ALVTAL <mark>SGR</mark> IFEL	
SBWRKY-IIA	42 DDSTVLVA <mark>TY</mark>	E G E <mark>H N H</mark> G Q P P	QHD <mark>GG</mark> RAARST.	AT AQAQ VASEA	AV	R <mark>P VAAP L P L</mark> QI	I <mark>P</mark> HQQQKQEAA		SSEVA - RKNLA <mark>E</mark> H	MAV <mark>T L T R D P G</mark> F K A	ALVSAL <mark>SGR</mark> IL <mark>E</mark> L	[r
ZmWRKY-IIA	42 DDR <mark>T</mark> ILVATY	E G D H N H A Q P P	RHNGG GSK	AATRST <mark>P</mark> AATE	AA	· · · · R PAPPPL PLQI	IQHQHQQKQEA	AT <mark>GP</mark>	SSEVA - RKNLAEH	MAVTLTRDPGFKA	ALVSALSGRILDL	
BIWRKY-NA	40 UDKALLVATY	EGDHNHAQPL	AIG	· · · · · TAAKNS	A5	KPAP - PLGLQ	QQKKQQQEAA	AMAA	ERE LAERKNLAEO	MAATLTRDPGFKA	ALVSALSGRILEL	11A4-
USWKKY-IIA	40 DUNIVLVATY	EGEHNHAQPP	HHUAG	· · · SKI AAAAK	пэ	QHQPPPSAAAA	AVVRQQQEQAA	· · · · · · · · · · · · · · · AAGP	SIL VAARKNLAEQ	MAATLIKUPGFKA	ALVIALSGRILEL	

Appendix 3

	СС-В		VPROF
	20 30 40	60 70 80 90	100 110
PtWRKY-IIB1a ELAVLR	· · · · AEVERMKVENL <mark>RLKG</mark> MLNHVTSN <mark>Y</mark> NAL	O · · · MDLVTLMODQNSHH · · · · · DGKNKDDG · · · · · · · · · · · · · · · · · ·	······································
GmWRKY-IIB1a EMADLQ	· · · · GDLERIKRENQKLEDTLDEVNTNYNAL	. <mark>a</mark> · · · · MH F MN MM <mark>a</mark> E RK · <mark>a</mark> E · · · · VS d <mark>a</mark> K VK E KK a <mark>a</mark> a S <mark>a a</mark> · · · · · · · · ·	·····
GmWRKY-IIB1b EMAVLQ	ADLERMKRENQKLRDSLDEVTTNYSAL	Q · · · MH FMN LMQE RK · GE · · · · EVYGGEKKQQLGESGGD · · · · · · ·	
VyWRKY-IIB1a KMVGLR			······································
GmWRKY-HB14 EFIALL	TELDOMNAENORLRELVDOVNNCNTL	R MQ L VK L T QKHHSHG KED	MVPRSFLDIGVAEKDE
GmWRKY-#81e KEVAML	· · · · SELGEMNAENORLRELVHNLNNKYNAL	H · · · KD LMK L THKQHENE · · · · · · · · · NDKRDD · · · · · · · · ·	·····MIIPRSLLDIGIATKED···
GmWRKY-IIB11* KFAAML · · · · · ·	AELHIINAENQHLRELVDQVNNKYNDL	HKDLTKLMQKQHKNEKDKRDD	MIISRSFLDIGIATKED
CpWRKY-11816 KLMVLR	· · · · AE LNRMKAENQ <mark>R</mark> LTGML DE ANHN <mark>Y</mark> NA L	Q · · · AH FAS LMORROSON · PL · YLYPYMING I QQR · · · · · · · · · · · · · · · · · ·	· · · · · · RHGARQFLDLGQAAMADE · ·
OSWRKY-IIB1a ELAAMQ	A <mark>E L GRMNEEN</mark> O <mark>R L RGML</mark> T O V T T S <mark>Y</mark> Q A L	Q · · · MH L VALMQQRP · · · · · · · · · PPP HQDGKAEG · · · · · · · ·	· · · · · AVVPROFLDLGPS · SGA · · ·
SBWRKY-IIB12 NTNEMQ	A <mark>E LARMN DE N</mark> O <mark>R L RGML</mark> TO VTS S <mark>Y</mark> OA L	<mark>.0</mark> MH <mark>LVALM</mark> OARAGGQPPTTDGAAAAV	····· MPLPROFLGLGPA······ ITD1
BdWRKY-IIB1a ELAAMQ	A <mark>E LG</mark> RMNEE <mark>N</mark> Q <mark>R LRGML</mark> TQVTNS <mark>Y</mark> QA L	. <mark>Q</mark> · · · MH <mark>LVALMQ</mark> QRT · · · · · · · · · PPTHED <mark>G</mark> KIE <mark>G</mark> · · · · · · · ·	····· AIVPROFLDLGPSGAGA··· IIDI-
VvWRKY-IIB1b EVAQLQ	V <mark>E</mark> LERMNA <mark>EN</mark> Q <mark>K L<mark>RG</mark>MLNQ VTNN<mark>Y</mark>ST L</mark>	. <mark>Q</mark> MH L VT L MQQQSQQN V I L KHE VA <mark>G</mark>	·····
PtWRKY-IIB1b ELAQLQ	VELQKMNAENQ <mark>RLK</mark> DMLSQVTNN <mark>Y</mark> SAL	Q · · · MH F V A L I Q Q · Q Q R N · · · · · · · · E E K K H E · · · · · · · · · · · · · · · · · ·	·····
PtWRKY-IIB1¢ KLAQLQ	MELQKMNTENQ <mark>R</mark> L <mark>K</mark> DMLGQVTTS <mark>Y</mark> SAL	O · · · MHFAALMOQHQQQN · · · · · · SEEKKHEDV · · · · · · ·	·····VVPROFMDLGP····SA···
GmWRKY-IIB1y ELAQLQ	· · · · VELQRMNAENKKLKEMLSHVTGNYTAL	.a MHLVTLMQQNQ · QR · VEDKNHGVGG ·	······GKVPROFLDIGPSGTAE···
GmWRKY-IIB1h ELAQLQ	· · · · VELQRMNAENKKLKEMLSHVTGNYTAL	Q MH L VT LMQQNQ - QR VE DKNHG VGG	·····GKVPROFLDIGPSGTAE···
GmWRKY-IIB11 ELARLK	· · · · EDL RNMNAENQKLKEML SHVSSNYANL	Q · · · MH LAAVLOQQQNQR · · · · · · · AEERKHGVGG · · · · · · ·	······GTVPROFLSLVPS···E···
AtWRKY-WR12 ELVELO	DELKKNTHONOKI PELLTOVSNSVTSL		TIVPPOFIDICETRAVC
AtWRKY, IIB 16 ENAOL R			ENVPROFIDI GP
AtWRKY-IIB1c ENAOLO	····EELKKMKIENORLEDMLSOATTNENAL		·····
GmWRKY-#81k* OMGMLR	····IKLEEAKKENEILKAMLNOVNEHCTTL		·····
OsWRKY-IIB22 AATTVE	<mark>GELRQAGEEN</mark> R <mark>RLR</mark> RRLEELTSS <mark>YG</mark> AL	YHQLVQAQQLHTKHQ <mark>P</mark> IAGVQLLD	ALAAA <mark>SP</mark> A <mark>S</mark>
ZmWRKY-IIB2a DVAAVE	<mark>GELRQAGEEN</mark> R <mark>QLR</mark> RMLEELTRS <mark>YG</mark> AL	YH · · · Q L I Q A Q A A Q H H H H · · · · · · · · · · P T S G V Q F M D · · · · · · · ·	· · · · · · · · <mark>P</mark> AGRV <mark>PPPP</mark> · · · · · · · · · · · · ·
BdWRKY-IIB22 AAATVQ	A <mark>E L</mark> RRVV <mark>E EN</mark> L <mark>R LRGML</mark> E E LT RS <mark>YG</mark> A L	YH · · · Q L L Q V T Q Q Q Q H P H · · · · · · · · R S P L T Q T H L T · · · · · · · · · ·	A <mark>P</mark> N <mark>TS</mark> A <mark>T</mark> - QQFLE <mark>P</mark> R
OSWRKY-IIB20 TTAAVD	· · · · VELRRVVEEN <mark>R LR</mark> GML DE LNRS <mark>Y</mark> SAL	YH · · · Q <mark>YL</mark> QVT <mark>Q</mark> QQNHRH · · · · · · · R <mark>P</mark> SLAQTHRT · · · · · · · ·	· · · · · · · · · AAT <mark>TT</mark> A <mark>T</mark> TQQFLE <mark>P</mark> R · ·
SBWRKY-IIB22 EASAVE	VELRRVVEENR <mark>R</mark> L <mark>R</mark> GMLEELNRS <mark>YG</mark> AL	YQ · · · QLLQVT Q · · · · · H · · · · · · · RSSLAHTHLTT · · · · · · ·	· · · · · · · TAASHNTSSTRQLLEAR · ·
ZmWRKY-IIB2b EASAVE	· · · · AELRRVVEENR <mark>R</mark> L <mark>R</mark> GILEELTRS <mark>YG</mark> AL	YQ · · · QLLLVT QGQHHQH · · · · · · · RPSLAHTHLTS · · · · · · ·	· · · · · · · TAAS <mark>QYS</mark> A <mark>S</mark> TRLLLEAR · ·
AtWRKY-IIB2a QISRLK	· · · · LELERLHEENHKLKHLLDEVSESYNDL	QR · · · RVL LAROTQVEGL · · · · · · · · · · · · · · · · · · ·	
CPWRKY-IIB2a ELSSLQ	· · · · KELERLHEENRKLRSTLDQITKKYSEL		
		DC	
MWRKY,UR2a EMSILE.	SEL REVOEENHKI RIMI FOITKSYSOL		VSFOIFINN.N
GmWRKY-IIB2a LLNTLQ		0LFIALOKOKOCO	
GmWRKY-#828 LLTTLQ	···· RESLRIKEENCKIRTMIDGITKNYNGI	0 LF ALOKOKOCO OKMETNLNGM	MFGQHLLDPRG
GmWRKY-IIB2c ELSSLE	NOLGKLOEENSKLRSVLDQITKNYAQL	QAQLFNALQKQKNVETK-IVD	
GmWRKY-IIB24 ELSSLE·····	· · · · SELLRLQEENNKLRNVLDH I TKSYTQL	QAQLFIALQNQNMETK-IVD	· · · · · · · · <mark>P G</mark> T S R K <mark>L</mark> D V N · · · · · · · · ·

	СС-В		acidic-2	
PuWRKY-IIR1a		HSDST	NSDRENLRAGSPLKSTV	
P#WRKY-II3b	ELMAAR · · · · · · · · TEITRLSEENDRFRSMLTQLSTEYRNLQ · · · IHVVASMQR	SSDTL	NPDIKKFHGLSSPRSAP	
PtWRKY-IIB3a	ELSGLQ · · · · · · · · · · · · · · · · · · ·	QQNMD	KDPQ I SLGLNANDNKAVQE	
VvWRKY-IIB3a	ELCVLQ · · · · · · · · · · · · · · · · · · ·	QQNK	• • • • • • • • KDLQ <mark>S L</mark> S L H <mark>G</mark> K D • RN L Q D	
AtWRKY-IIB32	ELESAK · · · · · · · · · AEMSEVKEENE <mark>K</mark> L <mark>KGML</mark> ER IESD <mark>Y</mark> KSLKLRFFD I IQQE <mark>P</mark> S	NTATK	DQER <mark>ELV</mark> S <mark>L</mark> S <mark>LG</mark> R	
PtWRKY-IIB3b	VIESAKS <mark>EMG</mark> DVR <mark>EENHR</mark> L <mark>K</mark> MMLERIEKD <mark>Y</mark> QSL <mark>Q</mark> LRFFDILQ <mark>H</mark> ETS	sк <u>q</u> sт	EESN <mark>ELV</mark> SLCL <mark>G</mark> R	
GmWRKY-IIB3a	KLESAK · · · · · · · · · AEMSKVKEENE <mark>R LK</mark> MM I ER V <mark>G</mark> KN <mark>Y</mark> HSLOLRFFD I LHRETS	NK <mark>G</mark> VE	· · · · · · · VEE <mark>P</mark> KLV <mark>SL</mark> CL <mark>G</mark> T · · · · · ·	
GmWRKY-IIB3b	KFKSAK • • • • • • • • • • T <mark>EMG</mark> EVK <mark>EENER</mark> L <mark>K</mark> MMLERVEKD <mark>Y</mark> HSL <mark>O</mark> LRFFDILHKDVS	KK <mark>G</mark> LA	· · · · · · TAE <mark>PELV</mark> SLCL <mark>G</mark> R · · · · · ·	
GmWRKY-IIB3c	K IKS AK • • • • • • • • • T EMGEVKEENER LKMMLER VEKDYHS LOLRFFD I HHED VS	KK <mark>G</mark> LA·····	· · · · · · TEE · · LVSLCLGR · · · · ·	
MWRKY-IIB3b	K <mark>P</mark> KC <mark>S</mark> K · · · · · · · · · EEMGEVKEENE <mark>R K</mark> TML SRVEKD <mark>Y</mark> NS LO RFFD VNKDVS	MKD E · · · · · · · · · · · · · · · · · ·	· · · · · · DEE <mark>PEFVSLCLG</mark> R · · · · ·	
PtWRKY-IIB3c	QLESAR · · · · · · · · · AEMGE VRKENQR LK IHLDR VVKDYRT LQVQFYE I IQQEET	KKSTD	· · · · · · · TEEH <mark>ELV</mark> SLTL <mark>G</mark> R · · · · · ·	
GmWRKY-IIB3d	QLETAK · · · · · · · · · AEMGVVREENQ <mark>R</mark> LKMCLNK IMNEYRT LEMQFQD I LKQQGT	KKNVD	· · · · · · LEESDLVSLCLGR · · · · · ·	
GmWRKY-IIB3e	QLETAK · · · · · · · · · AEMG VAREENQLIKTCINK IMNEYRTLEMQFQDILKQQGT	KKNAD · · · · · · · · · · · · · · · · · · ·	·····LEEADLVSLCLGR·····	
MWRKY-IIBJ2		KKNND	NAESDLVSLSLGR	
GRWRK 1-IIBJI		KUSSU	MERCELLELER	
CPWRK T-IIB32		IKKS <mark>P</mark>	FFFTDLLELELE	
GmWRKY-HD3#		KRSTD	DEFEIVELEIG	
GmWRKY.ttR1b			NEESELVSISIG	
GmWRKY-IIR3i		FDATN	NEESELVSI SIG	
OSWRKY-IIB3a		T	·····DDEPDDLVSLSLC·····	
BdWRKY-IIB3a	KLEATK · · · · · · · · · · · AEMGS VREENE RLKTLLSQI VRD VQSLHTHFLAAVKPKSS	0 P 0 P 0	AAAEDLVSLSLGR	
ZmWRKY-IIB3a	RLEAAR · · · · · · · · · · · AEMGE VREENER IK SML SR I VGQ YQSL QTH FLD VVK VRE L	ASSSA	DE <mark>P</mark> YDDLVSLSLG	IIB3-
VvWRKY-IIBJc*	·····	KESTE	· · · · · · · IEE <mark>PEFVSL</mark> SL <mark>G</mark> R · · · · · ·	
AtWRKY-IIB36*	·····	<mark>Р</mark> ТКFQ	• • • • • • • • • • • • • • • • • • •	·
VvWRKY-IIB3d*	· · · · · · · · · · · · · · · · · · ·	skr <mark>p</mark> v	SKEL <mark>ELV</mark> S <mark>L</mark> C <mark>LG</mark> R	
OsWRKY-IIB3b*	· · · · · · · · · · · · · · · · · · ·	к	· · · · · DIEE <mark>PELV</mark> SLRL <mark>G</mark> · · · · · · ·	
ZmWRKY-IIB3b	R L E A A R • • • • • • • • • • • A <mark>E M G</mark> E V R <mark>E E N E R</mark> L <mark>K</mark> S <mark>T</mark> L S R I V S Q <mark>Y</mark> Q S L <mark>Q</mark> M H F L D V V K V H <mark>E</mark> Q	A - S A A	T D D <mark>P</mark> D D L V S L S L <mark>G</mark>	
ShWRKY-IIB3a	R L E A A R • • • • • • • • • • A E MG E V R E E N E <mark>R L K</mark> SML SR I V SQ <mark>Y</mark> Q S L <mark>Q</mark> MH F L D V V K VQ E Q	A - S S A	· · · · · DD <mark>GP</mark> DDLVSLSL <mark>G</mark> · · · · · · ·	
BdWRKY-IIB3b	RLASTR · · · · · · · · · AEMGE VROENER IKT MLSR I IEDHRSLOKH FND VLOOGRE	к	····DVEET <mark>ELVSL</mark> SLG····	
AtWRKY-IIB3c*	ELLQLQ · · · · · · · · · I QME · VKEENT RLR · · · KLVEQ · · · TLEDYRHLEMKFPVI	DKTKK	LEMFLGVQGKRCVDITS	
GINWRKY-IIB3J*				
GRWRK T-IIBJR			KETEVEL SLODTATISCEG	
BAWRKY-IIB3C*	ELSENG		KET DVELSLGAT AAABAGC	
OsWRKY-IIB34	TFTLFR		SLGLSLGS SPDACOCH	
ZmWRKY-IIB3c	TTLVFK		VLSLSLGPAANANRAS	
SHWRKY-IIB36		D-EPAAGP	FLSLSLGP - ANALGAS	
ZmWRKY-IIB3d		··· · ·· ·· ·····		
GmWRKY-IIB3	V L V T AK · · · · · · · · · Y Q R AKWQKSNKR LERLEN V D D L I S Y Q M I F K K H N I R H M K N G	I H C T S · · · · · · · · · · · · · · · · · ·	· · · · · · · SDE <mark>P</mark> DLAEAELDD · · · · ·	
AtWRKY-IIB3d	ELDATK · · · · · · · · · AKVEK VREENEK LK LLLST I LNN YN SLOMQVSK VLGQQOG	A S S ME	· · · · · · NDYDVD I S L R LG · · · · · ·	
ZmWRKY-IIB4a	ELAL <mark>T</mark> K·····SEL <mark>G</mark> RLNEENKOL <mark>K</mark> NILTRLTSN <mark>SNPLO</mark> ···MQMQALTTM	QQRTSKDQE <mark>G</mark> SR <mark>GGG</mark>	· · · · · · HLLPQOFIGLSTPALSFDDP	
ZmWRKY-1184b	ENKQLN - · · · · · · · EENKQLNEENK <mark>Q</mark> L <mark>R</mark> SMLSRLTTSFN <mark>PIQ</mark> K <mark>P</mark> SMQMQLLSLM	QQQAT • • • • • • • • • DDQKASR <mark>GG</mark> A • • • • • • • • • • • • • • • • • • •	· · · · · · HTLPQQFI · · · SVGTAPDDL	
ShWRKY-IIB4a	ELALAK······SEL <mark>G</mark> RLNEENKOLKOMLSRMTIKFNAFO···VQMPVYTTL	MQQQQ - · · · · · · · DHQE <mark>G</mark> S - <mark>GG</mark> S - · · · · · · · · · ·	· · · · · · HLLPROFI · · SSLGTAPDDP	IIIB4- II
BdWRKY-IIB4a	ELAATK · · · · · · · · · SELARVREENK <mark>R</mark> LKSMLSSANSKCNSLH · · · MHL <mark>T</mark> HLQQQ	QQRSSHLDQL <mark>P</mark> L <mark>P</mark> TTTA	· · · · · LNMPROFISLGSA · · · PDEP	
OSWRKY-IIB4a	ALAVTK AEIGRLSEENKRLKNMLSNVTTKYNSLQ MQFVTLMQQ	RRS EQE <mark>G</mark> SQQQQQ	· · · · · · QLI <mark>P</mark> R <mark>QFI</mark> S <mark>LGS</mark> ASLQ <mark>P</mark> DVE	


	120	130	140	150	160	170	180	190	200	210	220	230
P#WRKY-IIB3a				GLVHALN - SQV	KVEVTT						R S R P R T G S A S I	PERSTSP
P#WRKY-II3b			QEQIAKQM	IGTRIFY-NQS	Q Q D T						KACQLQQRAS	vно <mark>s</mark>
PtWRKY-IIB3a	V <mark>P</mark> KAII		- PRQSGSSYI	QRHQAAASTGL	SLRLQITST					. s aa	EREEDMEENN	REDQTAN
VvWRKY-IIB3a			- PRRISKVLN	INDQILPSSGL	5 L R L K <mark>P</mark> N					TRE	EREEDGEANKI	Е <mark>Е</mark> ТVS
AtWRKY-IIB3a	RSSS <mark>P</mark> SD	KE	EKTDAISAEV	NADEELTKAGL	LG - INNGNGG	E <mark>P</mark> K					- EGLSMENRAN	NSGSEEA
PtWRKY-IIB3b	S - <mark>P</mark> S E	EE	KST NSAKS	RENE - ELKANL	LG - LDSKILT	s					- TETASNPSP	AE-SVEE
GmWRKY-IIB3a	• • • • \$ <mark>9</mark> • • • • • •	KD	- <mark>G</mark> IICNSSKH	KENEDLEAS - L	r L <mark>g</mark> - LDCK <mark>g</mark> vs	SKEQ·····					vsd	MNTSEEK
GmWRKY-IIB3b		KE	LAR I GYSNKP	KE - EDV <mark>gp</mark> n - L'	T L <mark>g</mark> - L D S K H L F							S <mark>E</mark> E <mark>P</mark> KEV
GmWRKY-IIB3c		KE	- AR I GNSNKL	K E D V <mark>g P</mark> N - L'	T L <mark>g</mark> - LDSKHLL	SMEV					VSDFS <mark>P</mark> MNS- 1	S <mark>E</mark> Q <mark>P</mark> KEA
MtWRKY-IIB3b	S <mark>P</mark> N	к р	AKN I EN <mark>P</mark> NK <mark>P</mark>	KEKEDMEVN - L:	SL <mark>g</mark> ·LDSKYMV	PMEL·····					V S D L S <mark>P</mark> MN S S I	E <mark>ELP</mark> KEV
PtWRKY-IIB3c	· · · I S S · · · · ·	RD	GKNNKTSSQG	KNHDEQVKESL	S L <mark>G</mark> - S L C T F E A	SKSA·····				т	NET L <mark>P</mark> N <mark>P</mark> SI	PVNSFG-
GmWRKY-IIB34	••••VPT•••	• • • • • • DE	KI•KVSNK <mark>P</mark> L	KDDE		<mark>G</mark>				F	NNEELTL <mark>G</mark> ·	LDCEV
GmWRKY-IIB3e	• • • • V <mark>P</mark> • • • • • •	DE	KI-KVSNK <mark>P</mark> L	KDDE <mark>g</mark> l • • • • • *	ſ L <mark>g</mark> ∙ L E C K F E T	SKS <mark>G</mark>				s	TNEAL <mark>P</mark> NN <mark>P</mark> SI	PEN <mark>S</mark> CEV
MtWRKY-IIB3a	• • • • v <mark>P</mark> SNN I • •	QE	KVNKVSKLAL	NNDEEFNKEEL	S L <mark>g</mark> - L E C K F E T	sкs<mark>g</mark>				s	T T E <mark>G L P</mark> N I <mark>P</mark> S I	P VN <mark>S</mark> S E V
GmWRKY-IIB3f	• • • L P T • • • • •	· · · · · NE	KV···NNK <mark>P</mark> L	KEEEKEDKE <mark>g</mark> l:	SL <mark>g</mark> - LDCKFET	SKS <mark>G</mark>				s	TTEHL <mark>P</mark> NQ - S	PNN <mark>S</mark> VEE
CpWRKY-IIB3a	V S N	RK	DYEKNKE	KQDEKDKLE <mark>g</mark> l:	SL <mark>G</mark> - LNYKYE <mark>v</mark>	sks				L	D T A L <mark>P</mark> N S S I	PANSISD
VvWRKY-IIB3b	VSSAE	DD	K K T S F L S <mark>G</mark> K <mark>G</mark>	K <mark>g</mark> de - Kmde <mark>g</mark> la	AL <mark>G</mark> - LECKFE <mark>P</mark>	A <mark>P</mark>					- TEHMMNAS <mark>P</mark> I	EN- <mark>S</mark> FE <mark>G</mark>
GmWRKY-IIB3g		E	MKKKKNRNEK	MRENEDLKDIL	AL <mark>G</mark> · LDIRFD ·						SSAIKNLS	T <mark>E</mark> S <mark>S</mark> C D <mark>G</mark>
GmWRKY-IIB3h		EK	MINN <mark>G</mark> IEK	REDEDVH-KRL	V L <mark>G</mark> - L D I N L D <mark>P</mark>	VDQD · · · · · ·					· · ELAANNST	PES <mark>S</mark> F-G
GmWRKY-IIB3i	· · · · <u>·</u> · · · · · ·	<mark>E</mark> K	ILKNKNGIEK	TEDEDVHNKRL	VL <mark>G</mark> - LDINLD <mark>P</mark>	VDQD · · · · · ·					· · ELTANNST	f <mark>E</mark> S <mark>S</mark> F V G
OsWRKY-IIB32	• • T R P NAAA • •	<mark>P</mark> S	5 <mark>66</mark>	<mark>gg</mark> dd <mark>g</mark> RL:	S L <mark>GG</mark> VAS D D D D	DKQ				ASRRA	L <mark>PP</mark> M <mark>P</mark> VLNLS	S <mark>D</mark> S <mark>S G</mark> D A
BdWRKY-IIB3a	S R <mark>P</mark> D A <mark>G</mark> A G -	AS	s s s <mark>g</mark>	TDTDRDDRL	5 L <mark>G -</mark> NDDD	DK				· · · · · · · KAT	· · · · <u>·</u> LLNLS	S G S <mark>S</mark> S A D
ZmWRKY-IIB3a	• • T RANSSGG •	· · · · · AS	S S <mark>G</mark> T ADE M <mark>P</mark> N	T - ADDERRHQI	S L <mark>GG</mark> T AAAADD	DK				• • • • • ASHDA	ST - APVLNLS	S D S <mark>S G</mark> S A
VvWRKY-IIB3c*	· · · V S S · · · · ·	· · · · · KD	EKNKTTS <u>-</u>	KVEDD <mark>g</mark> vk <mark>gg</mark> l:	5					<u>.</u>	LN <mark>P</mark> S	PEN <mark>S</mark> FG -
AtWRKY-IIB3b*	NSEV <mark>P</mark> S · · · · ·	· · · · · E	KNKDVEEAE <mark>G</mark>	DRNYDDNEKSS	IQ <mark>G</mark> - LSM <mark>G</mark> IEY	KALS				N <mark>P</mark>	NEKLEIDHNQI	ETM <mark>S</mark> LEI
VvWRKY-IIB3d*	S S <mark>P</mark> T D	DG	KSS <u>I</u> ASKA	KEDDDELNA <mark>g</mark> l'	r l <mark>g</mark> · l d sk f q v	SKL				D	VT E FASNSS P	F <mark>ens</mark> iee
OsWRKY-IIB3b*	· · T ST SK	· · · · · TT	S S E VK <mark>G</mark> S T E D	F L K I K <mark>G</mark> <mark>G</mark> L :	5 L <mark>G</mark> • • • C R V D A	NN				· · · SEKV	- Q P D V M T - L S I	PEG <mark>S</mark> FED
ZmWRKY-IIB3b	• • T RANS <mark>GGG</mark> •	· · · · · AS	S S <mark>G</mark> TADEMTT	A <mark>G</mark> ADDE <mark>G</mark> - HRL:	S L <mark>G</mark> S - T V A T D D	DK • • • • • • • • •				· · · · · ASHAS	· · · · · VRNLS	S D S <mark>S G</mark> S A
ShWRKY-IIB3a	T R A N S <mark>G G</mark>	· · · · · s s	SSGTAETTTA	ADADDQ <mark>G</mark> HHQL:	5 L <mark>G</mark> STTTATDD	DK				· · · · · ASHAS	T APVLNLT	S D S <mark>S G</mark> S A
BdWRKY-IIB3b	· · T S T S T · · · ·	···· SТ	I AE <mark>G</mark> Q <mark>G</mark> RQ L <mark>G</mark>	L L K I RQ <mark>G</mark> AA <mark>G</mark> I :	5 L <mark>G</mark> S Q L S <mark>G</mark>	АТ				DQKV	P R P D V L L S L S I	<mark>P E G </mark> S E E
AtWRKY-IIBJc*			· <mark>-</mark> к	ARKR <mark>G</mark> AERS <mark>G</mark> L:	SLSLEKK					о ко	EESKEAVQSHI	HQRYNSS
GmWRKY-IIB3j*			- <mark>P</mark> SRILEIFN	KQM - QSA <mark>P</mark> S <mark>G</mark> L:	SLRLQ <mark>P</mark> S					тsн	HKESDV <mark>G</mark> NN - H	KEDKNDQ
GmWRKY-IIBJk*	• • • • • • • • • • • • •		- PSRINEIFN	KQ I RQSA <mark>P</mark> S <mark>G</mark> L:	SLRLQ <mark>P</mark> S					тsн	HKE S D V <mark>G</mark> NNNH	K <mark>e</mark> dkndq
OsWRKY-IIB3c*			<mark>GGGFP</mark> EA	KSKEQAAWR <mark>G</mark> L:	S L S L <mark>G</mark> A S S S					· · · · · · · YDD	DQK - AVEAR <mark>P</mark> I	HDVDGAA
BdWRKY-IIB3c*			<mark>GP</mark> EQQQA	E A R R <mark>P</mark> S A G S G L :	SLSL <mark>G</mark> ASSS					· · · · · · · YED	QEKT <mark>G</mark> LEAAA	PGMAMS I
OsWRKY-IIB3d		· KD · · · · ·	· · · · EADA <mark>G</mark> N	<mark>ggg</mark> Dgy L/	ALALRCA <mark>P</mark> · · ·					<mark>AAG</mark> E <mark>P</mark> MVH	PKRQRATTNS	5
ZmWRKY-IIB3¢		• E <mark>G -</mark> • • • •	- RR <mark>G</mark> E I SASN	DADD <mark>GTG</mark> LI	AL <mark>G</mark> LRC					· · <mark>666</mark> · · · · ·	••••••••••••••••••••••••••••••••••••••	RORVVIV
SBWRKY-IIB3D 。		- R <mark>GG</mark> EAIN.	ASQA <mark>PP</mark> HAAS	NADD <mark>G</mark> <mark>G</mark> L /	ALGLRCDSD					<mark>GGG</mark> E <mark>P</mark> VLA	AVV <mark>G</mark> SAA <mark>G</mark> TKI	RORVAIL
ZmWRKY-IIB3d				· · · · · · · · · · · · · · · ·	· · <mark>·</mark> · · · · · · · · ·						MAKKR/	
GmWRKY-IIB3I	SLCLLAFS	AFSNSCEH	ITTNTDTFEE	R D R K Q S R R N K <mark>G</mark> I	MR <mark>G</mark> ILKIIEKL	I <mark>G</mark> EDEVE <mark>G</mark> VG	L VN FN			THVVL	PLEYAARNVT	MESLY <mark>P</mark> Q
AtWRKY-IIB3d		· · · · · · KK	EENKVDKIST	KNVE <u>ESK</u> DKRS/	4 L <mark>G</mark> - F <mark>G</mark> F Q I Q S	YEAS					· · · · KLDDLC	ROVKLAN

MMPLP

IIB3-

ZmWRKY-IIB4a	LRF•VASDVQ••••••••••••••••••••••••••••••••••••	
ZmWRKY-IIB4b	LRS · V <mark>G</mark> SHALR · · · · · · · · · GDDC SAST SNAE <mark>PPPP MASGKDNMP LP · AFELG NQ · · · · · · · · · · · · · · · · · · </mark>	
SbWRKY-IIB4a	LRS · V <mark>G</mark> SDAMH · · · · · · · · <mark>GGG</mark> NSS <mark>G</mark> SSTSNAE <mark>PPPP</mark> MVSSKEMM <mark>PLP</mark> · AFEH <mark>G</mark> HQQ <mark>P</mark> Q · · · · · · · · · · · · · · · · · · ·	\parallel IIR/
BdWRKY-IIB4a	<mark>PP</mark> ···L <mark>P</mark> ARAS······NGSLDCA <mark>P</mark> SSSN <mark>P</mark> VGVD·GSKAADUV <mark>PVPP</mark> AFDYHHHHG·································	IID 4 -
OsWRKY-IIB4a	A <mark>P</mark> ····HSVVVV········GGDV·CAPSSSN <mark>P</mark> D·····AAV <mark>P</mark> AMM <mark>PLP</mark> ·HFDHHNHHH <mark>P</mark> ··································	





	WRKY DNA	-binding domain	n								
									RyyLPC		
	360	370	380 390	400	410	420	430	440	450 460		
PtWRKY-IIB1a	YYRCTMAAG CPVQ	RCAED - RTILT	TTYEGNHSHPLPPAAT	AMA <mark>STTSS</mark> AARML	L <mark>S</mark> G	S-MSSTDG			MNS-NFLTRTI-		
GmWRKY-IIB1a	YY <mark>RCT</mark> MAAGCPV <mark>RKQ</mark> VQ	RCAED · RTVLI	TTYEGNHNHPLPPTAM	AMAQ <mark>TTSS</mark> AARML	L <mark>S</mark> G	<mark>s</mark> - M <mark>s</mark> sads			MNA - DFLTRTL -		
GmWRKY-IIB1b	YY <mark>RCSMA</mark> SACPVRKQVQ	RCAED - RTVLI	TTYEGNH <mark>NHP</mark> LPPTAM	AMAQ <mark>TTSS</mark> AARML	L <mark>S</mark> G	S-MSSADS			MNA - NFLTGTL -		
MtWRKY-IIB1a	YYRCTMALGCP VRKQ VQ	RCAED · KTILI	TTYEGHH I HALPPAAM	IE MVQ <mark>TTSS</mark> AARML	<mark>L S</mark> G	P - MT SADG		· · · · · · · · · · · · · · ·	. MN <mark>P</mark> · N <mark>YL</mark> T RA I ·		
MWRKY-IIB1b	CYRCSTSRGCS I OKOVO	RCALD - RTVAI	TTYEENRNLPLPAAAK	EMVQ <mark>TTS</mark> AAAKML	L <mark>S</mark> A	S-TSSNDG		••••••	LNA - N <mark>LL</mark> TRT <mark>P</mark> -		
GmWRKY-IIB1c	YYRCTMAAGCP VRKQ VQ	RCAED - RTILI	TTYEGNHNHP LPPAAM	IAMAQ <mark>TTSS</mark> AARML	L <mark>S</mark> G	S - MS SADG			. MNA - SFLTRTL -		
CPWRKY-IIB1a	YYRCTMAAVFETERQVQ	RCAED - RTILI	TTYEGNHNHPLPPAAM	AMASTTSSAARML	· · · · L SG	A - MSSADG			MNS - NFLARTL -		
VVWRKY-IIB1a	YYRCTMATACPVRKQVQ	RSAED - RTVLI	TTYEGHHNHPLPPAAM	AMASTTSAAATML	L <mark>S</mark> G	<mark>S</mark> - M <mark>P</mark> S S D G			IMSSSFHSRTMF		
GIIIWRKT-IIB10	YYRCEMOT ACRYPKOVO	REAED OFVIL	TTYECOUNNYL BRTAK	AMA STTEAVENI	1.00				LINDNILESTAS		
GmWRKY-UB4#	XXPCSMGTACPVPKQVQ	PNAED. LEVI I	TTYEGOWNWYI PPTAK	ALASTTSAAASMI		S. MISSDG			LIVPNILES.AS		
CuWRKY-UR1b	YYRCTMATS CPVRKOVO	RCAED KTIL	TTYEGNHNHPLPPAAM	AMASTTSAAASMI		S.MTSIDG	P				
OsWRKY-IIB1a	YYRCTMATGCPVRKQVQ	RCAED - RSILI	TTYEGTHNHPLPPAAM	AMASTTSAAASML		S - MP SADG	A		LMSSNELARTV-		1
ShWRKY-IIB1a	YYRCTMANGCP VRKQ VQ	RCADD - RSILI	TTYEGTHNHPLPPAAM	AMASTTSAAASML	L S G	S-MPSGD-			MMT SNFLARAV -		
BdWRKY-IIB1a	YYRCTMATGCP VRKQ VQ	RCAED · RTILI	TTYEGTHNHP LPPAAM	AMA <mark>STTS</mark> AAASML	L <mark>S</mark> G	S - MP SADG	s		LMSSNFLARTV-		IIR1
VvWRKY-IIB1b	YY <mark>rct</mark> mavgcpv <mark>rko</mark> vo	RCAED - R <mark>S</mark> ILI	TTYEGTHNHPLPPAAM	AMA <mark>STTS</mark> AAA <mark>N</mark> ML	L <mark>S</mark> G	<mark>s</mark> - M <mark>s</mark> sadg			MNP - NFLARTI-		11D1-
PtWRKY-IIB1b	YY <mark>RCT</mark> MAVGCPV <mark>RKQ</mark> VQ	RCAED - KTILI	TTYEGNHNHPLPPAAM	T MA <mark>STTT</mark> AAA <mark>T</mark> ML	L <mark>S</mark> G	S-MSSADG		. N	1MN <mark>P</mark> · N <mark>LL</mark> ARA I ·		
PtWRKY-IIB1¢	YY <mark>RCT</mark> MAVGCPV <mark>RKQ</mark> VQ	RCAED - RTILI	TTYEGNHNHP LPPAAM	IAMA <mark>STT</mark> AAAA <mark>S</mark> ML	L <mark>S</mark> G	<mark>S</mark> - M <mark>S</mark> SADG			MN <mark>P</mark> - NLLARAI -		
GmWRKY-IIB1y	YY <mark>RCT</mark> MAVGCPV <mark>RKQ</mark> VQ	RCADD - RTILV	TTYEGTHNHPLPPAAM	IAMA <mark>STT</mark> AAAA <mark>T</mark> ML	L <mark>S G</mark>	S-MSSADG		• • • • • • • • • • • • • • • • • •	/MN <mark>P</mark> · N L L ARA I ·		
GmWRKY-IIB1h	YYRCTMAVGCP VRKQ VQ	RCADD - RTILV	TTYEGTHNHPLPPAAM	IAMA <mark>STT</mark> AAAA <mark>T</mark> ML	L <mark>S</mark> G	S - MS SADG		••••••	/MN <mark>P</mark> · NLLARAI ·		
GmWRKY-IIB1i	YYRCTMAVGCP VRKQVQ	RCAED - RTILT	TTYEGTHNHPLPPAAM	AMA S T T VAA <mark>T S</mark> M L	L <mark>S</mark> G	S - MS SADG			(MNP - NLLTGAI -		
GmWRKY-IIB1j	YYRCTMAVGCP VRKQVQ	RCAED - RTILT	TTYEGTHNHPLPPAAM	AMASTTAAAASML	· · · · L SG	S-MTSADG			MNP - NLL TRAI -		
MWRKY-IIB1C	YYRCTMALGCP VRKQVQ	RCAED - RSILI	TTYEGTHSHPLPPAAM			S-MSSADG			MNP - NLLARIL -		
COWRANT-INDIC	YTRCIMAVGCP VRKUVU	RCAED RSILI	TTYECHNNHPLPPAAM			S-MSSAEG			MNP · NLLARAI ·		
AWRKY-UR16	YYRCTMAYGCP VRKOVO	RCAED . RTILI	TTYEGNHNHPIPPAAM	NMASTTT AAASMI		STMSNODG			MNPTNII ARTI.		
AtWRKY-IIB1c	YYRCTMAGGCP VRKOVO	RCAED - RSILI	TTYEGNHNHPLPPAAT	AMASTTTAAASMI		S-MSSODG			MNPTNLLARAI-		
GmWRKY-IIB1k*	VLEPTTVAAWE LHAQVQ	RCFKD·ETILI	TTYEGNHNHPLPPAAR	PLASSTSAALNME	L S G	SITSSHCT	т		LSNSPLFSSSPS		
								-			
									· · · · · · · · · · · · · · · · · · ·		
OsWRKY-IIB2a	YY <mark>RCTMASQCP</mark> VRKQVQ	RCAED - KSILI	TTYEGTHNHPLPPAAA	AMAK <mark>TTS</mark> AAAAMI	L • • • • • • • • L <mark>S (</mark>	PAV <mark>S</mark> RDA -	· H · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·	VA <mark>PPP</mark> F <mark>FH</mark> H <mark>P</mark>		
ZmWRKY-IIB2a	YYRCTMATGCP VRKQ VQ	RCAED - KAVLI	TTYEGTHSHQLPPQAA	AMAKTTSAAAAMI	L L <mark>S G</mark>	PAVSRDAG	AH · · · · · · · · · ·	· · · · · · · · · · · · · · · v/	AA <mark>P</mark> LFAQ <mark>Y</mark> H <mark>P</mark>		
BdWRKY-IIB2a	YYRCTMAT GCP VRKQVQ	RCAED - KTVL I	TTYEGSHNHQLPPAAF	T MANTTS AAAAMI	L <mark>S</mark> G	PATSRDGP			· · · TASFFHPHHQ	HYS	
OsWRKY-IIB2b	YYRCTMA IGCP VRKQVQ	RCAED - KTVL	TTYEGNHNHOLPPAAT	TMANTTSAAAAMI	L <mark>S</mark> G	PAASRDGA	A		· · · · HHHHHPAAM	FHQS	
SBWRKY-IIB2a	YYRCTMAVACP VRKOVO	RCAED - KTILV	TTYEGHHNHPLPPAAT	TMANTTSAAAAMI		PATSRDGA	A		···ALFHHSSS·		
	YYPCTMAVACP VPKOVO	RCAED TT II	TTYEONUNUPL PREAT	AMAATTS AAAAAMI		C C C C NI UO	T		A · AAL PHHSSSS		
CuWRKY-IIR?	YYRCT MAYGCP VPKOVO	RCADD - KS II	TTYEGNHNHPLPPAAT	AMASTTSAAAAMI		STS.TK.			PTS · AFLSS · · ·		нира 📗
PtWRKY-IIB2a	YYRCTMAVGCPVRKQVQ	RCAED - KTILI	TTYEGNHNHPLPPAAT	AMASTTSAAAAMI	L SC	STT SK			SSSSSFYPS		IIBZ-
PtWRKY-IIB2b	YYRCSMTVGCPVRKQVQ	RCAED - KTILI	TTYEGNHNHPLPPAAT	VMANTTSAAATMI		STS-SR		· · · · · · · · · · · · · · · · · · ·	LSSSSGFYPS···		
VVWRKY-IIB2a	YYRCTMAAGCP VRKQVQ	RCAED - KTILI	TTYEGNHNHPLPPAAT	AMANSTSAAAAMI	L <mark>S</mark> G	SNT - SK		A	VNSAGFFHP	· · · · · · · · ·	
MWRKY-IIB2a	YYRCTMAVGCP VRKQVQ	RCAED - KTILI	TTYEGNHNHPLPPAAT	A I AH <mark>TTS</mark> AAAAMI	L <mark>S</mark> S	<mark>s</mark> тs - sт		RI	KESAT <mark>GYL</mark> SN-S-		
GmWRKY-IIB2a	YY <mark>rct</mark> mavgcpv <mark>rkq</mark> vq	RCMED - KTVLI	TTYEGNHNHPLPPSAT	VMAN <mark>STS</mark> AAAAMI	L L <mark>S</mark> S	SCSISN		A	L SNT V <mark>G V F</mark> S S - · ·		
GmWRKY-IIB2b	YY <mark>rct</mark> mavgcpv <mark>rkq</mark> vq	RCMDD - K <mark>T</mark> VL I	TTYEGNHNHPLPPSAI	VMAN <mark>STS</mark> AAA <mark>S</mark> MI	F L <mark>S</mark> S	SCSTSNN -		A	L S N T V <mark>G V F</mark> S S		
GmWRKY-IIB2c	YY <mark>rct</mark> mavgcpv <mark>rkq</mark> vq	RCADD - KAVLI	TTYEGNHNHPLPPAAT	AMAN <mark>TTS</mark> AAAAMI	L L <mark>S G</mark>	SAAPSK		· · · · · · · · · A	L T N S A <mark>g y y</mark> S S S S -		
GmWRKY-IIB2d	YY <mark>r</mark> ctmavgcpv <mark>rkq</mark> vq	RSADD - KTVLI	T S Y E G NHNHP L P P A A T	AMANTTSAAAAMI	L L <mark>S G</mark>	SAA - SK	• • • • • • • • • • • •	· · · · · · · · · · · s	L T N S A <mark>G Y Y</mark> S S - T -		

WRKY DNA-binding domain
 30
 37
 30
 30
 30
 30
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40
 40< 360 370 380 390 PpWRKY-IIB3a PpWRKY-II3b PtWRKY-IIB3a VyWRKY-IIB3a AtWRKY-IIB3a PtWRKY-IIB3b GmWRKY-IIB3a GmWRKY-IIB3a GmWRKY-IIB3a GmWRKY-IIB3a MtWRKY-IIB3a PtWRKY-IIB3c GmWRKY-IIB3d GmWRKY-IIB3e MWRKY-IIB3a GmWRKY-IIB3f CpWRKY-IIB3a VVWRKY-IIR36 GmWRKY-IIB3g GmWRKY-IIB3h IIB3-GmWRKY-IIB3i OsWRKY-IIB3a BdWRKY-IIB3a ZmWRKY-IIB3a VvWRKY-IIB3c* AtWRKY-IIB3b* WWWRKY.HR14 OsWRKY-IIB3b* ZmWRKY-IIB3b ShWRKY-IIB3a BdWRKY-IIB3b AtWRKY-IIB3c* GmWRKY-IIB3j* GmWRKY-IIB3k* OsWRKY-IIB3c* BdWRKY-IIB3c* OSWRKY-IIR24 ZmWRKY-IIB3c SHWRKY.IIR16 ZmWRKY-IIB3d GmWRKY-IIB31 AtWRKY-IIB3d FLAR YYRCTMAAGCPVRKGVDRCAED - TTVVVTTYEGNHNH PLPPAAMPMASTTTASSML YYRCTMAAACPVRKGVGRCAED - TTVVVTTYEGNHNH PLPPAAMPMASTTAAAAML YYRCTMAAGCPVRKGVGRCAED - RTVLITYEGNHNH PLPPAAMPMASTTAAAASML YYRCTMAAGCPVRKGVGRCAED - RTVLITYEGNHNH PLPPAAMPMASTTAAAASML LMAGSNFLARAV - - - - -ZmWRKY-IIB4a ZmWRKY-IIB4b LMAGSNFLARAV - -LMAGSNFL IMAGSNFMARAV-----IIB4-SbWRKY-IIB4a BdWRKY-IIB4a MPSADGS LMAGSNFLARAV-

	RxxLPC	LDL-2							QSKF				alanine-rich-2		
	470 480	490	500	510 52	0 5:	30 540	550	560	570	580	590	600	610 6	20	
PtWRKY-IIB1a	LPCS-SSLAT	ISASAPFPTVTL	LTQNP5	PLOLPKOPIOF	FPFPNP P	QNLATASAAALL	PQILGQALYN			· · · · · · · · · · a	SKSFGLOMS -	- ADSLTTAT	AA I AA <mark>DPN</mark> FTAALAAA	1 T S I I G	
GmWRKY-IIB1a	•••• L P ••• C S • S SMAT	I I SASA <mark>P</mark> F <mark>PTVT</mark> L	D L T H S P N P · · · L	. Q F <mark>P</mark> R Q Q H <mark>P</mark> N Q L Q	a I G V P Q N	F A N S <mark>P</mark> S - S L M	PQIFGQALYN			. .	SKFSGLQMS -	<mark> v</mark> g.	AAIAT <mark>DPN</mark> FTSALAA#	1 <mark>T S</mark> I I <mark>G</mark>	
GmWRKY-IIB1b	····L <mark>P</mark> ··· <mark>C</mark> S· <mark>S</mark> SMAT	I I SASA <mark>P</mark> F <mark>P T V T</mark> L	D L T H S P N P · · · L	. Q F <mark>P</mark> R Q Q H <mark>P</mark> N Q L (Q I G V P Q N N	FANS <mark>P</mark> AASLL	PQIFGQALYNN - ·			· · · · · · · · · a	SKFSGLQMS -	- A <mark>D</mark> T <mark>VG</mark>	AA I AT <mark>DPN</mark> FTAALAA#	li <mark>ts</mark> ii <mark>g</mark>	
MtWRKY-IIB1a	· · · · L P · · · YS · SSIAT	I SASA <mark>P</mark> FPTVTLI	0 L T Q S P N Q N	IQF <mark>P</mark> - NNHSNQF	Q F P F P	QN F L	PQVFGQTLLN			· · · · · · · · · · · · · · · •	SKFSGLQMS -	- A <mark>D</mark> T VN	- A I AA <mark>DPN</mark> FTAALAAA	1 <mark>T S</mark> I I G	
MWRKY-IIB16	····LP···CS·SSIAT	I SASAP F PT I T I		TPORNPYOF	TPLITH S	· · · SANSSTSLI	POIP ··· NON···			· · · · · · · · · · · · · · · · · · ·	SKFSGLOMS -	- VQ I VN	AA I AANPN F PADL LAA		
GINWRKT-IIB1C		ISASAPFPIVIL			VPEPNP	ONI SNS <mark>P</mark> AAALL	POLESOAL YN.				SKESGLOMS.		AA I AAUPNFI AALAAA		
VvWRKY-IIB1a	PCSPSLAT	ISASAPFPTITL		LLQHQRPNAQF	1V	PFQNLPQNFAPG	SHAFNPVLHS				SKESALOSSP	ESDTVTAAT	AAITADPNFTAALVAA	ITSIIG	
GmWRKY-IIB1d	LPCSRN-MAT	L SASA <mark>P</mark> FPTITL	LTQNAT	N - SSQLQREI	P	L <mark>G</mark> QLG LL	SPLL - AHKF			M	S - VPKIFGQG	ANILNAAT	A DPN FTAALVAA	I <mark>T S</mark> VM <mark>G</mark>	
GmWRKY-IIB1e	LSCSQNTAAT	I SASA <mark>P</mark> F <mark>PTIT</mark> LI	DLTQSAT · · · · ·	NNSSQLLQ <mark>G</mark>	4 <mark>Р</mark>	· · QDNQH · · SLL	S <mark>PVL</mark> · · AQKF · · ·			M	SATNIFDQ -	- VDT VNAAT.	AA I TA <mark>DP</mark> KF SAALMAA	1 <mark>T S</mark> I I <mark>G</mark>	
GmWRKY-IIB11*	····LAT	I L ST S A <mark>P</mark> F <mark>P T I T L</mark>	0 L T Q S T T	NNSSQLLQ <mark>G</mark> /	А <mark>Р</mark>	· · QDNQHIYSLL	S <mark>P</mark> LL AQKF			· · · · · · · · · M	S A T N I F Y Q N	HVDT VNAAT	AA I T <mark>GDP</mark> KFSAAVMA#	I <mark>TS</mark> II <mark>G</mark> ···	
CpWRKY-IIB1b	· · · · · PA · · ACPPSLAT	I SASA <mark>P</mark> FPTVTLI	DL T ST T	SDPYSFSQK	(P	ILNNNNNSNNNL	SYMLGHALQN			· · · · · · · · · · L	SKFSGLF <mark>G</mark> MD	P S A VHD S T V	G <mark>ST</mark> TVEPSFTAALMAA	1 T AVM G	
OSWRKY-IIB12	····LP···CS·SSMAT	I SASAP F PT VT L	DLTHAPPG · · · A	ARPGAPAP - QFC		GGGMA - PA	FAV <mark>PP</mark> QVLYNQ -				SKFSGLOMSS	DSDTVSAAA.	AAITADPNETVALAAA		
SHWRKT-IIB12		ISASAPEPTVIL				GGGMAGPT							AATTADPNETVALAAA		IIBI-
VvWRKY-IIB1b		ISASAPFPTVTL		PLOYORPTSOF		QSFPSAQTS - SL	POVESOAL YN				SKESGLOLS	ADTVSAAT	AAITADPNETAALAAA		
PtWRKY-IIB1b	L PG CS - SMAT	I SASA <mark>P</mark> FPTVTL	LTQNTN	PLQFQRPPTQF	VPFPGQP	QNFALVTAP - QL	PQVFGQALYN			.	SKFSGLQLS-	- HOTLSAAT	AA I TA <mark>DPN</mark> FTAALAAA	1 5 5 1 1 G	
PtWRKY-IIB1c	L <mark>P</mark> AGS-SNMAT	I SASA <mark>P</mark> F <mark>PTVT</mark> L	DL <mark>T</mark> HN <mark>P</mark> N	PLQFQR <mark>PPPQF</mark>	Q <mark>VPFPG</mark> Q <mark>P</mark>	QNFSSVTT <mark>P</mark> - QL	PQVFGQALYN			.	SKFSGLQLS-	- QDTLSAAT	AA I TA <mark>DPN</mark> FTAALAA#	1 <mark>5 5</mark> 1 1 <mark>6</mark>	
GmWRKY-IIB1y	····L <mark>P</mark> ···CS·TSMAT	I L SASA <mark>P</mark> F <mark>P T V T</mark> L I	DL <mark>T</mark> HN <mark>P</mark> N	PLQFQR <mark>PG</mark> APF	QVPFLQA-QP	QNF <mark>GSG</mark> AT <mark>P</mark>	IAQALYN			· · · · · · · · · a	SKFSGLQLS -	- ADTVSAAA	<mark>S</mark> AITS <mark>DPN</mark> FTAVLAA#	1 <mark>5 5</mark> 1 1 <mark>6</mark>	
GmWRKY-IIB1h	····LP···CS·TSMAT	I SASA <mark>P</mark> F <mark>PTVT</mark> LI	DLTHNPN	PLQFQRPGAPF	VPFLQA-QP	QNF <mark>G</mark> SGAT <mark>P</mark>	AQALYN			· · · · · · · · · a	SKFSGLOLS -	- ADT VSAAA	SAITS DPN FTAVLAAA	1 <mark>5 5</mark> 1 1 <mark>6</mark>	
GmWRKY-IIB1i	L P C S - N - MAT	LSASAPFPTVTL	DLTHNPN · · · · ·	ALQQYQLRPQT	TPFLPS - PP	QNFMSGPTTPQL	PKLIAQVLYN			· · · · · · · · · · · · · · · · · · ·	SKFSGLQLS-	- TDTVS	· AITADPNFPAALTAA	1 <mark>5 5 1 1 6</mark>	
GINWRKY-IIB1		LSASAPFPIVIL		DNNGNE BEOF	PLIPS-QP	ONFOSOO IN TAGSTPPQL	POLLAGALYN				SKISGLULS-		- ATTADPNETAALVS		
CuWRKY-IIR1c	IPCS. SSMAT	ISASAPFPTITI		PIGFORIPTOF		ONFOSVT	POVENGALYN				SKESGIGISN			TSIIN	
AtWRKY-IIB1a	•••••LP••••CS•TSMAT	I SASA <mark>P</mark> FPTVTL	LTHSPP PPN	GSNP SSSA	TNNNHN - QQ	QQQMTNLPPGML	PHVIGQALYN			.	SKFSGLQFS-	- ADT I T	- ALTADPN FTAALAAN	1 <mark>5 5 M I N</mark>	
AtWRKY-IIB1b	L P C S - S SMAT	I SASA <mark>P</mark> F <mark>PTITL</mark>	0 L T E S P N · · · · ·	GNNP · · · TNNP	. MQ F S Q	R S G L V E L N Q S V L	PHMMGQALYYNQ				<mark>s</mark> k f s <mark>g l hmp</mark> -	- GESVSAAT	AA I ASN <mark>PN</mark> FAAALAA#	1 T S I I N	
AtWRKY-IIB1c	····LP···CS·SSMAT	I I SASA <mark>P</mark> F <mark>PT I T L</mark>	DL <mark>T</mark> NS <mark>P</mark> N	GNN <mark>P</mark> NMT TNN <mark>P</mark> I	. MQ F A Q	R <mark>PG FNP</mark> AVL	PQV <mark>VG</mark> QAMYNNQ(a	SKFSGLQL <mark>P</mark> -	- AESVSAAS	AA I AS <mark>DPN</mark> FAAALAA <i>I</i>	ITSIMN	
GmWRKY-IIB1k*	·····	FSHNATCPTVTL	DLTQPNN	YLQFQRATTS	DR <mark>H</mark> T <mark>P</mark>	S F F <mark>P</mark> L <mark>P</mark> L H <mark>G</mark> N	PQNYSEDLMHL -			w	YR <mark>VP</mark> L <mark>P</mark> TML -	- LAL VDVVS	E A I T K <mark>D P</mark> S L K A A L F <mark>S</mark> A	ISSLTE	
	PYAS														
OcWPKY-IIR2a		ISASAPEPTITI			VAESSI			/1 <mark>P</mark>				e vie T HT			
ZmWRKY-IIB2a		LSASAPFPTITL		PAAGLLQHRLP	PPVPAM	PEPMPYGEP	GAGGHRLAAAPV	AP		HP	PAGATILG	RSALDTMT		STIMA	
BdWRKY-IIB2a	FPYAS SMAT -	LSASAPFPTITL	DLTOPPAG	- LPPAASP	APAAMM	P LPPQLA	- MYLQQQ RAS	TTM		LP	PAGLTVOG	- OSVMDTVT	AAIAADPNESTALAA		
OsWRKY-IIB2b		L S A S A P F P T I T L	<mark>р L Т</mark>	. HR <mark>PP</mark> VIH <mark>PG</mark>	MAQAM	PFAVPPQLA	- MYLPQQ RAA			<mark>.</mark>	- AGLGGAG	- PSVMETVT	AALAADPNFTTALAA	A I S S V V A	
ShWRKY-IIB2a		L SASA <mark>P</mark> FPTITL	D L <mark>T</mark> Q A <mark>P B G</mark> L	. HR <mark>PPGGIHPV</mark> T/	а <mark>vр</mark> АМ	· · · <mark>P</mark> F <mark>P V P S P</mark> L A	SMF <mark>LP</mark> Q RA <mark>P</mark>	б <mark>Р</mark>		<mark>P</mark> M	PTGLQVAR · ·	- SVMMETVT	AA I AA <mark>DPN</mark> FT TALAA/	AI <mark>SS</mark> VMA	
ZmWRKY-IIB2b	·······	L SASA <mark>P</mark> F <mark>P T I T L</mark>	0 L T Q Q A P L	. H R <mark>P P G G</mark> M H <mark>P</mark> V V	A <mark>APA</mark> AAM	<mark>P</mark> F <mark>P A P</mark> S <mark>P</mark> L A	• MFLPQ • • • RAP	r		· · · · · · · · v	P T G L P V A P R Q	- QSVMETVT	AA I AA <mark>DPN</mark> FTTALAA/	1 <mark>5 5</mark> V M A	
AtWRKY-IIB2a	·····FPYTSTIAT ·	LSASAPFPTITL	DLTNPPRP	LQ <mark>PPP</mark> QFLS		QYGPAAFLP	NANQ I RSMNNN -			• • • NQQ	LLIPNLF <mark>GP</mark> -	- REMVDSVR	AA I AM <mark>DPN</mark> FTAALAA)	4 <mark>5 N</mark> <mark>6</mark>	
CpWRKY-IIB2a				MHLLR · · · SQT	SFPLP	· · · LHGSPQLLG	HPMFVAPH K -			F	PAMPPLQLA-	- PSMVETVT	AAIASDPNFTAALAA	A I <mark>S S I I 6</mark>	IID2-
PtWRKY-11829		ISASAPEPTITI		MPFPCTSPSPA	FPFP	LHGCPOIPG	NPMYVAP				PATPSVQLG-	- ASMVETVT	AATASUPNETAALAA	AT ST FMG	
VvWRKY-IIB2a	MPYLSTMAS	LSASAPFPTITL		MHFHRGPPSST	5 F P S P	· · · LHACPQLIG	QPLYAPP K-				PVLPSAOMO.	- BSMVETYT	AAITSDPNETAALAA	AISTIIG	
MtWRKY-IIB2a	FPYATMATST	LSASQPFPTITL	FTQNHNL S	MHHNRVPL			- P L F F SHK			L <mark>P</mark>	PLLQLGQP	- S - MVESVS	AAISSDPNFTTALAA	AISSIIG	
GmWRKY-IIB2a	· · · · · · · · · · · · · · · · · · ·	LSTSAPFPTITL	MTTNPMQ S	ALTSLL <mark>P</mark> LH		ATSIPQLLG	н <mark>р V I F Р</mark> НК			.	MPHPLGQQ	- LFLNETMS	AAIASNPNFTIALAA	A I <mark>S S</mark> I I G	
GmWRKY-IIB2b		L ST S A P F P T I T L	DMTTNP s	ALTS <mark>P</mark> L <mark>P</mark> LH	<u>.</u>	· · · ATTF <mark>P</mark> QLLG	Н <mark>Р V I F Р</mark> НК			· · · · · · · · · <u>-</u> ·	M <mark>PHP</mark> LLGQ	- L FTTETMS	AA I ASN <mark>PN</mark> FT I ALAA/	4 <mark>5 5</mark> <mark>6</mark>	
GmWRKY-IIB2c	· · · · · · · · IPYASMAT · ·	L SASA <mark>P</mark> F <mark>P T I T L</mark>	DLTQNPNN A	AMQLHRV <mark>P</mark> AGHG	• AT F P L P L	HAAAA <mark>G - P</mark> HLLG	H <mark>P L F</mark> FQQK - L <mark>P</mark> - 1			• • • • • • • • • • • <mark>P</mark> A	ALM <mark>P</mark> LLQR	- SSMVETVS	AAIAS <mark>DPN</mark> FTAALMA/	4 <mark>5 5</mark> <mark>6</mark>	
GmWRKY-IIB2d	······	LSASAPFPTITL	$\mathbf{DL}\mathbf{T}\mathbf{QNP}\mathbf{NN} \cdots \mathbf{A}$	AMQ L H R V <mark>P P G</mark> H <mark>G</mark> G	SAT FPLP L	HAAAS <mark>GGP</mark> HLL <mark>G</mark>	HPLFFQQKLL <mark>P</mark> -			<mark>P</mark> Q	ALMPLLQR	- S - MVETVS	AA I AS <mark>DPN</mark> FTAALAA	A I <mark>S S</mark> I I G	



Appendix 4

		10	20	30	40	50	60	70	80	90	100	110
	PtWRKY-IIC1a		MEEVIKSS	••••••••••••••••••	DHESEDELVRE	LLD			• • • • • DES <mark>P</mark> FFF	L <mark>P</mark> KE····		
	VvWRKY-IIC1a		MDELMA · · · ·	AWT	E <mark>G</mark> SED - ELVRE	LLD			DES <mark>P</mark> FLV	V <mark>P</mark> QE · · · · ·		
ANNONY COLD ILLEP ANNONY COLD ILLEP<	GmWRKY-IIC1a	M	MEEVVMTT	sws	E <mark>g</mark> s e d d d l v r e	LLD			D <mark>G</mark> S <mark>P</mark> L L I I	E <mark>PP</mark> N		
	GmWRKY-IIC1b								<mark>P</mark> L L I	E <mark>PP</mark> N · · · · ·		
OWNEY, COLUMPTIC LET LET LET TAAN. IN IN NUMBER LET DER. IN IN AAAADEG. OWNEY, COLUMPTIC LET ATAA. IN IN AAAADEG. IN IN AAAADEG. OWNEY, COLUMPTIC LET ATAA. IN IN AAAADEG. IN IN AAAADEG. OWNEY, COLUMPTIC LET ATAA. IN IN AAAADEG. IN I	AtWRKY-IIC1a		ИЕЕЕ <mark>б</mark> Ү	QW4	RRC <mark>G</mark> NNAVED <mark>F</mark>	FVY			E <mark>PP</mark> LFF	L <mark>P</mark> QD · · · · ·		
MANNEYS IF BOLLMARLLUE AATAA IF STANDARD IF STANDARD IF STANDARD MANNEYS ILLA AVANDAL LAS INCY ILLA AVANDAL LAS INCY ILLA AVANDAL LAS INCY ILLA AVANDAL LAS INCY MANNEYS IND BOVELLE FRANDAL VINT GEL LED ILLET FRANDY ALDE FRANDAL VINT GEL LED ILLET FRANDY ILLET FRANDY MANNEYS IND STANDAL VINT GEL LED ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDAL VINT GEL LED ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDAL VINT GEL THE ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDAL VINT GEL THE ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDY ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDY ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDY ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDY ILLET FRANDY ILLET FRANDY ILLET FRANDY MANNEYS IND STANDY ILLET FRANDY ILLET FRA	OsWRKY-IIC1a	G D V V G E E L L R E I I	LDETAAVH		NSNSNSNSNSNS	EAEEEDER··			· · · · EYFAAAA	ADEQ		
BURKEY VICE LILLEAV VIET LAE INKY CONDUCTOR TARGET AND	BdWRKY-IIC1a	FGFGQELMMRELI	LDEAATAA		VDAGAAAGYSS	YSSKDDEE · ·			· · · · EEEEYCR	SAAR		
CHMMERY DEF INTERDUCTION TOTAL CODUCT CALL DO LUT CALL DO	ShWRKY-IIC1a	LLLLEASVAVPP	LAESINKY ···		NGSDFAT - TEE				· · · · DEAALSS	s <mark>p</mark> vo		
WHWYYGE HUG BY VLINE PRAADVA. YWGE L DOL LIT LLED DILLT TEPEOVADD F CONSTRUCTS AAAAACCA ADWYYGE HUG DY VLINE PRAADVA. YWGE L DOL LIT LLED DILLT TEPEOVADD F CONSTRUCTS AAAAACCA ADWYYGE SSTOD D	CHWRKY-HC4a		HEEDONI R		DOGE EL VNE				NESPENV	VEKD		
AMMERY HER NO. VLARGAT AS ALE VALUE VALUE VALUE LINDEL LV. AAAL H. T.LYYST KASSAAAAAC CH.	Children Hors	HODODYOLUBER		V.M.	GELDEOLITEL	LEDDELLIAT			PDTCCASSADA	PCIE		
ANNEYYIC23 SSENORD DNG TINLD HINE Y GISSON ANNEYYIC23 SSENORD DNG TINLD HINE Y GISSON SAMMEYYIC23 SSENORD AAAAAAAAYAYYYAAFSONTYYAA TAAAAAAAAAYAYYYAAFSONTYYAA SAMMEYYIC23 SSENORD MAAASSSTYTYYAAFSONTYYAA TAAAAAAAAAAYAYYYAAFSONTYYAA SAMMEYYIC23 SSENORD MAAASSSTYTYYAAFSONTYYAA TAAAAAAAAAAAYAYYYAAFSONTYYAA SAMMEYYIC23 SSENORD MAAASSSTYTYYAAFSONTYYAA TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	O MERCY HOAK	MODODIGCHPER			OF LOPULIE		LIFFIQUURA	ACUP TLYYS				
ANNEYY ACCOUNT OF A CONTRACT ON A CONTRACT O	USWKK 1-11010		SALEVA <mark>uu</mark>		CELDUNLIGEL			ACHE	F UAU S SAAAAA	A <mark>P</mark> CN		
AMMENY-MEZY STATE OLD DIG V LID - HARLY SISTER												
ANNEYY MC22 - SSTSGRO - DHOR I LULD - HURLING SISTON ONNEYY MC22 - SSTSGRO - DHOR I LULD - HURLING SISTON SSTSGRO - DHOR I - LULD - DHOR I - LULD - HURLING SISTON SSTSGRO - DHOR I - LULD - DHOR I - LULD - HURLING SISTON SSTSGRO - DHOR I - LULD - DHOR I - LULD - HURLING SISTON SSTSGRO - DHOR I - LULD - DHOR I - LULD - DHOR I - LULD - HURLING SISTON SSTSGRO - DHOR I - LULD - DHOR I - LULD - DHOR I - LULD - HURLING SISTON SSTSGRO - DHOR I - LULD - DHOR I - L												
AMMENY ADD IST GOD DHOE ILLE. HURL THE ISONE ILMENY ADD SHMENY ADD STT GOD ADD ADD ATT TY APER TO THE ISONE ILALAH SHMENY ADD STT GOD STT GOD STT GOD SHMENY ADD ILALAH ILALAH SHMENY ADD FINDION FILL STT GOD STT GOD SHMENY ADD FINDION FILL FINDION FILL STT GOD SHMENY ADD FINDION FILL FINDION FILL ST GOD STANDARD SHMENY ADD FINDION FILL FINDION FILL ST ST ADD SHMENY ADD FINDION FILL FINDION FILL ST ST ADD SHMENY ADD FINDION FILL FINDION FILL ST ST ADD SHMENY ADD FINDION FILL FINDION FILL ST ST ADD SHMENY ADD FINDION FILL FINDION FILL ST ST ADD SHMENY ADD FINDION FILL FINDION FILL ST ST ADD SHMENY ADD FINDION FILL												
NUMBER VIEWS AND AND AND AND TY VER FOR THE PARTY OF A CONSTRUCTION OF A CONSTRUCTIO	AN169407 1100-		DUD C 1 NU		con <mark>n</mark>				1 chi <mark>n</mark> ter			
NUMBERY 1023 STATA	AWKK T-IIC29		· · · DHD <mark>g</mark> flnl	ACCCOUNT DT	SUR <mark>P</mark>				······································	AW		
SAMENY RULES S ST A A S DALARS ST T T T A C OB SERVER AT S E L A LANDENH SAMENY RULES NO ST A T T T T T T T T T T T T T T T T T	05WRK 1-11029		AAA <mark>F</mark> ATATTTV	Arsrunririr.								
ZUMENY MICLS S TAM GLARA S S I I TA A GOS STARTS S LALA S S I ZUMENY MICLS I TO MUNITY HILS TO MUNITY ALLS COTTING IN CONTRINCT (KI I S A A I TS S A I TO MUNITY ALLS COTTING IN CONTRINCT (KI I S A A I TS S A A I TS S A A I TS S A I TS S A I TS S A A I TS S A A I TS S A I TS S A A I TS S A I TS S A A I TS S A A I TS S A I	SBWRKY-IIC22	SSTAGAAGGM/	AAASSSST <mark>P</mark> T I	TFAFQ - PSPPP -	· · · T S · · · · ·				· · · GLALAH · ·	.н		
AMMENY ALESS IN THE TY PILLY PLO CHOSENESS AS A	ZmWRKY-IIC2a*	· · · S S T A <mark>G</mark> · · · · <mark>G</mark> M/	4 A S S S S S - T P T I	T F A F Q Q <mark>P</mark> S <mark>P P P P</mark>	P TATS				· · · SLALAH <mark>P</mark> HI	чн		
MMWRY-MC22 MOMBURY-LIP LONCHIGS LKG SALTES S.S.S.C.A.M. MMWRY-MC22 TOBLEL-NIEP MOLYTEL COTTINUES SALA S.S.S.C.A.M. GWWRY-MC23 S.S.F.F.R.D.D.F.G.T.H.S.S.LALE SLAKE S.S.S.C.A.M. GWWRY-MC23 S.S.F.F.R.D.D.F.G.S.S.C.G.G.B.K.F.N.F.M.LT.F.S.S.C.M.S.L.S.F.C.A.M.S.S.C.G.G.S.K.F. C.G.G.C.M.S.S.S.S.C.G.G.S.K.F. GWWRY-MC23 S.S.F.F.R.D.D.F.G.S.S.C.G.G.G.K.F. I.F.S. LAADASS GWWRY-MC23 S.S.F.F.N.LTW G.S.C.G.G.G.K.G.F.L.F.S. LAADASS GWWRY-MC23 S.S.F.F.N.LTW G.S.C.G.G.G.K.G.F.L.F.S. LAADASS GWWRY-MC23 S.S.F.F.N.LTW G.S.C.G.G.G.K.G.F.L.F.S. LAADASS GWWRY-MC23 MEKY O.T.L.W.G.G.S.C.G.G.K.G.F.L.F.S. LAADASS GWWRY-MC23 MEKY O.T.L.W.G.G.S.C.G.G.K.G.F.L.F.S. S.S.S.G.G.G.G.G.G.G.G.G.G.G.G.G.G.G.G.G	MtWRKY-IIC2a*	••••••••••••••••••••••••••••••••••••••	P F P T N L T F <mark>P</mark> S L	. D CNQS LKAFS	SIASS				LTSESDST	5N		
GWWRY/WIGZ TOMELYN LFYRDUY WELS COTYTES SALA S.S. SATS GWWRY/WIGZ TOMELYN LFYRDUY WELS COTYTES SALA S.S. SATS GWWRY/WIGZ TOMELYN LFYRDUY WELS COTYTES SALA S.S. SALA GWWRY/WIGZ TOMELYN LFYRDUY WELS COTYTES SALA S.S. SALA GWWRY/WIGZ TOMELYN LFYRDUY WELS COTYTES SALA S.S. LAISSODSSN GWWRY/WIGZ TOMELYN LFYRDUY WELS COTYTES SALA S.S. LAISSODSSN GWWRY/WIGZ TOMELYN LFYRDUY WELS COTYTES SALA S.S. LAISSODSSN GWWRY/WIGZ TOTE YN WSAR (COTYTES STERUTY END STERUWEL OLEC LYDDMAA GWWRY/WIGZ TOTE YN WSAR (COTYTES TOTE THE STELAN (FRUNCHMUNEL FRUNCHMUNEL FR	MtWRKY-IIC2b*	· · PMQMGFN · · I	P F <mark>P P</mark> NMT L <mark>P P</mark> L	DNCHHQSLK <mark>G</mark> IS	A T <mark>P</mark> S · · · · · ·	• • • • • • • • • • •			S L S <mark>P</mark> E A	AN	• • • • • • • • • • • •	
OWNEXT-MIC2 / TOULT NIL FITNUT FILL COTT VILLS SALAF	GmWRKY-IIC2a	* <mark>P</mark> TQM <mark>G</mark> LFN - I <mark>F</mark>	P F <mark>P P</mark> NQT Y <mark>P P</mark> L	. <mark>g</mark> - CQT <mark>g</mark> tlksis	A I A <mark>P</mark>				SLS-SA	r s		
GWWRYYG222 * T T G G F F . E F T TH T F T G C GG S L KATS I A SS . L U S GG S X M SS . L U S S S U S GG S	GmWRKY-IIC2b	* <mark>P</mark> TQM <mark>G</mark> LFNII <mark>I</mark>	P F <mark>P P</mark> NQT Y <mark>P P</mark> L	. <mark>g</mark> - CQTVTLKSIS	A A <mark>P</mark>					AN		
GWWRY/WC224* TONGE F - ST F TH F F LG CHOSS KATS IAF SIA SS LA IS GODSASN. GWWRY/WC22* ST GF IS YE VS SA GF RO F TE F THAUG HHAM L DF LOLL CH LA LON F OGD D GWWRY/WC22* ST GF IS YE VS SA GF RO F TE F THAUG HHAM L DF LOLL CH LA LON A SS WWRY/WC22* T G HU F Y S SA GF RO F TE F THAUG HHAM L DF LOLL CH LA ADA SS WWRY/WC22* T G HU F Y S SF BH TY TH G A GODSKAF HI F S LA ADA SS WWRY/WC22* T G HU F Y SS F BH TY TH G A GODSKAF HI F S LA ADA SS WWRY/WC22* T SN YO LO AY TS ST T T I G HUM WWRY/WC22* T SN YO LO AY TS ST T T I G HUM WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23* HE KY O I F I Y SD WWRY/WC23*	GmWRKY-IIC2¢	• • • <mark>P</mark> TQ <mark>G</mark> F F • • • <mark>I</mark>	P F <mark>P</mark> T N L T F <mark>P P</mark> L	. <mark>g</mark> - Chqsslkafs	S A <mark>P</mark> S S S				LVISQQDSV	5N		
Owners/Harzel Winsty-Harzel	GmWRKY-IIC2d	• • • <mark>P</mark> TQM <mark>G</mark> FF• • • 5	S F <mark>P</mark> T N L T F <mark>P P</mark> L	. G - CHQSSLKAFS	S A <mark>P</mark> - S S				LAISQQDSA	5N		
OWNRY/WIG22 • STOR IS YOU SAR & REOP & GET IT MOMEDING WE FEEL LE CHE. • KL LOME FOODED. WWRKY/WIG22 • FOODE & FY STENTLY NO SAR & CHOSLK & F NI FFS • LAADAR SS PWWRKY/WIG22 • TOWEFF & ST FFILLY NO SAR & CHOSLK & F NI FFS • LAADAR ST AWWRKY/WIG22 • TOWEFF & ST FFILLY NO SAR & CHOSLK & F NI FFS • LAADAR ST AWWRKY/WIG23 • FS NUT VIOLOUVISSST • TIGENNA PWWRKY/WIG24 • FS NUT VIOLOUVISSST • TIGENNA PWWRKY/WIG24 • FS NUT VIOLOUVISSST • TIGENNA PWWRKY/WIG24 • HE W • OI FL VS DE PWWRKY/WIG24 • HE W • OI FL VS DE PWWRKY/WIG24 • HE W • OI FL VS DE PWWRKY/WIG24 • HE W • OI FL VS DE PWWRKY/WIG25 • HE W • OI FL VS DE PWWRKY/WIG26 • HE W • OI FL VS DE PWWRKY/WIG26 • HE W • OI FL F A PWWRKY/WIG26 • HE W • HI F FA PWWRKY/WIG26 • HE NY • HI F FA PWWRKY/WIG26 • HE NY • HI F FA PWWRKY/WIG27 • HE NY • OI F FA PWWRKY/WIG28 • HE NY • OI FA FA<	OSWRKY-IIC2b*	SQLETACL <mark>P</mark>	AALYA <mark>p</mark> lc <mark>p</mark> yt	PPSPPSFLAPL	SLQHKL <mark>P</mark> ····				QL <mark>P</mark> QLVHDH	AA A		
VWWRY/WD22* F G GU FY S F HULLY YOUS COUSLIKAF -NU PTS	OsWRKY-IIC2c*	STQFQISY <mark>G</mark> VS	S S A F <mark>G</mark> F R C D F G	TGFITHGMGDHH	GMPLPPPLQLP	СНР			KLLQMPFDQ	ED		
PWWRKY/0223 F TOWEFF SF SPLITY 028AS CHOSKEF 1 PTS. LAADAF STT. AWWRKY/0234 F SN YD LQOYTSSST T I DENNN FLYPE FLYPE ETN. FLYPE ETN. PWWRKY/0234 HEKY O I FL VSDF. FLYPE ETN. PWWRKY/0235 HEKY O I FL VSDF. STSS QWWRKY/0236 HEKY O I FL VSDF. STSS QWWRKY/0236 HEKY O I FL VSDF. STSS QWWRKY/0236 HEKY O I FL VSDF. STSS QWWRKY/0237 HEKY O I FL VSDF. STSS QWWRKY/0236 HEKY O I FL F G SS. STSS QWWRKY/0236 HENY HIN F AA STASSAA SWWRKY/0237 HENY HIN F AA STASSAA WWRKY/0236 HENY HIN F AA STASSAA WWRKY/0236 HENY HIN F AA STSS WWRKY/0237 HENY SNL F STS SS WWRKY/0236 HENY SNL F STSS WWRKY/0237 HENY SNL F STSS SS SS WWRKY/0238 HENY SNL F STSS SS SS SS SS WWRKY/0246 HENY SNL F SS	VVWRKY-IIC2a*	P SQLGFY	SFHPNLSFPOV	G COOSLKAF -	NIPPS				· · · · LAADAPS			
PWWRKY/JIC23 • ETOMO FF.SPESHLTY DOLESS CHOSS KG F. I I EES LAADA ESTT. PWWRKY/JIC23 • FSNYDLOOVTSSST. • TIOENMIN. PWWRKY/JIC23 • MEKY OIFLPYSD PWWRKY/JIC23 • MEKY STUVYSG PWWRKY/JIC23 • MEKY STUVYSG PWWRKY/JIC23 • MEKY STUT PWWRKY/JIC23 • MEKY STUT PWWRKY/JIC23 • MEKY STUT PWWRKY/JIC24 • MEKF OILTPYFA	PtWRKY-IIC2a*	PTQMGFF	FPPHLTYPOS	AS - CHOSLKGF -	11 PP S				· · · · LAADAPS	г т		
ANWRKY/JC23 . NEKY. OIFLPVSD PHWRKY/JC23 . NEKY. OIFLPVSD CWWRKY/JC23 . MEKY. OIFLPVSD STSS . STSS WWRKY/JC23 . MEKY. . STSS WWRKY/JC23 . MEKY. . STSS WWRKY/JC23 . MENY. . HILFGAA. ZUWRKY/JC23 . MENY. . STGS JUWRKY/JC24 . MENY. . STGS ZUWRKY/JC25 . MENY. . STGS ZUWRKY/JC26 . MENY. . STGSS ZUWRKY/JC26 . MENY. . STGSSSS ZUWRKY/JC26 . MENY. . STGSSSSSSS ZUWRKY	PtWRKY-IIC?h*	PTOMOFF	SEPSHITYPOL	GS CHOSSKOF	11 PP S				I AADAPS'	т		
PHWRKY/IIC33 IMERY 0 F V S D VWRKY/IIC33 IME H 0 I P G S S GWRKY/IIC33 IME H N I F A SWRKY/IIC33 IME Y ST V F B SWRKY/IIC33 IME Y ST V F B SWRKY/IIC33 IME Y ST V F B SWRKY/IIC33 IME Y H F A SWRKY/IIC33 IME Y H F A SWRKY/IIC33 IME Y H F A SWRKY/IIC33 IME F A ST G F SA SWRKY/IIC33 IME F A A SWRKY/IIC34 IME F A A SWRKY/IIC35 IME F A B SWRKY/IIC35 IME F A B SWRKY/IIC35 IME F A B SWRKY/IIC36 IME Y D I/ F A SWRKY/IIC36 IME Y D I/ F A SWRKY/IIC36 IME Y	AtWRKY-IIC26*	- ESNYDLOOVTS							ELVPEEF			
PWWRKY-#C23 ME KY 0 I F L V S D K VWWRKY-#C23 ME KY 0 I L F Q S S QWWRKY-#C23 MB K ST M S S PWWRKY-#C23 ME NY MMI F Q A ZWWRKY-#C23 ME NY MI F F A OWWRKY-#C23 ME NF O L F F A PWWRKY-#C23 ME NF O MI F F S QWWRKY-#C23 ME NF O MI F F S QWWRKY-#C24 ME NF O MI F F S QWWRKY-#C23 ME NF O MI F F S QWWRKY-#C24 ME NF O MI F S S <		/ SHIDLAUVIS		· · · · · · · · · · ·								
AtWRKY-IIC43 MDRE D - IN MLSRLD AtWRKY-IIC44 MED DHTN MLTLEE BWRKY-IIC44 MEN FLNG BWRKY-IIC43 MENS FLNG BWRKY-IIC44 MENS FLNG BWRKY-IIC43 MENS TLNG CV ROPATP PL ATT CLAPL ZmWRKY-IIC43 MENS TLNG CV ROPATP PL VATT CLAPL ZmWRKY-IIC43 MENS OF EDEL P VOQF QULA EMANGE GMWRKY-IIC44 MES UD GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P GMWRKY-IIC44 MES OP PNPPP QNN P I FT P SWRKY-IIC44 MES OP PNPPP QNN P I FT P SWRKY-IIC44 <	GpWRKY-IIG3a PtWRKY-IIG3a SbWRKY-IIG3a BdWRKY-IIG3a BdWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a AtWRKY-IIG3a GmWRKY-IIG3a GmWRKY-IIG3a	MDXY MENY MENY MENY MENF MENF MENF MENF MENF MENY MENY MENY MENY MENY MENY	STV HIM STV G G G G G G G G G G G G G G G G G G G	V F S G LL F G A A LL F G R F LF LF F F A L F P F A L F P C F L F P C F L F P C S L F P S L F S V S LL F P S S S S S S S S S S S S S S					ST SS ST HASSA. ST OF SSA. AT OF SS ST AT OF SS ST SS PN ST SS PN SSA PPPA NSSSYP I NSSSYP I	a		
OSWRRY:///C44 - MEN LQLQGDD - HDDEALPH FPYFAVPSP SWRRY:///C44 MEN	AtWRKY-IIC4a AtWRKY-IIC4a BdWRKY-IIC4a SbWRKY-IIC4a BdWRKY-IIC4a BdWRKY-IIC4a AtWRKY-IIC4a GmWRKY-IIC4a WWRKY-IIC4a PtWRKY-IIC4a	MDRE MERS	LHQED - D LHQED - EAH LHQ VIRQPNT LHQGSCQQPT RGRHMSMQPED - ODPPN - ODPPN - ODPPN - HLQI	INPMLSRLD ITNPMLTLEE. PPLATGTCLAPU GLDTASCLAPU DELPPVQGFQYLG LVD IVD PPPQNNPFLFT PPPQNNPFLFT PPPQNNPFLFT PPPQLPPLQ	T							
SWWRY/UC40 - MEN - OHLOGDESSSHALSFY FAVES	OsWRKY-IIC4a	MEN LO	LQ <mark>G</mark> DD - HDDE	ALPHEPYEAVPS	P							
	ShWRKY-IIC46	MEN 01	ILOGDESSSSH	ALPSEPYEAVES	P							
	ZmWRKY-IIC46	MENQ		IT FPSFPYFTVPS	P							

IIC1-

IIC2-

IIC3-

IIC4-

	10		20	30	40	. 5	0	60		70	80	90	100	. 1	10	
AtWRKY-IIC5a			MNY <mark>P</mark> SN <mark>P</mark> N	PSST	 			DF1	TEFFKF	DDFDDTFEM	IMEEI <mark>G</mark> -					
OsWRKY-IIC5a		- MYMAA	AAAGASTP	F	 	N F	CRHGSI	HAEYDAVF	s <mark>g</mark> sw		MARR	APHGGGASGS	sgsgy			
BdWRKY-IIC5a		- MAVAA	AEAAGTA -		 	· · · · A)	RYHHPI	HAAADAVA	CR5		MASSPYS	s As				
ShWRKY-IIC5a		MAL	s s .		 		R S S	FAAADVLL	P AAM		AYROPCS	GGGGGGP AT S S \	F G S R			
ZmWRKY-IIC5a		- MHMAL	s s		 		R S S	FAA - DVLL	РАТМ		SYRQPCS	G ASS)	L <mark>G</mark> SQ			
BdWRKY-IIC5b		- MAAVG	AA <mark>P</mark> VLFY-		 		QQPAP	AAA <mark>P</mark> ALAA	GDAIG-	C F F S <mark>P</mark> S S	PMSSFFS	SH <mark>GHGG</mark> SSSTA	GSSPA			
OsWRKY-IIC5b		- MAAVG	AHAAVYH-		 		HP	V S G L S A P A	GDAAY -	.	SMSSYFS	HGG SS-TS	SSAS			
ZmWRKY-IIC5b		- MAAVG	AH <mark>p</mark> vlyh-		 		H <mark>P</mark>	AP A	GDAS		SMSSYFS	H <mark>GG</mark> SSTTS	SSAS			
ShWRKY-IIC5b		- MAAVG	AR <mark>p</mark> vlyh-		 		H <mark>P</mark>	APA	GDAA		SMSSYFS	Q <mark>GG</mark> SSTTS	SSASA			
BdWRKY-IIC5c		- MAAS -	L <mark>G</mark>		 		- LI <mark>P</mark> EA	ADLFSSAY			- AHGDFT	S - P LQEY				
OsWRKY-IIC5c		- MAAS -	v <mark>g</mark>		 		- LN <mark>P</mark> E	AFFFSNSY	s y s		- SS <mark>P</mark> FMA	. S Y T <mark>P</mark> E · · · · · ·				
SHWRKY-IIC5c		IQMAAS -	L <mark>G</mark>		 		- LN <mark>P</mark> E	ALFASYSS	AYSS		- SS <mark>P</mark> FVS	DYAAS				
ZmWRKY-IIC5c		- MAT S -	L <mark>G</mark>		 		- LN <mark>P</mark> EI	DLFTSYSS	SYYS		- S <mark>PP</mark> FMS	DYAAS				
BdWRKY-IIC5d		- MAASS	LGR <mark>P</mark> AGL -		 	e	LEAYS	S S F S A <mark>P</mark> L Y :	5 F <mark>P P</mark> L -		MEAGNFC	<mark>P P</mark> D D A V L <mark>P G</mark> F A	PELD			
OsWRKY-IIC5d		- MAAS -	L <mark>G</mark>		 		- LCHE	T S Y A Y :	SY <mark>P</mark> AS -		NTSSSLC					
ZmWRKY-IIC5d		- MAAS -	L <mark>G</mark>		 		- LAHE	AA <mark>P</mark> FYAAA`	Y <mark>PP</mark> AG -		AAYAYLA	<mark>P P P G</mark> D				
ShWRKY-IIC5d		- MAAS -	L <mark>G</mark>		 		- LAHD	AS - CYAAY	PPAAA-		AASSYF	S <mark>PPPPG</mark> D····				
GmWRKY-IIC5a		- MT DKN	PR <mark>PP</mark> D -		 		S <mark>P</mark> D	- DDFTNQW	P - L E L S	EYLNFDDDG	W <mark>P</mark> DDY <mark>P</mark> E	S F V				
CpWRKY-IIC5a		- MSSST	DRQE 5 -		 		<mark>P</mark> EAI	N F E A <mark>P</mark> NLF:	SFFYED	EWVRVED	DDLS	S L V F • • • • • •				
VvWRKY-IIC5a		- MADTT	A <mark>G</mark> SQDS <mark>P</mark> -		 		EAE	S D F E L <mark>G</mark> D T 1	SNFELS	EYLLFDD - L	MEEDHSA	FLA				· ·
PtWRKY-IIC5a		- MSS <mark>P</mark> K	FSTQQDT -		 		<mark>P</mark> EL:	SDFAHDQL	SNFELS	EFLTFDEWI	ALEDAPS	S A				• •
GmWRKY-IIC5b		- MT DK I	P <mark>kppp</mark> d-		 		T <mark>P</mark> D :	SDDFTNQW	P - FELS	EYLKFDDNG	WMHD <mark>G</mark> LE	S F A · · · · · · · ·				
GmWRKY-IIC5c			<u>.</u>		 						- MHDDLE	SVA				
AtWRKY-IIC5b		- MNDAD	T N L <mark>G</mark>		 			SSFSDDTH	S V F E F <mark>P</mark>	ELDLSDE	WMDDDLV	SAV				· ·
SbWRKY-IIC5e		- MSSSY	s s		 	•••LL\$	B RAH	ADDHRVLL	GGDA		- DDDDMA	AVS				· ·
OsWRKY-IIC5e		- MSSLY	P S		 	L L S	S L S E S <mark>P</mark> /	AEYRQV <mark>GG</mark>	GRYA <mark>G</mark> -	••••EDV\	DDDDDMA	AVADAVS				· ·
GmWRKY-IIC5d		- MDYYF	GNLN <mark>P</mark> NP -		 	YYH -	HS	AVVNMAS <mark>P</mark> :	S S E F M -		- LSDYLV	LEDALVVD				· ·
GmWRKY-IIC5e	MTDTYLAIT	SMDYYF	GN <mark>PNP</mark> K <mark>P</mark> -		 	Y D N F	εHS	AVVNTES <mark>P</mark> :	5 5 E F M -		- LSDYLV	LEDA - • VD - • •				
GmWRKY-IIC5f		- MDFYF	GNS <mark>PP</mark> Y <mark>P</mark> -		 	NN	- YAHN:	SLN - MALS	S <mark>P</mark> E I A -		- LSDYLM	LDDY VD				
GmWRKY-IIC5g		- MDFYF	GN <mark>PHP</mark> Y <mark>P</mark> -		 	N I H G	YHAHN	SVVSMT <mark>P</mark> S:	S <mark>P</mark> E I A -		- LSDYLM	LDDY VD				
CpWRKY-IIC5b		- MEVIE	GAS <mark>PP</mark> TS -		 		YN	<u> </u>				· · · <u>· ·</u> · · <u>·</u> · · ·				
AtWRKY-IIC5c		- MNISQ	N <mark>P</mark> S <mark>P</mark> NFT -		 	Y F S C	DENFIN	P F M D N N D F 3	SNLMF -		FD	I DE <mark>GG</mark> NN <mark>G</mark> L				· ·
VvWRKY-IIC5b		- MHIYI	SSDSTTR-		 	• • F F • •		<u></u>			F L F	vwikv <mark>g</mark>				
VvWRKY-IIC5c		- MDYYY	\$\$\$\$		 	S <mark>P F</mark>	NPHPN	P S P D P N C A S	SQFEF -		- LYHYLM	L E D <mark>G</mark> S • • • • • •				• •
VvWRKY-IIC5d					 											

IIC5-

						TPNSS		
	10	20 30	40 50	60	70 80	90	100 110	
MIWRKY-IIC6a						AT <mark>P</mark> VSVS	····	
GmWRKY-IIC6a					<mark>.</mark> <mark>P</mark> ,	AS <mark>P</mark> NY		
GmWRKY-IIC6b					• • • • • • • • • • • • • • • • • • •	AS <mark>P</mark> NS		
GmWRKY-IIC6c				• • • • • • • • • • • • • • • • • • • •	· · · · · · · · · · · · · · · A,	ASPNL		
GMWRKY-IIC68						ASPNF		
VyWRKY-IIC6a						ATPNS		
CpWRKY-IIC6a					<mark> P</mark>	PT <mark>P</mark> NS		
PtWRKY-IIC6a					• • • • • • • • • • • • • • • • • • •	AT <mark>P</mark> NSS		
PtWRKY-IIC6b		• • • • • • • • • • • • • • • • • • • •			••••••••••••••••••••••••••••••••••••••	AT <mark>P</mark> NSS		
ATWRKY-IIC60					NPP	TTPNS		
GmWRKY-IICGe						ATPNS		
GmWRKY-IIC61					<mark>.</mark> <mark>P</mark>	T T <mark>P</mark> N S		
VVWRKY-IIC6b					N <mark>P P</mark>	AT <mark>P</mark> NS		
GmWRKY-IIC6g					· · · · · · · · · · · · · · · · · · ·	AT PNS		
GMWRKY-IIC68						AT PNS		
PtWRKY-IIC6d					NOP	ATPNS		
MtWRKY-IIC6b						AT <mark>P</mark> NS		
CpWRKY-IIC6b					• • • • • • • • • • • • • • • • • • •	AT <mark>P</mark> NS	· · · · · · · · · · · · · · · · · · · · · · ·	
BdWRKY-IIC6a					· · · · · · · · · · · · · · · L	LT PNSSLSFS		
ZmWRKY-IIC6a						HTPNSSVSIS		
ShWRKY-IIC6a					MI	HT PNSSVSLS		
ZmWRKY-IIC6b					<mark>P</mark>	LT <mark>P</mark> NTTSSMS		
ZmWRKY-IIC6c					• • • • • • • • • • • • • • • • • • •	VT <mark>P</mark> NTTSSMS		
ShWRKY-IIC6b					· · · · · · · · · · · · · · · · · · ·	VTPNTTSSMS		I IICO
BAWRKY-IIC68						GTPNTTSSMS		
OSWRKY-IIC6c						MTPNSMS · VS · · ·		
SBWRKY-IIC6c					<mark> P</mark>	L T <mark>P</mark> NSMS - MS - · ·		
ZmWRKY-IIC6d		• • • • • • • • • • • • • • • • • • • •		• • • • • • • • • • • • • • • • • • • •	••••••••••••••••••••••••••••••••••••••	LT <mark>P</mark> NSMS - MS - · ·		
ATWRK Y-IIC68	GP SAYNANMI SSSOVGE	DI PSKNI SPOGAFFI GFI	I SPSSSDEENPSI DOI		HEVYG. DGCATIK	SEVENSAS <mark>P</mark> UVS		
AtWRKY-IIC61	LPSQNGFNPSTYSFTDC	· LQSS <mark>P</mark> AAYESLLQKTF	LSPSSEVFNSSIDQ	P · · · · · · · · · NRDVTI	NDVIN GG - ACNE	TETRVS <mark>P</mark> SN		
GmWRKY-IIC6i		· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·	. <mark>.</mark>	· · · · · · · · · · · · · · · •	st <mark>p</mark> nssvs		
GmWRKY-IIC6j		• • • • • • • • • • • • • • • • • • • •			· · · · · · · · · · · · · · · · · · ·	ST <mark>P</mark> NSSVS		
CPWRKY-IIC6¢					· · · · · · · · · · · · · · · · · · ·	STPNSSVS		
PtWRKY-IICGe						STPNSSVSN		
PtWRKY-IIC61						ST <mark>P</mark> NSSISN		
VvWRKY-IIC6d					· · · · · · · · · · · · · · A	GT <mark>P</mark> NSSISS	• • • • • • • • • • • • • • • • • • • •	
PtWRKY-IIC6y		• • • • • • • • • • • • • • • • • • • •			• • • • • • • • • • • • • • • • • • •	AT PNSSLSF	•••••	
GmWRK Y-11668					· · · · · · · · · · · · · · · · · · ·	A I <mark>F</mark> NSSVSF		
GmWRKY-IIC6I						ATLNSSISS		
ShWRKY-IIC6d		•••••••ARS <mark>G</mark> DA <mark>G</mark> AS				· · ssss <mark>g</mark> D · · · · ·		
ZmWRKY-IIC6e		· · · · · · · · · · ARS <mark>G</mark> DA <mark>G</mark> AS			• • • • • • • • • • • • • • • • • • • •	· · SSSS <mark>G</mark> D · · · · ·	• • • • • • • • • • • • • • • • • • • •	
AtWRKY-IIC6g		····T····STNNN <mark>P</mark> SAT	SS		• • • • • • • • • • • • • • • • • • • •	- SSEDPAE		
GmWRKY-IIC6n		······································	SS			- SSEDPPE		
CpWRKY-IIC6d		VSTSN-PSV	\$\$\$			- SSED <mark>PP</mark> E		
PtWRKY-IIC6i		<mark>v</mark> sttny <mark>p</mark> sv:	ss			- SSED <mark>PP</mark> E		
PtWRKY-IIC6j	•••••	····V····STSNY <mark>P</mark> SVS	ss		• • • • • • • • • • • • • • • • • • • •	- SSEDQ <mark>P</mark> D	• • • • • • • • • • • • • • • • • • • •	
ZmWRKY-IIC6e	FTADDDAYGADSSAMEA	A ST SN - PSV -	FIMMSMII <mark>P</mark>			- SSED <mark>PP</mark> E	<u> </u>	
ShWRKY-IIC61*	FTADD - AYYADSSAIFA	ELAG CWAAGGGGTTTI			PPEDHHAT PPPPI	CLASTPTRAVSVI)	
BdWRKY-IICGe*	FS <mark>G</mark> DE DSSALLA	E L <mark>G</mark> W- • • • • AAS • • F V VI	DCT - LQL <mark>PP</mark> L		· · · · · ELH <mark>PPPP</mark> DE	GGG AAASS		
OsWRKY-IIC6d*	FAADE AYADSSAIFA	EL <mark>GW · · · · · ANG · · LAV</mark>	DAVGELLPPLD	· · · · · · · · · · · · · · · · · · ·	· · PPGELAT PPPPP	LDL <mark>P</mark> ET <mark>P</mark> A····		
OsWRKY-IIC6e*	FDGDAAFAEYSS - AVLA			·····AALDLPVD		EA <mark>P</mark> ARSGD		
BAWRKY-NC64*	RSAASGGGAGAGAGAGAGAGAG	ELG ·····GWSGELPSI	GELODPVYASSPICO		BOVA	A 5 5 5 5 5 <mark>6</mark> D		
	No ne over a serie a s	<u>-</u>						

278

IIC7-

					YSGPTI						
	120 130	140	150	160	170	180	190	200	210	220 230	
PtWRKY-IIC1a	KNQSK <mark>P</mark> S <mark>P</mark> TNEQ	TINGLTSK					VYSG	TIQDIENALS-		TSSRED-QSQAVSQ	
VVWRKY-IIC1a	TVES-KLMSSSSEQ	AISRLVST -			• • • • • • • • • • •		····AYTG	PTIEDIESALS-	• • • • • • • • •	LTT <mark>G</mark> KC QSEENSQ	··
GmWRKY-IIC1a GmWRKY-IIC1a	····TTTKLASSSDDQDQ	AFNRFISN					· · · · · · · · · · · · · · · · · · ·	PTISDIENALS- PTISDIENALS-		VTNQRD HFPQLSS VTNQRDH - HFPQLSS	
AtWRKY-IIC1a	QHHMH <mark>G</mark> LM <mark>P</mark> NEDFI	ANKFVTST-					· · · · · · L Y S <mark>G</mark>	RIQDIANALA-		LVE <mark>P</mark> LTH <mark>P</mark> VREISK -	IICI-
OSWRKY-IIC1a	···· LQVEAPCGRRRRES	MVNKLIST-					·····VYSG	PTISDIESALS-		FTAAGDHQLLADGHN	FA
SBWRKY-IIC1a	QQEEEE <mark>P</mark> -RRQRES	MLNKLIST -					VYSG	PTISDIESALS-		- FTGADQAAAVDAHI -	••
CpWRKY-IIC1a	HRVI-NHGESKS-K	SNSHLISST	LY	• • • • • • • • • • • • •	•••••		•••••NYSG	PTIQHVQDALL-	• • • • • • • • •	LTASR	••
ShWRKY-IIC1b OsWRKY-IIC1b	GGGTAAEHRELLPOPE	AVSRALCS -					· · · · · · · · · · · · · · · · · · ·	PTIRDIEKALS- PTIRDIEKALS-		T T R P Y P WS WS S S R Y S S A S P R P P Y P S G R R -	
AtWRKY.IIC2a		OKNNHOSEL	NAUNNU		NDDD						
OsWRKY-IIC2		EEGDHRGRP	QM <mark>G</mark> NK <mark>G</mark> E · · · · ·	AAAAMGAM	GIND			- AGNNTAAA		- AAAQHHL	
ShWRKY-IIC2a	••••••••••••••••••••••••••••••••••••••	LLLDHHHHP	TTTTTSS · · A · ·	- ASSSHAASSI	TLHH			- HLHGHAAA	• • • • • • • • •	AA <mark>P</mark> HASLS <mark>PP</mark> TRAS <mark>P</mark>	P P
ZmWRKY-IIC2a* MtWRKY-IIC2a*	·····GVLGYGSAC	NP-OKSKE-	T A T S S A S <mark>P</mark> S H - - Y L T S S F <mark>G - - G</mark> - -	- AAPSSVVPAP - STPFLSLHGS			WA I	- H F H <mark>G</mark> H A A <mark>G</mark> L <mark>G G</mark> E V I S N C		A <mark>p</mark> h vslspptrasp MNN s <mark>g</mark> krngvdhrdn	<u> </u>
MtWRKY-IIC2b*	FA-ETLLAT	AV-QKTREF	DDHLTSSF <mark>GG</mark>	- GGQLLSLNRS	KVNS			WAWGEVTGS		L I <mark>G</mark> KRSS <mark>GG</mark> DDQHHH	нн
GmWRKY-IIC2a*	••••••••••••••••••••••••••••••••••••••	AV-QRPRE-	DLTSNLLA <mark>GG</mark>	- GGQLLSLNRS	RVNS			WAWEEVSDC		LMGKR - IGGDDN	·· IIC2
GmWRKY-IIC2e*	••••••••••••••••••••••••••••••••••••••	TA-QKSRE-	DLTSSFR	GAQFLSLHRL	SVNPCWDS		IMNEF	RAL <mark>G</mark> EVADC		••• FSTT•••••	
GmWRKY-IIC2d*	• • • • • • • • • • • LTAETLFST	TA-QKSRE-	D L T S S F <mark>G</mark>	- <mark>GG</mark> QFLSLHRS	S VN P			WAL <mark>G</mark> EVAEC		FSSKR - S <mark>G</mark> FDDH	
OSWRKY-IIC2b* OSWRKY-IIC2c*	····ATGT··NHGVMFSSD	HGCLYPLLP HCGLYPL-P	GIPFCLDSGCG -	- AAACDDDK <mark>P</mark> A	GFAHLGSA		·····EADTS	AAAARVDSE GGAFFVATT		IAAAAT - ATTCH <mark>gp</mark> n Vtka <mark>g</mark> nfsttcngst	SW TW
VvWRKY-IIC2a*	····LTEALLSS	APT - KQRED	• T <mark>P</mark> HL <mark>G</mark> • • • • •	- GAQLLSLQRS	TANL			WAW <mark>G</mark> EVNEC · · ·		LSSKRSI <mark>GG</mark> DD	
PtWRKY-IIC2a*	·····NLTETLLLS	SVTNKQRED	T ASDL <mark>G</mark>	- GPHLLSLQRS	SANL			WAWGEVNEC	• • • • • • • • •	LNSKRS <mark>g</mark> sggd	••
AtWRKY-IIC2b*	·····V···LTFFSSSS	SSSLSS P	S F <mark>P</mark> I HN S S	- STTTTHAPL	GFSNNLQG		GGP	LGSKVVNDD		QENF <mark>GGG</mark> TNNDAHSN	sw.
PtWRKY-IIG3a VVWRKY-IIG3a GpWRKY-IIG3a PtWRKY-IIG3a SaWRKY-IIG3a BdWRKY-IIG3a BdWRKY-IIG3a PtWRKY-IIG3a AtWRKY-IIG3a GmWRKY-IIG3a GmWRKY-IIG3a	SS SS SS SS SS SS SS SS SS SS SS SS SS S	SSLNMANP PP-NMANFH SFNEFKVTR SGGGGGGGFH TSGGGGGGFH TSGGGGGGCFH SSHDFN	IYFHG ENEN - A - STONG FLGLDH CMG - LGLDH OHDRGGGGHSGH HHHGQ CNS - HHHGQ CNS - HHHGQ CNS - SOG FLG - SOG FLG - SOG FLG - SOG FLG - NVH - SNELLGSDDH	10000 - SSSSF I AGG V P 5 SSSSF 6 GG SL SHG L F 10 GG G SL SHG L F	AELSSN AE <mark>R</mark> S		N N N RM L	GVR ESRT MN I BROLMINS DDSKEYGGA TSKOSGA NSKOGOD SKKENGA UNSKOGOD SKKENYHU LKTETGH V LKTETGH V GNSIENGFL DNGG GGGO		- QHS DAKHS D CTSR. - YKS AG. - YKS DRIKINGGS. - ASSPP AAS GR. - G DE GARSSPH GR. - VYG GAT SPH GR. - C DE GARSSPH GR. - YVD GAT RSPH GG. - PRT SEIN - KDL - PRT SFEIN - KDL - C ST PR SFEIN - C	IIC3-
CpWRKY-IIC3b	ASSSFPC	S L MNMAS <mark>P</mark> H	T <mark>g</mark> ynilh <mark>g</mark> ddq -					- GMKKMNN I		TRADDQDLDAKKT	
								DW	xxL		
AtWRKY-IIC4a	· · · · · · · · · · · · · · · · · · ·	FVDKTLMMM	PPS-TFSGE				• • • • • • • • • • •	VE <mark>P</mark> SSSSSW		Y <mark>P</mark> ESFHVHA <mark>P</mark>	
ATWRK Y-IIC4b BdWRKY-IIC4a	·····GENNNPFSS	PDDQQCPLA	A <mark>P</mark> SLIFSGD Elqq <mark>p</mark> qpppW	VCN <mark>PP</mark> LN				• • • • • • • • • • • • • • • • • • •		I <mark>M</mark> AGYHLSAQLENFR FQT <mark>PGP</mark>	
ShWRKY-IIC4a	· · · · · · · · · · · · · PPAVAAPQPP	EQHACSSDA	TTTSLV <mark>PG</mark> AATM	IM S C			P	PAAVDWASL		· · L L <mark>P</mark> R A <mark>P G</mark> T L H V G T T <mark>P</mark>	PP
ZmWRKY-IIC4a	······	QQHACT - DA	PPTSLAPGAA	MSCPP			· · · · · · · · · · P	PAAVDWASL		· · L L P · · · P ST S · · G · · ·	
GmWRKY-IIC49 GmWRKY-IIC4a	······································	QLQASSA PQ	<mark>G</mark> LHDD	.			!!	IDWVSL		L L . L . S . S . S . S . S . S . S . S	
AtWRKY-IIC4c								· · · · · · <u>-</u> · · · · ·		· · · · <mark>·</mark> · · · <mark>·</mark> · · · · · · · · ·	·· IIC4-
GmWRKY-IIC4b GmWRKY-IIC4c		PQE	• • • QSGLCD • • •					· · · IDWGNL · · ·		• • FSAQN <mark>G</mark> LLLN • • • • •	
VvWRKY-IIC4a	······································		AQILGD					I DWAT L		LSAQTGLSD	
PtWRKY-IIC4a	·····PSYLLTPLLL	PSSSLQYPS	IIE <mark>P</mark> QVL <mark>P</mark> D···				••••••	I DWV <mark>G</mark> L		L S G Q S Q L G E	
USWRKY-IIC4a ShWRKY-IIC4b	·····PPLAVAPAAS	AITSDGHQH ••DQHSTLI	GPLEVLEQP-PC TALQQQPSSSAC	SNN-LHPDGLV	PD		A	MISLDWQSL - <mark>P</mark> MVDWSAL		LQI CLQVP	
ZmWRKY-NC4b	·····QPSLLPPPSE	SDDQRNTLA	TALQQQPSSAC	NNNDLQLLPL -				VDWSAL		LQQHASLMGP	

279

IIC3-

IIC4-

	120	130 140	150	160	170	180	190	200	210	220	230
AtWRKY-IIC52	R	EDHSSS <mark>P</mark> TLSWSS:	SEKLVAAEITS					- PLQTSLA-		т s <mark>р</mark> м s	
OsWRKY-IIC5a		GAASYVA <mark>P</mark> TFGAAI	F RQQ	HLDLLDYLSDDQ	G V P		A <mark>P</mark> <mark>P P</mark>	AAV <mark>P</mark> SAS		- YVT <mark>P</mark> A P	AMAPAE
BdWRKY-IIC5a		AAPSFCPAPFNG-	v aa	QLDVLDYFSDD -	GGP		SA V <mark>P</mark>	GTFDT <mark>P</mark>		LQL <mark>PP</mark>	- RAPTEA -
ShWRKY-IIC5a		PAAPF FPFGTA	A	QLDVFECLSDE <mark>g</mark>	GAV		<mark>P</mark> A <mark>P P</mark>	AAV <mark>PG</mark> A		FAT <mark>PPPP</mark> L	. PLMPAE
ZmWRKY-IIC5a		PAAP F P SAAF GAV	A	QLDVFDCLSSDE	g v g		V <mark>P</mark>	AAV <mark>PG</mark> A		• • F A <mark>P P P</mark> • • •	PLMPAE
BdWRKY-IIC5b		FS <mark>P</mark> AL <mark>P</mark> TQ <mark>PPP</mark>	VTD <mark>P</mark> AA	QFDISEYLFDD <mark>g</mark>	I F A · · · · ·		AAT DAA	A P P S G A		- AVAAAMD <mark>G</mark> -	
OsWRKY-IIC5b	· · · · · · · · · · s	F SAALAAATT PPL	<mark>P</mark> D <mark>P</mark> S G S	QFDISEFFFDDA	P		<mark>P</mark> A	AVFNG A -		- PTAALPDG-	
ZmWRKY-IIC5b	· · · · · · · · · · · s	FTAALAPTTT · · ·	ALAE	HFDISEFLFDDA	AGA		<mark>GVAG</mark> A <mark>P</mark>	GVFADGAA-		- R <mark>P</mark> VVL <mark>P</mark> VP -	
ShWRKY-IIC5b	· · · · · · · · · · · s	FSAALA <mark>P</mark> TTT	T L A E	QFDISEFLFDDA	G		· · · · · · VA <mark>G</mark> A <mark>P</mark>	GVFADGSA-		- PVVVSDAA -	
BdWRKY-IIC5c		YQHRY <mark>P</mark> • • • • • • • • •		•• SDLEY••••				• • \$ AT <mark>P</mark> • • • •		• <mark>P</mark> V F L <mark>P G A G</mark> -	
OsWRKY-IIC5c		FSAAAIDANLFS-		E L D F D C S · · · ·				L <mark>P</mark> A		• <mark>P</mark> AQEY <mark>P</mark> • • •	
SHWRKY-IIC5c		F PAAVDSATAFS -	A	E L D D L H H • • • • •				• • F D Y S <mark>P</mark> A • •		• <mark>P</mark> I F T A V <mark>G</mark> A -	
ZmWRKY-IIC5c		FTPAAGDSTAFS - ·	s	ELDDLHH			<u>.</u>	• • F D Y S <mark>P</mark> A • •		- PIVTAAGA -	
BdWRKY-IIC5d	· · · · · · · · · · v v	L <mark>PPP</mark> LDDQYSLT - ·	LETFS <mark>G</mark>	GGG <mark>P</mark> VT <mark>GG</mark> NN			- H D L D <mark>g</mark> mvmdm	YRMS <mark>G</mark> S <mark>G</mark> F - ·		- <mark>P L P</mark> AT <mark>P P</mark> S -	
OsWRKY-IIC5d		FPPLMADHIVD		<mark>gggg</mark> csf <mark>g</mark> ef	<u>.</u>		• L E L <mark>G</mark> • • • • • •	HSVYSL -		- <mark>P</mark> L <mark>PPPP</mark> SQ -	
ZmWRKY-IIC5d		LVVEF <mark>PP</mark> RAAT		MADDCCYQF	GQE····		- M <mark>G</mark>	GAYCS <mark>P</mark>		- <mark>P</mark> VFDN <mark>G</mark> MTS	;
ShWRKY-IIC5d	· · · · · · · · · · · · · · · · · · ·	LVAEF <mark>PP</mark> TAAAT - ·	A	MADDYYYYFQF	GEE····		- M <mark>gg</mark> ara <mark>pg</mark> cg	GGYCS <mark>PP</mark> A-		- <mark>P</mark> AFDN <mark>G</mark> MS-	
GmWRKY-IIC5a	····s <mark>g</mark> HV	FSHNNQ - ANEVG -						- N F <mark>G G</mark> S S T		• H F E E S <mark>S</mark> • • •	
CpWRKY-IIC5a	<mark>G</mark> S N T	QNSLH I <mark>P</mark> TQ <mark>P</mark> VQF						PQINTTIT -		- T I VT Q <mark>P</mark> SR -	
VvWRKY-IIC5a	· · · · · SEFA	QN <mark>P</mark> IH <mark>P</mark> - GNEVD						- K <mark>P G</mark> S S S S		- QHER <mark>P</mark> A	
PtWRKY-IIC5a	S S Y A	CN <mark>P</mark> VYR - AHVV <mark>G</mark>						- E S <mark>G G</mark> S S S		- <mark>P</mark> REELS	
GmWRKY-IIC5b	SENV	SNQVHQ - VSNA <mark>G</mark>						- E F <mark>G G G</mark> S S		- HFE <mark>G</mark> LRCL -	
GmWRKY-IIC5c	SENV	LNQVHQ - ASNLG -						- E F <mark>G G G</mark> S S		- H F E <mark>G</mark> S S S I -	
AtWRKY-IIC5b	····S <mark>G</mark> MN	Q S Y <mark>G</mark> Y Q - T S D V A <mark>G</mark>	<u>.</u>	<u>.</u>				A L F S <mark>G</mark> S S S - ·		- C F S H <mark>P</mark> E S <mark>P</mark> -	<u> </u>
SbWRKY-IIC5e		SYLSLDDIVDDVV	G G E	WY R <mark>P</mark> L A E E S S S A	AA			AELQ <mark>P</mark> E <mark>P</mark> L - ·		- LFATLQAED) <mark>g</mark> ycvs <mark>g</mark> s -
OsWRKY-NC5e		SYLSFD MDDV -	E	Ү Ү Т <mark>Р</mark> Е V <mark>G</mark> F H S				KQHN <mark>PPP</mark> V-		- AAA <mark>P</mark> LEA <mark>GO</mark>	GREQSR
GmWRKY-IIC5d	HHQES	WSQSTETESSEKA						- TSSDASH -		- <mark>G F G</mark> D A T	
GmWRKY-IIC5e	N - Q E S	WSQSTETESSEK <mark>G</mark>						- NSSDVSH -		- <mark>G F G</mark> D A T F	
GmWRKY-IIC5f	H - QDS	RS-SQSTESSEKA						• T FNDATH • •		- <mark>G</mark> FST <mark>G</mark> A	
GmWRKY-IIC5g	H - QDS	RS - SQSTESSEKA						- T F C D <mark>P</mark> T H		- <mark>G</mark> F S T <mark>G</mark> A	
CpWRKY-IIC5b		· · · · PTPTSSTNK · ·									
AtWRKY-IIC5c	IEEE	ISS <mark>P</mark> TSIVSSE						- T F T <mark>G</mark> E S <mark>G</mark>		- <mark>G</mark> S <mark>G</mark> S A T T	
VVWRKY-IIC5b		•••KS <mark>G</mark> LVTRV•••				• • • • • • • • • • • •		· IFFYLFL · ·		• F F • • • • • • •	
VvWRKY-IIC5c	· · · · EE	DSCSQTTAAASAV						• T V T <mark>G</mark> • • • •		T <mark>G</mark> H I DQ	
VvWRKY-IIC5d											

IIC5-

		TPNSS										
	120	130	140	150	160	170	180	190	200	210	220	230
MIWRKY-IIC62		ssisss		E				ENEHE	HEEELEGD		EAEEKDEDH	
GmWRKY-IIC6a		5515555	EAAAAAAAA	ATGNDND -					TIDAAAGR		<mark>G</mark> EED	
GmWRKY-IIC6b		ssisss	EAAAATTANK	TTGNDNEE				EEDET	AIDATAGR		· · · · EEEDH · · · ·	
GmWRKY-IIC6c		ssisss	EATVNNNAEQ	SLKPGNDQ				· · · · · · · · EEAEA	EEEGGNGG		RGDD	
GmWRKY-IIC6d		5515555	EATVNNNSEQ	QSGKEL				· · · · · · · · · · · · · · · · · · ·	EEEGGNGG		· · · · RGDN · · · · ·	
AtWRKY-IIC6a		TSVSSSS	EAANDNNSGK						EVTVKDQE		· · · · E G D Q Q · · · ·	
VvWRKY-IIC6a		5515555	EAANDD						QTKAAE		· · · · EEE · · · · · ·	
CpWRKY-IIC6a		SSISSYS	NE AGNDDDHHH						HHQQSTKD		· · · · QEDE · · · · ·	
TWRKY-IIC62		ASISSS	IEAGEDA						FQQVKAGD		· · · · QEEE · · · · ·	
PtWRKY-IIC6b		ASISSS							FQQVKTGD		· · · · QEEE · · · · ·	
AtWRKY-IIC6b		SSISSAS	SEALNEEK · · ·						- PKTEDNE		EE <mark>GG</mark> EDQQ -	
AtWRKY-IIC6c		· · · · · s s	SEAVN <mark>G</mark> D						- DEEEEDG		· · · · EEQ · · · · · ·	
GmWRKY-IIC6e		SSISSAS	SEVFYDEQ						- NKT VDQA		<mark>Р</mark> ЕН	
GmWRKY-IIC61		SSISSAS	BEVLYDEQ						- NKT VDLA		<mark>Р</mark> ЕН	
VvWRKY-IIC6b		SSVSSAS	6HEQ <mark>g</mark> sk						AVEEE		···· EEDEE ····	
GmWRKY-IIC6g		SSISSAS	DAVNDEQ						- KKT L DQA		· · · · EEDDDDDD -	
SmWRKY-IIC6h		SSISSAS	DAVNDEQ						NKTLLDQA		· · · · EEDDDEE	
PtWRKY-IIC6c		SSISSAS	SE <mark>GLNDE</mark> PA						- NKGVDNE		· · · · VEE · · · · ·	
PtWRKY-IIC6d		SSISSAS	SE <mark>GLNDE</mark> PV						- NKSVDDE		· · · · VEE · · · · · ·	
MtWRKY-IIC6b		SSISSAS	BEAINDEH						- NKT VDQT		· · · · NNQ · · · · · ·	
CPWRKY-IIC6b		SSISSAS	BEAVNDE						KKVEED		· · · · QEED · · · · ·	
dWRKY-IIC6a		T S G G	DDGEG					KSRRCH	K-QVLAKD		· · · · N · EEEEEE	G
DsWRKY-IIC6a			EAEGG					SSRRCH			EDDKD	,
mWRKY-IIC6a		S S D R	EGEG					QLRRCF	KGRPKAED		D - AEGDEKE	
WRKY-IICGa		S S D R	. E G E G G					QPRRCH	KGRPKAED		AEGDEKE	
MWRKY-IIC6b		· · · s s s s	EAAGAAAGGG	GGGS FG				EEESQS	LGRCKKED		D - GDGEESH	DKEA
mWRKY-IIC6c		· · · s s s s	EAAGG - AAGG	666R				EEDSP-	QGRCKKEE			(A
WRKY-IIC68		· · · s s s s	EAAGGGAAGG	GGGGSFGG				EEDSPH	QGRCKKEE			(A
SWRKY-HC6D		TSSS	EGVGGGGGGG	AGAG					- ARCKKEE		D-ENKEEGH	(
dWRKY-IIC6b		· · · · s s s s	EAGGGGGGGS	6666				GDQGDS	AGRCKKEE			(
SWRKY-IIC6c		· · · ST S S	EACGVGGGAG						AGKCKKEE		E - GDGGDDD	GKEGS
WRKY-IICGc		· · · ST S S	EACGAGAGA-						AGKCKKED		· · · · · · · · GEEEGLE	SKDDGS -
mWRKY-IIC6d		· · · ST S S	EACGAGAG					EELAT	AGKSKKEE			SKDGDA-
tWRKY-IIC6d		· · · · s s s	4EGE					<mark>Р</mark> КЕ	NTNDKSDQ		· · · · MEDNEGDLH	.
tWRKY-IIC6e		· · · · s s s s	ADHH <mark>P</mark>						SGKIRKKR		EVRDG	
AtWRKY-IIC6f		· · · · s s s s s	ADH <mark>P</mark>						SGKSRRKR		· · · · ELV · · · · · ·	
GmWRKY-IIC6i		· · · · s s s	IEAEAII					EED	STKSQKDK		QPKGCE-E-	
GmWRKY-IIC6i		s <mark>P</mark> s						EED	STKSQKDK		QPKACE-AD	
CuWRKY-IIC6c		· · · · s s s	4D					ED	G SKKEK		QSKGSED	
VVWRKY-IIC6c		· · · · s s s	EAI					TED	SGKSKHKP		DLQGGGCED	
PtWRKY-IIC6e		· · · · s s s	10 <mark>6</mark> AT					EED	ракзккнк		QPKEST-ED	
PtWRKY-IIC6f		· · · · s s s						EED	SGKIKKDK		····QPKES··ED	
VVWRKY-IIC6d		S S T	AGA					EED	SSKSKKDR		QAKVSEL - 0	
PtWRKY-IIC6#		· · · · s s s s	AGG					DED	SGKTKKET		QPSRPED	
WRKY-IIC6		· · s s s s	AGG					DED	SGKTKKET		QPEKPED	
GmWRKY-IIC6k		· · · · s s s s	AGA					· · · · · · · · · · EED	SGKSKKER		····QVKTEE···	
SmWRKY-IIC6I		· · s s s s	AGA					EED	SGKSKKER		····QVKTEE···	
ShWRKY-IIC6d		- GDGAAP	NDD						- DDRKAA		P AAEAAG	
ZmWRKY-IIC6e		GAAPO	NGDGD						DDRKAA		PAAEAAG	
tWRKY-IIC6#		NST	ASAEKT						<mark>PPP</mark> E		т <mark>р</mark> vк	
SmWRKY-IIC6m		· · · · · KST	/SDD						KPPE		· · · · I PSK · · · · ·	
SmWRKY-IIC6n		····кsт	/SDE						· · · · K P P E · ·		· · · · I PSK · · · ·	
WRKY-IIC6d		кsт	s <mark>g g</mark>						· · · · K P P E · ·		· · · · I PNK · · · · ·	
tWRKY-IIC6i		кsт	s						· · · · KPPE · ·		· · · · I PST · · · ·	
WRKY-IIC6		кэт	s						KPPE		I <mark>P</mark> STK	
WRKY-IICGe		кsт	s s g g g						KPPE		IPSK	
mWRKY-IIC6f*		AAVS SI	RSSSTD						DDGAPAAA		· · · · ATTEAAS · ·	
SHWRKY-IIC61*		- GGGASS	ASST DD GAAAA	AAAQEEEEE					DDGAPAAA		· · · · AATEAAS · ·	
BdWRKY-IIC6e*		SSIDD	AAT PEP					TAG	ADG - KPAT		GATEAAS	
o		GSSAD	AASSCS						ADGGKPAA		· · · · ASTEAAS · ·	
USWKKY-11C6d*												
JSWRKY-IIC60* JSWRKY-IIC60*		GAAA/	AASSSSS <mark>G</mark>					E <mark>P</mark> A	APDKRPAA		AEAAPAA	
USWRKY-IICGd* OSWRKY-IICGe* BdWRKY-IICGd*		<mark>G</mark> AAA/ <mark>G</mark> ASA/	4 A S S S S S <mark>G</mark> 4 A E N A D						APDKRPAA KPAS		AEAA <mark>P</mark> AA AAAEAAS	

PpWRKY-IIC7a	SSTCTDAQLLADRSHS <mark>G</mark> T <mark>P</mark> ITSESLSSSLVSD <mark>G</mark> KTESK···································	
PpWRKY-IIC7b	PTTSSAELRLEEVYHSQT <mark>P</mark> IDSASLSSSFLSDSKTSFK	IIC7
PpWRKY-IIC7c	••••• <mark>P</mark> F MAVHE AD <mark>G</mark> SAAT Q <mark>P</mark> NN VE L SQN S L S L S <mark>G SGG</mark> D S D E EWVRNAV V •••••• SSE A <mark>G</mark> E R S <mark>G</mark> K R K S <mark>P</mark> NS SAT <mark>G</mark> •••••• E F R E AD F D K <mark>P</mark> S S D A E L	IIC/-

Appendix 4 continued KYTxK 320 260 280 290 310 330 340 ----- ARISLLQKGLSK MENN-----PtWRKY-IIC1a GmWRKY.UC1a GmWRKY-IIC1b ARVSILERGLSK IEN AUMAN', 1967 OSWRKY, 1967 BWRRY, 1967 BWRRY, 1967 IIC1- CpWRKY.IIC19 R F AMGRG FRQ S E S SbWRKY.IIC1b S P TMH LGR LGALS R AP E OSWRKY-IIC1b AtWRKY-IIC2a 0.50% RYY 1/1029 50% RYY 1/1029 50W RYY 1/1029 - HPWSTA <mark>0.60 PAP</mark>AHDRQA <mark>0.00 GR PP R0</mark>K GAAP VISE <mark>6</mark> SAAAAL GVGAV RMKKA GGGGGGGG GRKK VR EPRFCFK *GmWRKY-IIC2a** AR<mark>RK</mark>V<mark>REPR</mark>FCFH IIC2-PHWRKY.IIC34 VVWRKY-IIC3a ATNYNS CSTE GKAE GK. KKEE DK C#WRKY-IIC3a PtWRKY-IIC3b SHWRKY-HC3a ZmWRKY-IIC3a BdWRKY-IIC3a IIC3-OsWRKY-IIC3a PtWRKY-IIC3c PHWRKY-IIC3d VVWRKY-IIC3b - - RYAFO IRKPRYAFQ GmWRKY-IIC3b ELKEK<mark>R</mark>SR<mark>KVPR</mark>IAFH AtWRKY-IIC4a GGEMGG L VSNNSNNSDHNKN AtWRKY-IIC40 RdWRKY-IIC4a VASEVESGGS - SAVT VAGSSASATAAGEGDNNYKAGA ShWRKY-IIC4a IIC4-GORMK<mark>R</mark>TRVPRFAFQ GmWRKY-IIC4b OWNERY-ING4e ORANA I ECASS SS FVAQINEVACEEEKE ORANA I ECASS SS FVAQINEVACEEKE ORANA I ECASS SS FVAQINEVACEEKE VWREY-VIG4e OLANDA I ECASS SS FVAQINEVACEEKE NEEKEK ORANA I TEVER FAFQ VWREY-VIG4e LVT RAE I TS VMA EEEKES I KDREK ORANA I TEVER FAFQ VWREY-VIG4a KREY VTE SSMINA EN GAEKE ORANA I TEVER FAFQ

SEWRKY-IIC4E ZmWRKY-HC4b

	240	250	2	60	27		280	29	0	300	310	32	۱. <u>.</u>	330	340		
AtWRKY-IIC52		 				FE 10	DKDEIF	ĸĸĸĸĸĸċ	0 <mark>P</mark>							I I <mark>H</mark> VFK	
OsWRKY-IIC5a		 						GYPRSVA	AA			· · · AAAVA	E G		R D R T	TTD <mark>KIAF</mark> R	
BdWRKY-IIC5a		 				- PIVPO	VGGYFT	AHPRSTA	AA			AVAA				TTD <mark>KIAF</mark> R	
SbWRKY-IIC5a		 				- RVVPD	AAAGYS	SHARSAA	AA			AAGEG	<mark>РР</mark>		R	RTDRIAFR	
ZmWRKY-IIC5a		 				- RVV <mark>P</mark> D	AAA <mark>g</mark> ys	SHTRS - A	AA			• VAGEG	s		R	TT <mark>HR</mark> IAFR	
BdWRKY-IIC5b		 				VGAS	AVAAL	RS <mark>P</mark> ADQQC	QQ			QAAVEF	P			RTERIAFR	
OsWRKY-IIC5b		 				AAAN	I AT R	SAAEAV <mark>P</mark> A	P A			P AAVEF	P			RTERIAFR	
ZmWRKY-IIC5b		 				DAAG	GGAIIG	AAA <mark>gg</mark> aa/	AA			SEV PEF	P			RTTRIAFR	
SbWRKY-IIC5b		 				AAAG	GGGAIS/	AAA <mark>G</mark> SAAA/	AA			···EAV <mark>P</mark> EF	<mark>Р</mark>			R T E <mark>R I A F</mark> R	
BdWRKY-IIC5c		 				- DHH <mark>gg</mark>	EEEEE	KTRANSK	кк			AR - AIGGGF				<mark>I G F</mark> R	
OsWRKY-IIC5c		 					ENEM	IT MMRYE SE	E			EKMRAF			'	VN <mark>GRIGF</mark> R	
SBWRKY-IIC5c		 				<mark>G</mark> AGG	- DRNEF	(MMMWCE <mark>G</mark>	36 · · · · ·			DEKRLF			:	S S <mark>G R</mark> I <mark>G F</mark> R	
ZmWRKY-IIC5c		 				<mark>G</mark> AGG	GDRNEF	(- MMWCEGO	3G			DERRLF				S N <mark>G R</mark> I G F R	
BdWRKY-IIC5d		 				YYYH	I C C <mark>G</mark> NHN	INA <mark>G</mark> RKEEI	L D			· · · AAAF				G H R <mark>R I G F</mark> R	
OsWRKY-IIC5d		 				PVVV	A A G G NN E) Q Y <mark>G</mark> V S S S S	s s			· · · · · AAAT	<u>.</u>			- T <mark>S R</mark> I G F R	
ZmWRKY-IIC5d		 				LLNY <mark>G</mark> -	VEGDGF	R R P V G G P A C	9 <mark>6</mark>			TG NGGGF	<mark>P</mark> R			P A <mark>S R</mark> I G F R	
SbWRKY-IIC5d		 				LLSYGG	I V D G D G F	RR <mark>P</mark> MS <mark>GP</mark> A/	4 <mark>6</mark>			T GGN GGGG F	P		· · · · ·	P A <mark>S R</mark> I G F R	IIC3-
GmWRKY-IIC5a		 			· · · · <u>·</u> ·	S R D	VGNE	REKKE	VRD							· · · <mark>RV</mark> AFK	
CpWRKY-IIC5a	· · SA · · · ·	 · · · · T	ΤΑΤΤΙΝ	ко	• <mark>P</mark> (F <mark>P</mark> IIGE	IER <mark>E</mark> - ·	· · · · DKKE	VKE							<mark>RY</mark> AFK	
VvWRKY-IIC5a		 				S RN	ISES <mark>G</mark> - ·	· · · QKKKE#	AKE							· · · <mark>R</mark> VAF I	
PtWRKY-IIC5a		 				<mark>G G</mark> E	GEEG- ·	· · · REKKE	AKE							· · · <mark>R</mark> VAFK	
GmWRKY-IIC5b	· · YI · · · ·	 · · · · н	FQLLVC	LSTIN	IDMNLI	RCLYED) T S S <mark>G</mark>	· · · RENRE	VRE							· · · <mark>R</mark> VAFK	
GmWRKY-IIC5c	•••• • • • • •	 · · · · н	EMMEVO	F	I C S I	RCLYSH	IFQMTKO	MFRLKRSO	CKE							· · · <mark>R</mark> VAFK	
AtWRKY-IIC5b	· · ST · · · ·	 · · · · к	ΤΥΥΑΑΤ	AT		ASAD	NONK	· · · KEKKK	I К <mark>G</mark>			· · · · · · · · · · ·				· · · <mark>R</mark> VAFK	
SbWRKY-IIC5e		 				<mark>g g</mark> e q s s	AALAND	ONHORIDLI	ra			DGGSRF	L		· · LRSEI	H <mark>G - KIAF</mark> K	
OsWRKY-IIC5e		 				REAA	VN L <mark>G</mark> KM	ADR <mark>GP</mark> APVS	5 <mark>G</mark>			<mark>G</mark> AAT <mark>G</mark> G	v		<mark>P R</mark> S K I	N <mark>g sk</mark> i Afk	
GmWRKY-IIC5d		 				SNTNMH	IIKCQNS	6 G I K <mark>G</mark> KNA ∙				EVSC				· · · <mark>R I T</mark> F R	
GmWRKY-IIC5e		 				SNTNMH	IIKCENN	1 <mark>6</mark> I KRKKE ·				EVSC				· · · MITFR	
GmWRKY-IIC5f		 				TSKNNN	IINCKN	• <mark>G</mark> INENK <mark>G</mark> •				<mark>G</mark> VGP				<mark>RIAF</mark> R	
GmWRKY-IIC5g		 				TSKNNN	IMQCKN	• <mark>G</mark> I NENKR •				<mark>GVGP</mark>				· · · <mark>RIAF</mark> R	
CpWRKY-IIC5b		 				INIR	CKSYLE	E <mark>G</mark> VKKEKM∙				····EVGC				· · · <mark>RVAF</mark> R	
AtWRKY-IIC5c		 				L SKK	ES-TNF	R <mark>G</mark> SKE SDQ1	гк			••••ЕТ <mark>G</mark> ⊦				· · · <mark>RVAF</mark> R	
VvWRKY-IIC5b		 				L Y F C	RECKD	GAKRKKT				D L G F				<mark>R</mark> VAFK	
VVWRKY-IIC5c		 				· · L HT	ATPTHE	GVRRSKE S	S D			D <mark>G</mark> AF				· · · VVAFR	
VVWRKY-IIC5d		 					· · · · · N	1 <mark>G</mark> R								VFALR	

	240	250	260	270	280	290	300	310	320	330	340		
MtWRKY-IIC62				VENQN	I D Q D Q MKKQ L K <mark>P</mark>	кккн					· · · KN <mark>KK</mark> L	RPARVT FK	
GmWRKY-IIC6a				· · · · · · · · a	ADADKTKKALK <mark>P</mark>	ккк · · · · ·					Na <mark>кк</mark> a	REPREAFM	
GmWRKY-IIC6b					DODKTKKOLKP	ккк · · · · ·					NQKKQ	REPREAFM	
GMWRKY-IIC6c					DODKTKKOLK <mark>P</mark>	KKK · · · · ·					NQKKQ	REPREAFM	
AtWRKY-IIC62					FORGTER	***					NOKKA	REARFAFI	
VvWRKY-IIC6a					DQEKNKQQLK	ккк					NQ KRQ	REPREAFI	
CpWRKY-IIC6a				· · · · · · · · a	DODKSKKOLK <mark>P</mark>	ккк					NQ <mark>КК</mark> Q	REPRFAFM	
PtWRKY-IIC6a				· · · · · · · · a	I D Q D K N K K Q L K <mark>P</mark>	KKR • • • • •	• • • • • • • • • • • • •		• • • • • • • • • • •		SQ <mark>KR</mark> L	REPRFAFM	
PtWRKY-IIC66					DHDKTKKQLK <mark>P</mark>	ккк					NQKRQ	REPREAFM	
ATWRKY-IIC60				· · · · · · · · · · · · · · · ·	- KSHIKKQLKA	**************************************					QKRQ	KVPKVSE I	
GmWRKY-IIC6e					OKTKEQLKA	KKTN					0KR0	REPREAFM	
GmWRKY-IIC6f					QKTKEQLKA	ккти					0 KR 0	REPREAFM	
VVWRKY-IIC6b					- EKKTKKELK <mark>P</mark>	кктт s					a <mark>kr</mark> a	REPRFAFM	
GmWRKY-IIC6y				B	G Q H K T K K Q L K P	кктп					Q <mark>KR</mark> Q	REPREAFM	
GmWRKY-IIC6h				E	GOOKTKKOLKP	KKTN····					QKRQ	REPREAFM	
PTWRKY-IIC6d					- OFKNKKELKP	KKTN						REPREAEM	
MIWRKY-IIC6b					LNKQLKA	KKTN					QKKP	REARIAFM	
CPWRKY-IIC6b					- QKKTTKQLKA	<u>ккти</u>					0 KR 0	REPRFAFM	
BdWRKY-IIC6a				VKDLQD	D <mark>g</mark> enskk <mark>g</mark> nnk	<mark>6</mark> кк					- QKAE <mark>KR</mark> Q	R L P R V <mark>S</mark> F L	
OsWRKY-IIC62				EE	DGENSKK <mark>P</mark> NK <mark>P</mark>	K					- KKAE <mark>KR</mark> Q	ROPRVAFL	
ZmWRKY-HC6a				QE		··ĸ····					- KKAEKRO	PRVAFL	
ZmWRKY-UC6b				MKGFFD	DADTGKKGS - A	AGKGK					- GKGEKRO	ROPRIAFM	
ZmWRKY-IIC6c				LDKGEE	DADKGKKGSAA	AGKGK					- GKGEKRO	ROPREAFL	
SHWRKY-IIC68				· · · · LDK <mark>g</mark> ee	DADK <mark>g</mark> kk <mark>g</mark> s <mark>p</mark> a	AAKGK					- GKGEKRQ	ROPRFAFM	
OsWRKY-IIC6b				<mark>G</mark> EED) E <mark>G</mark> HKNKK <mark>G</mark> S - A	AK <mark>GG</mark> KA					- GKGEKRA	R Q P R F A F M	
BdWRKY-IIC6b				GDDD	E G D K S K K G A A A	GKGGKA					- GKGEKRP	ROPRFAFM	
ShWRKY-IICSC					GEGEEKNKKGG	ANKGK					- GKGEKRP	ROPREAEM	
ZmWRKY-IIC6d				··· AAAKGD ·	AQGQEKSKKGA						- VKGEKRP	ROPREAFM	
AtWRKY-IIC6d					GVGESSKQLTK	Q <mark>G</mark> K					- KK <mark>g</mark> e <mark>kk</mark> e	REVRVAFM	
AtWRKY-IIC6e					GEDDQRSQKVV	ктк					- KKEE <mark>KK</mark> -	KE PRVSFM	
AtWRKY-IIC61					GEEDQISKKVG	ктк					- KTEVKKQ	REPRVSFM	
GMWRKY-IIC6I					GDEKSKKERFR						- KKKEKKP	REPRESEL	
C#WRKY-IIC6c					GD-HDKSKKMM	NKP					- KK - EKRO	REPREAFL	
VVWRKY-IIC6c					GDENSKKAN	KSK					- KKGEKRP	KEPRFAFM	
PtWRKY-IIC6e					<mark>g</mark> d i d <mark>g</mark> dakkvs	ктк					- KK - E <mark>KR</mark> Q	KEPRFAFL	
PtWRKY-IIC61					G DGDAKKVS	ктк	• • • • • • • • • • • • •				- KK - E <mark>KR</mark> Q	KEPRFAF L	
VVWRKY-IIC6d					GGDGSKKVN	KPK					- KKAEKRO	REPREAFM	
PIWRK T-IICEB					GGENSDKKD	KAK					- KKAEKRO	KEPREAEM	
GmWRKY-IIC6k					GEENSKKGN	KEK					KKGEKKO	KEPRFAFM	
GmWRKY-IIC6I					GGENSKK <mark>G</mark> N	KEK • • • • •					- кк <mark>бекк</mark> а	KEPRFAFM	
ShWRKY-IIC6d					· · · RK <mark>P</mark> AAATA ·	• • • • • • • •					- KKGQKRP	ROPRFAFM	
ZmWRKY-HC6e					RK <mark>P</mark> AA						- KK <mark>G</mark> QKRA	ROPRFAFM	
GmWRKY-HC6g											- KKAQKRI	ROPREAEM	
GmWRKY-IIC6n					<mark>g</mark> K						- KKGQKRI	ROPREAEM	
CpWRKY-IIC6d					VR						- KK <mark>GQKR</mark> V	ROPRFAFM	
PtWRKY-IIC6i					· · vĸ · · · · ·						- KK <mark>GQKR</mark> I	R Q P R F A F M	
PtWRKY-IIC6j					··· VK · · · · · · ·						- KKGQKRI	RQPRFAFM	
VvWRKY-IIC6e					GR	 МТКСС.					- KKGQKRI	ROPREAEM	
∠mwkkt-1-1067* SbWRKY-1104**					··· KPPARGKTT	• TKSS					- SAGUKRA	ROPREAFM	
BdWRKY-IIC6c*					- KPAPAPGR	- к <mark>с</mark> о						ROPREAFM	
OsWRKY-IIC6d*					KSLT P						GKKRA	ROPRFAFM	
OsWRKY-IIC6e*					· · AATATATA-						- KKGQKRA	RQPRFAFM	
BdWRKY-IIC6d*					MKPAAAT		• • • • • • • • • • • • •		• • • • • • • • • • • •		- RKGQKRA	RQARFAFM	
USWRK Y-IIC6f				···EEEEEI	DUE <mark>G</mark> SAVQSCK						- KKGAKKE		

PpWRKY-IIC7a	· · · · · · · · · · · · · · · · · · ·	
PpWRKY-IIC7b	AT I SAENKEQERSSE ···KE ···· ANVHNEQ	IIIC7-
PpWRKY-IIC7¢	MQSKYSESATLRRRKKNVKTLRK <mark>P</mark> I <mark>Yai</mark> e	<u> </u>

284

Appendix	x 4 continued	WRKY DNA-binding domain		
		· · · · · · · · · · · · · · · · · · ·		
	KYTxK			
	360 370	380 390 400	410 420 430 44	10 450 460
PtWRKY-IIC1a	LKRCGD GMADDGYKWRKYGOKS I KN	<mark>ISPH</mark> P RSYYRCTNARCSA <mark>KKQ</mark> VERCSED	P - DTLVITYEGLHLHYTYPYFLSN	<mark>0P</mark> 0
VvWRKY-IIC1a	IKSCGN GMADDGYKWRKYGOKSIKN	SPNP RSYYRCTNPRCSAKKOVEKSSED	P · DTLIITYEGLHLHFAYPFFLIS · · · · · ·	
GmWRKY-IIC12 GmWRKY-IIC15	IKCFGN GMGDDGYKWRKYGOKSIKN	SPNP RSYTRCINPRCSARROVERSNED	P.DTLIITYEGLHLHFAYPYFLMG	
AtWRKY-IIC1a	VKNNSNGMCDDGYKWRKYGQKSIKN	<mark>ISPNP RSYYKCTNP</mark> I CNA <mark>kkqver</mark> s I de	S · NTYIITYEGFHFHYTYPFFLPD · · · · · ·	
OsWRKY-IIC1a	M <mark>KS</mark> CGN <mark>GGLADDGYKWRKYGQKS</mark> IKN	NSPNP RSYYRCTNPRCNAKKQVERAVDE	P - DTLIVTYEGLHLHYTYSHFLHSTSSSSST	T T Q Q Q <mark>L</mark> Q <mark>P</mark> Q MMT N -
BdWRKY-IIC1a ShWRKY-IIC1a	IKTCGN GLADDGYKWRKYGQKSIKN	SPNP RSYYRCTNPRCNAKKOVERSTEE	P · DTLLVTYEGLHLHYTYSHFLQAHHHPA · · ·	
CpWRKY-IIC1a	IKSCGN GMADDGYKWRKYGQKS IKN	SPNP RSYYKCTNPRCSAKKQVEKSNEE	P - DTLIVTYEGLHLHFVYPFSLV	
SHWRKY-IIC1b	V <mark>RS</mark> CGGK <mark>TP</mark> SDGY <mark>KWRKYGQKS</mark> IKN	INPHP RSYYKCTSSRCGAKKHVEKSTED	<mark>P</mark> - EMLMV <mark>TYEGPH</mark> L <mark>HGP</mark> Q <mark>P</mark> LF <mark>P</mark> RRQWLSIDLS	GAAAAAAAASKT
OsWRKY-IIC1b	VRSCGG KMP ADGYKWRKYGQKS I KN	NPHP R CAT RSIID	<u>P</u> · I <u>· · · · · · · · · · · · · · · · · </u>	
			DMAT	
			RM VI	
AtWRKY-IIC2a	TLSEVDVLDDGYRWRKYGOKVVKN	TQH PRSYYRCTQDKCRVKKRVERLADD	P-RMVITTYEGRHLHSPSNHLDDDSLS	T S H L H P P L S N F F W-
OsWRKY-IIC22	TMSDVDVLDDGY <mark>K</mark> WRKYGQKVVKN	<mark>TTQH</mark> PRSYYRCTQDNCRVKKRVERLAED	<mark>P</mark> - RMV I TTYEGRHVHSPSRD - DDD	••••••••••••••••••••••••••••••••••••••
SBWRKY-IIC2a	TMSDVDVLDDGY <mark>KWRKYGQ</mark> KVVKN	TQH PRSYYRCTODNCRVKKRVERLAED	P - RMV ITTYEGRHVHSP SRDDDD	AARANAEMSFIW-
ZmWRKY-IIC2a* MWRKY-IIC2a*		TOH PRSYYRCTODNCRVKKRVERLAED	P - RMV ITTYEGRHVHSPSRDEDDD	
MtWRKY-IIC2b*	TMSDVDVLDDGYKWRKYGQKVVKN	TOH PRSYYRCTODNCRVKKRVERLAED	P - RMV ITTYEGRHAHSP SNELEE	·····sqTqSELTNFFW-
GmWRKY-IIC2a*	* TMS DVD VL DDG YKWRK YGQK VVKN	NTQH PRSYYRCTQDNCRVKKRVERLAED	P - RMV ITTYEGRHVHSPSNELED	•••••SQ <mark>TP</mark> SELSNFL₩•
GmWRKY-IIC2b*	TMSDVDVLDDGYKWRKYGQKVVKN	TOH PRSYYRCTODNCRVKKRVERLAED	P - RMV ITTYEGRHVHSPSNELED	·····SQTPSELSNFLW-
GmWRKY-IIC2¢*		TOH PRSYYRCIQDNCRVKKRVERFAED	P - IKMV ITTYEGRHVHSPSNDLED.	
OsWRKY-NC2b*		SLH PRSYYRCTHNNCRVKKRVERLSED	C - RMV ITTYEGRHTHTPCSDDAT	TG
OsWRKY-IIC2c*	TRSEVDVLDDGYKWRKYGQKVVKN	<mark>NSLH</mark> <mark>PRSYFRCT</mark> HSNCRV <mark>KKR</mark> VERLSTD	C - RMV I TTYEGRHTHSPCDDNSS	
VvWRKY-IIC2a*	TMSEVDVLDDGY <mark>KWRKYGQK</mark> VVKN	NTQH PRSYYRCTQDNCRVKKRVERLAED	P · RMV ITTYEGRH I HSP SHDLEE · · · · · · · ·	· · · · · · · · · · · · · · · · SQ <mark>AP</mark> SHLNNFFW·
PtWRKY-IIC2a* PtWRKY-IIC2b*		TOH PRSYYRCTODNCRVKKRVERLAED	P-RMVITTYEGRHAHSPSHDLEE	
AtWRKY-IIC2b*	TKS DVDVLDDGYKWRKYGQKVVKN	SLH PRSYYRCTHNNCRVKKRVERLSED	C - RMV ITTYEGRHNH I PSDDSTS	
				FEHH
		•	V	FEHIL
PtWRKY-IIC3a	TRS····QVD I L DDGYRWRKYGQKT VKA	IS KF PRSYYRCT NNGCNV <mark>KKQVQR</mark> NSKD	E · E I VVTTYEGMHTHP I EKCIDN · · · · · · ·	FEHIL
PtWRKY-IIC3a VvWRKY-IIC3a	TRSQVD I LDDGYRWRKYGOKTVKN TRSHVD I LDDGYRWRKYGOKAVKN	<mark>18</mark> k F <mark>Prsyyrct</mark> nngcnv <mark>kkovor</mark> nsko NK F <mark>Prsyyrct</mark> ykocnv <mark>kkovor</mark> lsko	E - E I VVTTYEGMHTH <mark>P</mark> I EKCT <mark>DN</mark>	FEHIL IEDILROMOTY FEHILROMOSYF <mark>P</mark> I
PtWRKY-IIC3a VVWRKY-IIC3a CpWRKY-IIC3a PWRKY-IIC3a	TRSBUD IL DDG YRWRK YGG KTUK M TRSHUD IL DDG YRWRK YGG KTUK M TRSRUD IL DDG YRWRK YGG KAUK M TRSRUD IL DDG YRWRK YGG KRUK M	NSKF PRSYYRCTNNGCNVKKQVGRNSK NKF PRSYYRCTNNGCNVKKQVGRLSK NKF PRSYYRCTHKGCNVKQQILSKC NKF PSYYRCTHKGCNVKQQIII.	E - E IVVTTYEGINHTHP I EK CI <mark>DN</mark>	FEHIL IEDILROMOTY FEHILROMOSYF <mark>F</mark> I
PtWRKY-IIC3a VyWRKY-IIC3a CµWRKY-IIC3a PtWRKY-IIC3a SbWRKY-IIC3a	TRSBVD I LDDGYRWRKYGOKTVK TRSRVD I LDDGYRWRKYGOKAVK TRSRVD I LDDGYRWRKYGOKAVK TRSRVD I LDDGYRWRKYGOKAVK TRSRVD I LDDGYRWRKYGOKAVK	NSK F PRSYYRCTNNGCNVKKQVORNSKO NK F PRSYYRCTYKDCNVKKQVORLSKD NK Y PRSYYRCTYKDCNVKKQIGLISKD NK F PRSYYRCTHGCNVKKQVORLISKD	E E IVVTTYEGMHTHPIEKC10N E E IVVTTYEGIHTHPVEK <mark>e</mark> ten E Givvttyegihskoinston E Gvvvttyeghhskoieks <mark>e</mark> on C Ovvvttyeghtheikskon	FEHIL IEDILROMOTY FEHILROMOSYF <mark>F</mark> I FEHILSOMOIYS. FEHILSOMOIYS. FEHILSOMOIYS.
PtWRKY-IIC3a VVWRKY-IIC3a CpWRKY-IIC3a PtWRKY-IIC3a SbWRKY-IIC3a ZmWRKY-IIC3a	TRSBVD I LDDGYRWRKYGO KIVK TRSRVD I LDDGYRWRKYGO KAVK TRSRVD I LDDGYRWRKYGO KAVK TRSRVD I LDDGYRWRKYGO KAVK TRSQVD I LDDGYRWRKYGO KAVK TRSQVD I LDDGYRWRKYGO KAVK	ISK F PRSYYRCTNNGCNVKKQVQRNSKD NKF PRSYYRCTNNGCNVKKQVQRLSKD NKF - PRSYYRCTHKGCNVKKQVQRLSKD NNF PRSYYRCTHQGCNVKKQVQRLSRD NNF - PRSYYRCTHQGCNVKKQVQRLSRD	E - E IVVTTYEGMHTHPIEKCIDN E - E IVVTTYEGIHTHPIKKPIEN - GIVVTTYEGHSMOIGNSIDN - GVVVTYEGHSMOIKSMOIKSPDN - GVVVTTYEGTHTHPIEKSNDN - GVVVTTYEGTHTHPIKSNDN	FEHIL IEDILRQMQTY FEHILRQMQSYF <mark>H</mark> I FEHILRQMQSYFB- FEHILSQMQIYS- FEHILSQMQIYS <mark>0</mark> - FEHILTQMQIYS <mark>0</mark> -
PtWRKY-IIC3a VyWRKY-IIC3a GpWRKY-IIC3a PtWRKY-IIC3a SaWRKY-IIC3a BdWRKY-IIC3a BdWRKY-IIC3a	TRSGVD ILDDGYRWRKYGGKTVKN TRSHVD ILDDGYRWRKYGGKAVKN TRSRVD ILDDGYRWRKYGGKAVKN TRSRVD ILDDGYRWRKYGGKAVKN TRSGVD ILDDGYRWRKYGGKAVKN TRSGVD ILDDGYRWRKYGGKAVKN	NEKF PRSYYRCTNNGCNYKKQYGRNSK NKF PRSYYRCTYKDCNYKKQYGRLSKD NKF PRSYYRCTYKDCNYKKQYGRLSKD NKF PRSYYRCTHQCNYKKQYGLLSKD NKF PRSYYRCTHQCNYKKQYGLLSRD NKF PRSYYRCTHQCNYKKQYGLLSRD	E - E I VVTTYEG MHT HP I EK CTDN E - E I VVTTYEG I HT HP VEK FT EN. E - G I VVTTYEG I HT HP VEK FT EN. E - G VVVTYEG I HT HP I EK STDN. E - G VVVTTYEG TH TH P I EK STDN. E - G VVVTTYEG TH TH P I EK STDN. E - G VVVTTYEG TH TH P I EK STDN.	FEHIL IEDILROMOTY FEHILROMOSYFFI FEHILTOMOIYTS - FEHILTOMOIYS - FEHILTOMOIYS - FEHILTOMOIYS - FEHILTOMOYS -
PtWRKY-IIC3a VvWRKY-IIC3a CgWRKY-IIC3a PtWRKY-IIC3a SaWRKY-IIC3a BdWRKY-IIC3a OsWRKY-IIC3a BdWRKY-IIC3a	TRS OVDILDDGYRWRKYGOKTVK TRS HVDILDDGYRWRKYGOKPVK TRSRVDILDDGYRWRKYGOKPVK TRSRVDILDDGYRWRKYGOKPVK TRSQVILDDGYRWRKYGOKAVK TRSQVILDDGYRWRKYGOKAVK TRSQVILDDGYRWRKYGOKAVK TRSQVILDDGYRWRKYGOKAVK	ISKF PRSYYRCTNNGCNYKKQVGRNK NKF PRSYYRCTYKDCNVKKQVGRLSKD NKF PRSYYRCTHKGCNVKKQIGRLFKD NKF RSYYRCTHAGCNVKKQVGLLFKD NNF - PRSYYRCTHAGCNVKKQVGRLSRD NNF PRSYYRCTHAGCNVKKQVGRLSRD NNF PRSYYRCTHAGCNVKKQVGLSRD NKF PRSYYRCTHAGCNVKKQVGLSRD	E - E I VVT TYEGMHTHP I EKCTDN. - E I VVT TYEG I HTHPVEKPTEN. - G I VVT TYEG I HSHQ I ONSTDN. - G I VVT TYEGHHSHQ I EKSTON. - OVVTTYEGTHTHPIEKSNON. - OVVVTTYEGTHTHPIEKSNON. - OVVVTTYEGTHTHPIEKSNON. - OVVVTYEGTHTHPIEKSNON. - OVVVTYEGTHTHPIEKSNON. - TVVVTTYEGTHTHPIEKSNON.	FEHIL IEDILROMOTY FEHILROMOSYFEI FEHILROMOSYFEI FEHILSOMOINTS FEHILSOMOINTS FEHILSOMOINTS FEHILSOMOINTS FEHILSOMOINTS FEHILSOMOINTS FEHILSOMOINTS FEHILSOMOINTS FEHILTOMOINTS FEHILTOMOINTS FEHILTOMOINTS FEHILTOMOINTS FEHILTOMOINTS FEHILTOMONINTS FEHILTOMOINTS FEHILTOMONINTS FEHILTOMOINTS FEHILTOMONINTS FEHILTOMONINT FEHILTOMONINTS
PłWRKY-IIG3a VWWRKY-IIG3a PłWRKY-IIG3a PłWRKY-IIG3a SaWRKY-IIG3a BaWRKY-IIG3a OsWRKY-IIG3a PłWRKY-IIG3a	TRSBVD I LDDGYRWRKYGDKTVK TRSRVD I LDDGYRWRKYGDKAVK TRSRVD I LDDGYRWRKYGDKAVK TRSRVD I LDDGYRWRKYGDKAVK TRSQVD I LDDGYRWRKYGDKAVK TRSQVD I LDDGYRWRKYGDKAVK TRSQVD I LDDGYRWRKYGDKAVK TRSQVD I LDDGYRWRKYGDKAVK TRSQVD I LDDGYRWRKYGDKAVK	NKF PRSYYRCTNNGCNVKKQVORNSKO NKF PRSYYRCTYKDCNVKKQVORLSKD NKF PRSYYRCTYKDCNVKKQIGLISKD NKF PRSYYRCTHAGCNVKKQVORLSKD NNF PRSYYRCTHAGCNVKKQVORLSRD NNF PRSYYRCTHAGCNVKKQVORLSRD NKF PRSYYRCTHAGCNVKKQVORLSRD NKF PRSYYRCTHAGCNVKKQVORLSRD NKF PRSYYRCTYAGCTVKKQVORLSRD	E - E IVVTTYEGMHT HP I EKC10N E - E IVVTTYEG I HT HP VEKFTEN E - G IVVTTYEG HS NG I ANSTON E - QVVVTTYEG HS NG I EKSTON E - QVVVTYEG TH TH I EKSNON E - QVVVTYEG TH THP I EKSNON E - QVVVTYEG TH THP I EKSNON E - TVVVTTYEG TH THP I EKSNON E - QVVVTYEGMHT HP I EKSNON	FEHIL IEDILROMOTY FEHILROMOSYFEI FEHILSOMOTYS FEHILSOMOTYS FEHILSOMOTYS FEHILTOMOTYS FEHILTOMOTYS FEHILTOMOTYS FEHILTOMOTYS FEHILTOMOTYS FEHILTOMOTYS FEHILTOMOTYS FEHILSOMOTYS FEHILSOMOTY FEHILSOMOTY
PWWRKY-IIC3a CyWWRKY-IIC3a PWWRKY-IIC3a PWWRKY-IIC3a SaWWRKY-IIC3a BdWWRKY-IIC3a OsWWRKY-IIC3a PWWRKY-IIC3a VWWRKY-IIC3a	T R S B. VD I L DDGY RWRKYGBKTVKN T R S H VD I L DDGY RWRKYGBKAVKN T R S R VD I L DDGY RWRKYGBKAVKN T R S R VD I L DDGY RWRKYGBKAVKN T R S R VD I L DDGY RWRKYGBKAVKN T R S R VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN T R S G VD I L DDGY RWRKYGBKAVKN	NKF PRSYYRCT NNGCNYKKQVGRNSK NKF - PRSYYRCT YKDCNYKKQVGRLSKD NKF - PRSYYRCT YKDCNYKKQVGLLSKD NKF - PRSYYRCT HACKNYKKQVGLLSKD NNF - PRSYYRCT HACKNYKKQVGLLSKD NKF - PRSYYRCT HACKNYKKQVGLLSKD NKF - PRSYYRCT HACKNYKKQVGLLSKD NKF - PRSYYRCT HACKNYKKQVGLLSKD NKF - PRSYYRCT YACGTYKKQVGLLSKD NKF - PRSYYRCT YACGTYKKQVGLLSKD	E - E I VVTTYEG MHT HP I EKC TDN. E - E I VVTTYEG I HT HP I EKC TDN. E - E I VVTTYEG I HT HP VEK TTEN. E - OVVTYEG MS K0 I GNSTDN. E - OVVTYEG MS K0 I EKSTDN. E - OVVTYEG HT HP I EKSTDN. E - OVVTTYEG TH TH PI EKSTDN. E - OVVTTYEG TH HP I EKSTDN. E - OVVTTYEG HT HP I EKSTDN. E - OVVTTYEG MS HP I EKSTDN. E - OVVTYEG HP I EKSTDN	FEHIL IEDILROMOTY FEHILROMOSYF FEHILROMOSYF FEHILOSMOIYS FEHILTOMOIYS FEHILTOMOIYS FEHILTOMOYS FEHILTOMOYS FEHILTOMOYS FEHILTOMOYS FEHILTOMOYS FEHILTOMOYS FEHILTOMOYS FEHILTOMOYS FEHILSOMOIN FEHILSOMOIN FEHILSOMOIN FEHILSOMOIN
PtWRKY-IIC3a VVWRKY-IIC3a GpWRKY-IIC3a PtWRKY-IIC3a BaWRKY-IIC3a BdWRKY-IIC3a BdWRKY-IIC3a PtWRKY-IIC3a PtWRKY-IIC3a PtWRKY-IIC3a AtWRKY-IIC3a	TRS GVD I LDDGYRWRKYGGKTVKN TRS HVD I LDDGYRWRKYGGKTVKN TRS HVD I LDDGYRWRKYGGKAVKN TRS RVD I LDDGYRWRKYGGKAVKN TRS QVD I LDDGYRWRKYGGKAVKN	NEF PRSYYRCT NNGCNYKKQYGN NS NKF PRSYYRCT YKDGNYKKQYGN LSKD NKF PRSYYRCT YKDGNYKKQYGN LSKD NKF PRSYYRCT HQGCNYKKQYGN LSKD NKF PRSYYRCT HQGCNYKKQYGN LSKD NKF PRSYYRCT HQGCNYKKQYGN LSKD NKF PRSYYRCT HQGCNYKKQYGN LSKD NKF PRSYYRCT YQGCNYKKQYGN LSKD	E - E I VVTTYEG MHT HP I EKCTON. E - E I VVTTYEG I HTHPVEK FTEN. E - E I VVTTYEG I HTNPVEK FTEN. E - E VVVTTYEG MS K0 I GNSTON. E - E VVVTTYEG HTHPIEKSNON. E - E VVVTTYEG THTPIEKSNON. E - E VVVTTYEG THTPIEKSNON. E - TVVVTYEG HTNPIEKSNON. E - TVVVTYEG HTNPIEKSNON. E - E VVVTYEG HTNPIEKSNON. E - VVVTYEG HTNPIEKSNON. E - VVVTYEG HTNPIEKSNON. E - VVVTYEG HTNPIEKSNON. E - VVVTYEG HTN	FEHIL IEDILRAMATY FEHILRAMASYF FEHILRAMASYF FEHILSAMAIYS FEHILTAMAIYS FEHILTAMAIYS FEHILTAMAIYS FEHILTAMAIYS FEHILTAMAIYS FEHILTAMAIYS FEHILTAMAIYS FEHILTAMAIYS FEHILTAMAYS FEHILTAMAYS FEHILTAMAYS FEHILTAMAYS FEHILSAMAIY FEHILSAMAIY FEHILSAMAIY FEHILSAMAIY FEHILSAMAIY FEHILSAMAIY FEHILSAMAIY FEHILSAMAIY FEHILSAMAIY
PtWRKY-IIG3a VvWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a SBWRKY-IIG3a BdWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a VvWRKY-IIG3a VvWRKY-IIG3a	TRS BVD ILDDGYRWRKYGOKTVK TRS HVD ILDDGYRWRKYGOKTVK TRS RVD ILDDGYRWRKYGOKAVK TRS RVD ILDDGYRWRKYGOKAVK TRS GVD ILDDGYRWRKYGOKAVK	ISKF PRSYYRCTNNGCNYKKQVGRNSK NKF PRSYYRCTNKGCNYKKQVGRLSKD NKF PRSYYRCTHKGCNYKKQIGRLSKD NKF RSYYRCTHAGCNYKKQVGLLSKD NNF PRSYYRCTHAGCNYKKQVGLSRD NNF PRSYYRCTHAGCNYKKQVGLLSRD NKF PRSYYRCTHAGCNYKKQVGLLSRD NKF PRSYYRCTYGCTVKKQVGLLSRD NKF PRSYYRCTYGCTVKKQVGLLSRD NKF PRSYYRCTYGCNYKKQVGLLSRD NKF PRSYYRCTYGCNYKKQVGLLSRD NKF PRSYYRCTYGCNYKKQVGLLSRD NKF PRSYYRCTYGCNYKKQVGLLSRD NKF PRSYYRCTYGCNYKKQVGLLSRD	E - E I VVT TYEGMHTHP I EKCTDN. - E I VVT TYEG I HTHP VEKPTEN. - G I VVT TYEG I HSHQ I QNSTDN. - G I VVT TYEG I HSHQ I QNSTDN. - OVVTTYEG HTHP I EKSNDN. - OVVTTYEG THTP I EKSNDN. - OVVTTYEG THTP I EKSNDN. - OVVTTYEG THTP I EKSNDN. - OVVTTYEG THTP I EKSNDN. - OVVTTYEG MISHP I EKSNDN. - OVVTTYEG MISHP I EKSNDN. - OVVTTYEG MISHP I EKSNDN. - OVVTTYEG MISHP I EKSNDN. - OVVTYEG VITHE I EKSTDN. - OVVTYEG VITHE I EKSTEN. - E OVVTYEG VITHE I EKSTEN. - SIVVTYEG VITHEN - SIVVTYEG VITHEN I EKSTEN. - SIVVTYEG VITHEN - SIVVYEG VITHEN - SIVVTYEG VITHEN - SIVVYEG VITHEN - SIVVYEG VITHEN - S	FEHIL IEDILROMOTY FEMILROMOSY FEMILROMOSY FEMILSOMOTYS FEMILSOMOTYS FEMILSOMOTYS FEMILSOMOTYS FEMILSOMOTYS FEMILTOMOTYS FEMILTOMOTYS FEMILTOMOTYS FEMILTOMOTYS FEMILSOMOTY
PtWRKY-11G3a VVWRKY-11G3a GpWRKY-11G3a PtWRKY-11G3a BaWRKY-11G3a BaWRKY-11G3a PtWRKY-11G3a PtWRKY-11G3a PtWRKY-11G3a VVWRKY-11G3a GmWRKY-11G3a GmWRKY-11G3a GmWRKY-11G3a	T R S B. VD I L DDGY RWRKYGBKTVKN T R S HVD I L DDGY RWRKYGBKTVKN T R S RVD I L DDGY RWRKYGBKAVKN	NKF PRSYYRCT NNGCNYKKQVGRISK NKF - PRSYYRCT YKDCNYKKQVGRISKD NKF - PRSYYRCT YKDCNYKKQVGRISKD NKF - PRSYYRCT HACKNYKQVGRISKD NNF - PRSYYRCT HACKNYKQVGRISKD NKF - PRSYYRCT HACKNYKQVGRISKD NKF - PRSYYRCT HACKNYKQVGRISKD NKF - PRSYYRCT HACKNYKQVGRISKD NKF - PRSYYRCT YACKNYKQVGRISKD NKF - PRSYYRCT YACKNYKQVGRISKD NKF - PRSYYRCT YACKNYKQVGRISKD NKF - PRSYYRCT HACKNYKQVGRISKD NKF - PRSYYRCT MACKNYKQVGRISKD NKF - PRSYYRCT MACKNYKQVGRISKD	E - E I VVTTYEG MHT HP I EKC TDN. E - E I VVTTYEG I HT HP I EKC TDN. E - E I VVTTYEG I HT HP VEK TTEN. E - OVVTTYEG MHS NG I GNSTDN. E - OVVTTYEG MTS NG I EKSTDN. E - OVVTTYEG TH TH FI EKSTDN. E - OVVTTYEG TH TH FI EKSTDN. E - OVVTTYEG TH TH FI EKSTDN. E - OVVTTYEG MHS NFI EKSTDN. E - OVVTTYEG MHS NFI EKSTDN. E - OVVTTYEG MHS NFI EKSTDN. E - OVVTTYEG WT HP I EKSTDN. E - SI VVTTYEG WT HP I EKSTDN. E - SI VVTTYEG WT HP I EKSTDN. E - SI VVTTYEG WT HP I EKSTDN. E - OVVTTYEG WT HP I EKSTDN. E - OVVTTYEG WT HP I EKSTDN. E - OVVTTYEG WT HP I EKTTDN. E - OVVTTYEG WT HP I EKTDN. E - OVVTTYEG WT HP I EKTDN. E - OVVTTYEG WT HP I E - OVVT HP I WT HP	FEHIL IEDILROMOTY FEHILROMOSYF FEHILGOMOSYF FEHILGOMOSYF FEHILGOMOSYF FEHILTOMOIYS FEHILTOMOIYS FEHILTOMOIYS FEHILTOMOIYS FEHILTOMOIYS FEHILTOMOIYS FEHILSOMOIY
PtWRKY-IIC3a VVWRKY-IIC3a GWRKY-IIC3a PtWRKY-IIC3a ZmWRKY-IIC3a BdWRKY-IIC3a BdWRKY-IIC3a PtWRKY-IIC3a PtWRKY-IIC3a VVWRKY-IIC3a GmWRKY-IIC3a GmWRKY-IIC3a	TRS	NEF PRSYYRCT NNGCNYKKQVGNXK NKF PRSYYRCT YKDCNYKKQVGNXK NKF PRSYYRCT YKDCNYKKQVGLSKD NKF PRSYYRCT HQCNYKKQVGLSKD NKF PRSYYRCT HQCGNYKKQVGLSKD NKF PRSYYRCT HQCGNYKKQVGLSKD NKF - PRSYYRCT HQCGNYKKQVGLSKD NKF - PRSYYRCT YQCGNYKKQVGLSKD NKF - PRSYYRCT YQCGNYKKQVGLSKD NKF - PRSYYRCT YQCGNYKKQVGLSKD NKF - PRSYYRCT HQCNYKKQVGLSKD NKF - PRSYYRCT HQCNYKKQVGLSKD NKF - PRSYYRCT HQCNYKKQVGLSKD NKF - PRSYYRCT HQCNYKKQVGLSKD NKF - PRSYYRCT HQCNYKKQVGLSKD	E - E I VVTTYEG MHT HP I EKCI DN E - E I VVTTYEG I HTHPVEK FITEN E - CIVVTTYEG I HTHPVEK FITEN E - OVVTTYEG MHS NG I GNST DN E - OVVTTYEG MTHPI EKSN DN E - OVVTTYEG THTHPI EKSN DN E - OVVTTYEG MTHPI EKSN DN E - OVVTTYEG MTHPI EKSN DN E - OVVTTYEG MSNH I EKSN DN E - OVVTTYEG MSNH I EKST DN E - OVVTTYEG MSHPI EKST DN E - OVVTTYEG VHTHPI EKST DN E - OVVTTYEG HTHPI CAN ON LED	FEHIL IEDILROMOTY FEHILROMOSYF FEHILOMOSYF
PtWRKY-IIG3a VVWRKY-IIG3a GWWRKY-IIG3a ZmWRKY-IIG3a ZmWRKY-IIG3a BdWRKY-IIG3a SbWRKY-IIG3a PtWRKY-IIG3a VWWRKY-IIG3a VWWRKY-IIG3a GmWRKY-IIG3a GmWRKY-IIG3a	TRS GVD I LDDGYRWRKYGGKTVK TRS HVD I LDDGYRWRKYGGKTVK TRS RVD I LDDGYRWRKYGGKAVK TRS GVD I LDDGYRWRKYGGKAVK	NKF PRSYYRCT NNGCNYKKQYGR NSK NKF PRSYYRCT YKDONYKKQYGR LSKD NKF PRSYYRCT YKDONYKKQYGR LSKD NKF PRSYYRCT HOGCNYKKQYGR LSKD NKF PRSYYRCT HOGCNYKKQYGR LSKD NKF PRSYYRCT HOGCNYKKQYGR LSKD NKF PRSYYRCT YGCNYKKQYGR LSKD NKF PRSYYRCT HOGCNYKKQYGR LTKD NKF PRSYYRCT YKGCNYKKQYGR LTKD	E - E I VVTTYEG MHT H P I EKCT DN. E - E I VVTTYEG I HTH VKK T TN. E - G VVVTTYEG I HTH VKK T TN. E - G VVVTTYEG MTSHQ I EKST DN. E - G VVVTTYEG MTSH I EKST DN. E - G VVVTTYEG MTH P I EKST DN. E - G VVVTTYEG MTH P I EKST DN. E - G VVVTTYEG MTH P I EKST DN. E - G VVVTTYEG MTSH I EKST DN. E - G VVTTYEG MTSH I EKST DN	FEHIL IEDILROMOTY FEHILROMOSYFEI FEHILSOMOIYS FEHILSOMOIYS FEHILSOMOIYS FEHILTOMOIYSG FEHILTOMOIYSG FEHILTOMOIYSG FEHILTOMOIYSG FEHILTOMOIYSG FEHILTOMOIYSG FEHILSOMOIY FEHILSOMNIY
PtWRKY-11C3a VYWRKY-11C3a CgWRKY-11C3a PtWRKY-11C3a BdWRKY-11C3a BdWRKY-11C3a PtWRKY-11C3a PtWRKY-11C3a VYWRKY-11C3a VYWRKY-11C3a GmWRKY-11C3a CgWRKY-11C3a CgWRKY-11C3a	T RS	ISK PREYYRC] NNGCNYKKQYGN SK NKK - PREYYRC] YKOCNYKKQYGN LSKD NKK - PREYYRC] YKOCNYKKQYGL LSKD NKK - PREYYRC] HAGCNYKKQYGL LSKD NNF - PREYYRC] HAGCNYKKQYGL LSKD NNF - PREYYRC] HAGCNYKKQYGL LSKD NKF - PREYYRC] HAGCNYKKQYGL LSKD NKF - PREYYRC] YGCNYKKQYGL LSKD NKF - PREYYRC] HAGCNYKKQYGL LSKD NKF - PREYYRC] YKCCNYKKQYGL LSKD NKF - PREYYRC] YKCCNYKKQYGL LSKD NKF - PREYYRC] YKCCNYKKQYGL LSKD	E - E IVVTTYEG MHT HP I EKCTDN E - E IVVTTYEG HT HP I EKCTDN E - E IVVTTYEG HT HP VEKTTEN E - OVVTTYEG HT HP I EKSTDN E - OVVTTYEG HT HP I EKTDN E - OVVTYEG	FEHIL I ED I L ROMOTY FEH I L GAMOSY F FI FEH I L GAMOSY F FI FEH I L GAMOSY F S FEH I L GAMOSY S FIN J S F
PtwRKY-IIG3a VYWRKY-IIG3a GWRKY-IIG3a SWRKY-IIG3a SWRKY-IIG3a SaWRKY-IIG3a OsWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a GWKKY-IIG3a GWKKY-IIG3a GWKKY-IIG3b	T R S	NEF PRSYYRCT NNGCNYKKQYGR NSK NKF PRSYYRCT YKDCNYKKQYGR LSKD NKF PRSYYRCT YKDCNYKKQYGR LSKD NKF PRSYYRCT HQCNYKKQYGL LSKD	E - E I VVTTYEG MHT NP I EKCTDN E - E I VVTTYEG I HTNP VEKTTEN E - C I VVTTYEG I HTNP VEKTTEN E - C VVTTYEG I HTNP I EKSTDN E - C VVVTTYEG THTP I EKSTDN E - C VVVTTYEG THTP I EKSTDN E - C VVTTYEG THTP I EKSTDN E - C VVTTYEG THTP I EKSTDN E - C VVTTYEG I HTNP I EKSTDN E - C VVTTYEG VIT I EKSTDN E - C VVTTYEG I HTNP I EKTDN E - C VVTTYEG I HTNP I EKTDN E - C VVTTYEG I HTNP I CKTDN E - CVVTTYEG I H	FEHIL I ED I L ROMOTY FEHI L ROMOSY F FI FEHI L TOMOI Y 5 - FEHI L SOMOI Y -
PtWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a PtWRKY-IIC3a BdWRKY-IIC3a BdWRKY-IIC3a BdWRKY-IIC3a PtWRKY-IIC3a PtWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a	TRS GVD I LDDGYRWRRYGGRYVR TRS HVD I LDDGYRWRRYGGRAVK TRS RVD I LDDGYRWRRYGGRAVK TRS RVD I LDDGYRWRRYGGRAVK TRS GVD I LDDGYRWRRYGGRAVK TRS	IN KE PRSYYRCT NNGCNYKKQYQR NSKQ NKF PRSYYRCT YKDGNYKKQYQR LSKD NKF PRSYYRCT YKDGNYKKQYQR LSKD NKF PRSYYRCT HQGCNYKKQYQR LSKD NKF PRSYYRCT HQGCNYKKQYQR LSKD NKF PRSYYRCT HQGCNYKKQYQR LSKD NKF PRSYYRCT YQGCNYKKQYQR LSKD NKF PRSYYRCT YQGCNYKKQYQR LSKD NKF PRSYYRCT HQGCNYKKQYQR LSKD NKF PRSYYRCT HQGCNYKKQYQR LSKD NKF PRSYYRCT HQGCNYKKQYQR LSKD NKF - PRSYYRCT HQGCNYKKQYQR LSKD	E - E I VVTTYEG MHT H P I EKCTON. E - E I VVTTYEG I HTH VEK F TEN. E - E I VVTTYEG I HTH VEK F TEN. E - E VVVTTYEG MHS NG I GNSTON. E - E VVVTTYEG MTS NG I EKSTON. E - E VVVTTYEG THTH I EKSNON. E - E VVVTTYEG THTH I EKSNON. E - E VVVTTYEG MTS H I EKSTON. E - O VVTTYEG H H I EKSTON. E - O VVTTYEG MTS H I EKSTON. E - O VVTTYEG MTS H I EKSTON. E - O VVTTYEG H H H I I O VI I I I I I I I I I I I I I I I	FEHIL IEDILRAMATY FEHILRAMASY FEHILGAMASY
PtWRIKY-IIG3a VYWRIKY-IIG3a GWWRIXY-IIG3a PtWRIXY-IIG3a BdWRIXY-IIG3a BdWRIXY-IIG3a DowWRIX-IIG3a PtWRIXY-IIG3a VYWRIXY-IIG3a GWWRIXY-IIG3a GWWRIXY-IIG3a GWWRIXY-IIG3a GWWRIXY-IIG3a GWWRIXY-IIG3a BdWWRIXY-IIG3a AtWRIXY-IIG3a AtWRIXY-IIG3a	T RS BVD I L DDGY RWRK YGGKT VKN T RS HVD I L DDGY RWRK YGGKAVKN T RS RVD I L DDGY RWRK YGGKAVKN T RS RVD I L DDGY RWRK YGGKAVKN T RS QVD I L DDGY RWRK YGGKAVKN T RS	IN KE - PREYYRCTNNGCNYKKQVGN ISKO NKK - PREYYRCTYKOCNYKKQVGN ISKO NKK - PREYYRCTYKOCNYKKQVGN ISKO NKK - PREYYRCTYACCNYKKQVGN ISKO NNF - PREYYRCTYACCNYKKQVGN ISKO NNF - PREYYRCTYACCNYKKQVGN ISKO NKF - PREYYRCTYACCNYKKQVGN ISKO	E E IVVITTYEG HT HP IEKCT DN E E VVVTTYEG HT HP IEKCT DN E C IVVTTYEG HT HP IEKST DN E C VVVTTYEG HT HP IEKST DN E C VVTTYEG HT HP IEKST DN E C VVTTYEG HT HP IEKST DN E G IVVTTYEG HT HP IEKST DN E G IVVTTYEG HT HP IEKT DN E G IVVTTYEG HT HP IEKT DN E G IVVTTYEG HT HP IEKT D	FEHIL I E D I L ROMOTY FEH I L GAMASY F H FEH I L
PtWRKY-IIG3a VYWRKY-IIG3a GWRKY-IIG3a PtWRKY-IIG3a BWWRKY-IIG3a DWWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a GWWRKY-IIG3a GWWRKY-IIG3a GWWRKY-IIG3a AtWRKY-IIG3a BWWRKY-IIG3a BWWRKY-IIG3a	TRS GVD I LDDGY RWRKYGOKTVKN TRS HVD I LDDGY RWRKYGOKAVKN TRS RVD I LDDGY RWRKYGOKAVKN TRS RVD I LDDGY RWRKYGOKAVKN TRS GVD I DDGY RWRKYGOKAVKN TRS DDDV DDGY RWRKYGOKAVKN TRS DDV DV DDGY RWRKYGOKAVKN TRS DDV DV DDGY RWRKYGOKAVKN TRS DDV DV DDY RWRYGWRKYGOKAVKN TRS DDV TDY DY DY RWRYKYG	INK PRSYYRCT NNGCNYKKQVGRASK NKY PRSYYRCT YKDCNYKKQVGRASK NKY PRSYYRCT YKDCNYKKQVGRASK NKF PRSYYRCT HQCNYKKQVGLSKD NKF PRSYYRCT HQCNYKKQVGLSKD NKF PRSYYRCT HQCNYKKQVGLSKD NKF PRSYYRCT HQCNYKKQVGLSKD NKF PRSYYRCT YQCGTYKKQVGLLSKD NKF PRSYYRCT YQCGTYKKQVGLLSKD NKF PRSYYRCT YQCGTYKKQVGLLSKD NKF PRSYYRCT YQCGTYKKQVGLLSKD NKF PRSYYRCT YQCGNYKKQVGLLSKD NKF PRSYYRCT YQCGNYKKQVGLLSKD NKF PRSYYRCT YQCGNYKKQVGLLSKD NKF PRSYYRCT YQCGNYKKQVGLLSKD NKF PRSYYRCT YYCGNYKKQVGLLSKD NKF PRSYYRCT YYCGNYKKQVGLLSKD NKF PRSYYRCT YYCGNYKKQVGLLSKD NKF PRSYYRCT YYCGNYKKQVGLLSKD NKF PRSYYRCT YYCGNYKKQVGLLSKD NKF PRSYYRCT YYCGNYKKQVGLLSKD NKF PRSYYRCT YYCGNYKKQVGLLAKD NSF PRSYYRCT YYCGNYKKQVGLLAKD NSF PRSYYRCT YYCGNYKKQVGLLAKD NSF PRSYYRCT YYCGNYKKQVGLAKD	E - E I VVTTYEG MHT HP I EKCTDN	FEHIL IEDILROMOTY FEHILROMOSYF FEHILGOMOSYF
PtWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a PtWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a PtWRKY-IIC3a PtWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a CyWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a DaWRKY-IIC3a	TRS GVD I LDDGY RWRKYGGKTVKN TRS HVD I LDDGY RWRKYGGKAVKN TRS RVD I LDDGY RWRKYGGKAVKN TRS RVD I LDDGY RWRKYGGKAVKN TRS GVD I LDDGY RWRKYGGKAVKN	NEK F PRSYYRCT NNGCNYKKQYGR NSKO NKF PRSYYRCT YKDGNYKKQYGR LSKD NKF PRSYYRCT YKDGNYKKQYGL LSKD NKF PRSYYRCT HQGCNYKKQYGL LSKD NKF PRSYYRCT YKGCNYKKQYGL LSKD NKF PRSYYRCT YHT CNYKKQYGL LSKD NKF PRSYYRCT YHT CNYKKQYGL LSKD NKF PRSYYRCT YHT CNYKKQYGL LAKD NKF PRSYYRCT YHT CNYKKQYGL LAKD SAF - PRSYYRCT HHT CUYKKQYGL LAKD SAF - PRSYYRCT HHT CUYKKQYGL LAKD	F. F. E. V.V.T.T.YEG.MHT.H.P. IEK.C.I.D.N. E. E. I.V.V.T.Y.YEG.H.T.H.P. IEK.S.T.D. E. G. V.V.T.T.YEG.MHT.H.P. IEK.S.T.D. E. O.V.V.T.T.YEG.T.H.P. IEK.S.D. E. O.V.V.T.T.YEG.M.S.H.F.K.S.N.D. E. O.V.V.T.T.YEG.W.H.H.P.IEK.S.T.D. E. O.V.V.T.T.YEG.W.H.H.P.IEK.T.D. E. O.V.V.T.T.YEG.W.H.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.V.H.N.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.V.H.N.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.W.H.N.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.W.H.N.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.W.H.N.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.W.H.N.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.W.H.N.H.P.C.K.L.M.E.T. P. N.V.V.T.T.YEG.W.H.N.H.P.C.K.L.M.E.T. P.N.V.V.T.T.YEG.W.H.H.H.E.K.T. P.N.V.V.T.Y.EG.W.H.H.H.E.K.T.	FEHIL IEDILROMOTY FEHILROMOSYF FEHILGOMOSYF
PtWRKY-IIC3a VYWRKY-IIC3a GWWRY-IIC3a PtWRKY-IIC3a BWWRKY-IIC3a BWWRKY-IIC3a PtWRKY-IIC3a PtWRKY-IIC3a VYWRKY-IIC3a GWWRKY-IIC3a GWWRKY-IIC3a GWWRKY-IIC3a BWWRKY-IIC4a AtWRKY-IIC4a BdWWRKY-IIC4a BdWWRKY-IIC4a BdWWRKY-IIC4a BdWWRKY-IIC4a	T RS BUD I L DDGY RWRK YGBK TVKN T RS HVD I L DDGY RWRK YGBK TVKN T RS RVD I L DDGY RWRK YGBK AVKN T RS RVD I R RVRK YGBK AVKN T RS RVD I R RVRK YGBK AVKN T RS RVD I R RVR RVR RVR RVRK YGBK AVKN T RS RVD I R RVRK RVRK RVRK YG	NKF PREYYRC] NNGCNYKKQVGN LSKD NKF PREYYRC] YNGCNYKKQVGN LSKD NKF PREYYRC] YNGCNYKKQVGL LSKD NNF PREYYRC] HAGCNYKKQVGL LSKD NNF PREYYRC] HAGCNYKKQVGL LSKD NNF PREYYRC] HAGCNYKKQVGL LSKD NKF PREYYRC] YNC MAKKQVGL LSKD NKF - PREYYRC] YNC MAKKQVGL LAKD NSF - PREYYRC] YNC MAKKQVGL LAKD NSF - PREYYRC] YNC MAKKQVGL LAKD NSF - PREYYRC] HHT CUYKKQVGL LAKD NSF - PREYYRC] HHT CUYKKQVGL LAKD NSF - PREYYRC] HHT CUYKKQVGL LAKD	F. F. UVUTTYEG MHT HP I EKCTON	FEHIL IEDILROMOTY FEHILGOMOSYF
PtWRKY-11G3a GWWRY-11G3a GWWRY-11G3a SWWRY-11G3a SWWRY-11G3a BWWRY-11G3a PtWRY-11G3a PtWRY-11G3a PtWRY-11G3a PtWRY-11G3a GWWRY-11G3a GWWRY-11G3a GWWRY-11G3a BWWRY-11G3a SWWRY-11G3a BWWRY-11G3a SWWRY-11G3a BWWRY-11G3a	TRS	NER PRSYYRCT NNGCNYKKQVGRASK NKK PRSYYRCT YKDCNVKKQVGRASK NKK PRSYYRCT YKDCNVKKQVGRASK NKF PRSYYRCT HQCNVKKQVGLASK NNF PRSYYRCT HQCNVKKQVGLSKD NKF PRSYYRCT HQCNVKKQVGLSKD NKF PRSYYRCT HQCNVKKQVGLSKD NKF PRSYYRCT YQCGTVKKQVGLLSKD NKF PRSYYRCT YCGCUVKKQVGLLSKD NKF PRSYYRCT YCGCUVKKQVGLLSKD NKF PRSYYRCT YHT CNVKKQVGLLSKD NKF PRSYYRCT YHT CNVKKQVGLLSKD NKF PRSYYRCT YHT CNVKKQVGLLAKD NAF PRSYYRCT HHT CNVKKQVGLLAKD NAF PRSYYRCT HHT CNVKKQVGLLAKD NAF PRSYYRCT HHT CNVKKQVGLLAKD	F. F. F. VUTTYEG MHT HP I EKCTON	FEHIL IEDILROMOTY FEHILGOMOSYF
PtWRKY-IIG3a CyWRKY-IIG3a CyWRKY-IIG3a PtWRKY-IIG3a BdWRKY-IIG3a BdWRKY-IIG3a BdWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a CyWRKY-IIG3a CyWRKY-IIG3a CyWRKY-IIG3a CyWRKY-IIG3a BdWRKY-IIG3a BdWRKY-IIG3a BdWRKY-IIG3a BdWRKY-IIG4a BdWRKY-IIG4a BdWRKY-IIG4a BdWRKY-IIG4a CyWRKY-IIG4a CYWRKY-IIG4a CYWRKY-IIG4a CYWRKY-IIG4a CYWRKY-IIG4A CYWRK	TRS GVD I LDDGY RWRKYGGKTVKN TRS HVD I LDDGY RWRKYGGKAVKN TRS RVD I LDDGY RWRKYGGKAVKN TRS RVD I LDDGY RWRKYGGKAVKN TRS GVD I LDDGY RWRKYGGKAVKN TRS GUD I LDDGY RWRKYGGKAVKN TRS DDDVLDDGY RWRKYGGKAVKN TRS DDDVLDDGY RWRKYGGKAVKN TRS DDDVLDDGY RWRKYGGKAVKN TRS DDDVLDDGY RWRKYGGKAVKN TRS DDDVLDDGY RWRKYGGKAVKN TRS DDDVLDDGY RWRKYGGKAVKN TRS DDDULDDGY RWRKYGGKAVKN TRS DDDULDDGY RWRKYGGKAVKN TRS DDDULDDGY RWRKYGGKAVKN TRS DDULDDGY RWRKYGGKAVKN TRS DDULDGY RWRKYGGKAVKN TRS DDULDDGY RWRKYGGKAVKN TRS DDULDGY RWRKYGGKAVKN TRS DDULDDGY RWRKYGGKAVKN TRS DDULDDY RWRKYG	INKE PRSYYRCT NNGCNYKKQYGR NSKO NKE PRSYYRCT YKDGNYKKQYGR LSKD NKE PRSYYRCT YKDGNYKKQYGR LSKD NKE PRSYYRCT HQGCNYKKQYGR LSKD NKE - PRSYYRCT HQGCNYKKQYGR LSKD NKE - PRSYYRCT HQGCNYKKQYGR LSKD NKE - PRSYYRCT HYT GNYKKQYGR LAKD NKE - PRSYYRCT HYT GNYKKQYGR LAKD SAE - PRSYYRCT HYT GNYKKQYGL LAKD	F F F VVVTTYEG HTHP F <td< td=""><td>FEHIL IEDILROMOTY FEHILGAMOTY FEHILGAMOTY<!--</td--></td></td<>	FEHIL IEDILROMOTY FEHILGAMOTY FEHILGAMOTY </td
PtWRKY-IIG3a VYWRKY-IIG3a GWWRY-IIG3a PtWRKY-IIG3a BWWRKY-IIG3a BWWRKY-IIG3a OSWRKY-IIG3a VYWRKY-IIG3a VYWRKY-IIG3a GWWRKY-IIG3a GWWRKY-IIG3a GWWRKY-IIG3a GWWRKY-IIG3a BWWRKY-IIG3a BWWRKY-IIG3a BWWRKY-IIG3a BWWRKY-IIG3a BWWRKY-IIG4a	T RS BVD I L DDGY RWRK YGBKT VKN T RS HVD I L DDGY RWRK YGBKT VKN T RS RVD I L DDGY RWRK YGBKA VKN T RS RVD I L RVN RWRK YGBKA VKN T RS RVN RVN RVN RWRK YGBKA VKN T RS RV	INKE - PREYYRCINNGCNYKKQVGR LSKD NKK - PREYYRCIYNGCNYKKQVGR LSKD NKK - PREYYRCIYNGCNYKKQVGR LSKD NKF - PREYYRCIYNGCNYKKQVGL LSKD NNF - PREYYRCIYNGCNYKQVGL LSKD NNF - PREYYRCIYNGCNYKQVGL LSKD NKF - PREYYRCINHT NNYKQVGL LSKD	F. E. IVVTTYEG MHT HP IEKCTON	FEHIL IEDILROMOTY FEHILGOMOSYF LSOLGOFIG LSPLLROLGFLSS LSPLIKOLGFLSS <
PtWRKY-IIG3a VYWRKY-IIG3a GWRKY-IIG3a SWRKY-IIG3a DeWRKY-IIG3a DeWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a GWWRKY-IIG3a GWWRKY-IIG3a GWWRKY-IIG3a BdWRKY-IIG3a ChWRKY-IIG3a ChWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3a CWWRKY-IIG3A CWWRKY-	TRS	INK PRSYYRCT NNGCNYKKQVGR SK NKY PRSYYRCT YKDCNYKKQVGR SK NKY PRSYYRCT YKDCNYKKQVGR SK NKF PRSYYRCT HQCNYKKQVGR SK NKF PRSYYRCT YQCNYKKQVGR SK NKF PRSYYRCT YQCNYKKQVGR SK NKF - PRSYYRCT YCCNYKKQVGR SK NKF - PRSYYRCT YCCNYKKQVGR SK NKF - PRSYYRCT YCCNYKKQVGR SK NKF - PRSYYRCT YCCNYKKQVGR SK NKF - PRSYYRCT YKCCNYKKQVGR SK SK - PRSYYRCT YKCNYKCNYKCN SK SK - PRSYYRCT YKCNYKCNYKCN SK SK - PRSYYRCT YKCNYKCNYKCN SK SK - PRSYYRCT YKCNYKCNYKCNN SK SK - PRSYYRCT YKCNYKCNN SK SK - PRSYYRCN SK SK -	F. E. IVUTTYEG MHT NP IEKCTDN	FEHIL IEDILROMOTY FEHILGOMOSYF
PtWRKY-IIG3a CyWRKY-IIG3a CyWRKY-IIG3a PtWRKY-IIG3a DawRKY-IIG3a DawRKY-IIG3a DawRKY-IIG3a DawRKY-IIG3a PtWRKY-IIG3a PtWRKY-IIG3a CyWRKY-IIG3a CyWRKY-IIG3a CyWRKY-IIG3a DawRKY-IIG3A DawRKY-IIG3A DAWRK	TRS BUD I L DDG Y RWRK YGB K TV KN RS RUD I L DDG Y RWRK YGB K AV KN TRS	NEK F PRSYYRCT NNGCNYKKQYGR NSK NKK PRSYYRCT YKDGNYKKQYGR NSK NKK PRSYYRCT YKDGNYKKQYGL LSK NKF PRSYYRCT HQGCNYKKQYGL LSK NKF PRSYYRCT HQGCNYKKQYGL LSK NKF PRSYYRCT HQGCNYKKQYGL LSK NKF PRSYYRCT HQGCNYKKQYGL LSK NKF - PRSYYRCT HHT CNYKKQYGL LSK NH - PRSYYRCT HHT CNYKCY NG LSK NH - PRSYYRCT HHT CNYKKQYGL LSK NH - PRSYY	F. E. IVUTTYEG MHT HP IEKCTON. E. UVUTTYEG HT HP VEKPTEN. E. UVUTTYEG HT HP IEKCTON. E. OVUTTYEG HT HP IEKSTON. E. OVUTTYEG TH THP IEKSTON. E. OVUTTYEG TH THP IEKSTON. E. OVUTTYEG MS HP IEKSTON. E. OVUTTYEG WITH PIEKTON. F. OVUTYEG WITH PIEKTON. F. SIVUTYEG H	FEHIL IEDILROMOTY FEHILGMOSYFI FEHILGMOSYFI FEHILGMOSYFI FEHILGMOSYFFI FEHILGMOSYFFI FEHILGMOSYFFI FEHILGMOSYFFI FEHILGMOSYFFI FEHILGMOSYFFI FEHILGMOSYFFI FEHILGMOSYFFFI FEHILGMOSFFFI FEHILGMOSFFFFI FEHILGMOSFFFI
PtWRIKY-IIC3a CyWRIKY-IIC3a CyWRIKY-IIC3a PtWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a PtWRIKY-IIC3a PtWRIKY-IIC3a CyWRIKY-IIC3a CyWRIKY-IIC3a CyWRIKY-IIC3a CyWRIKY-IIC3a CyWRIKY-IIC3a CyWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a BdWRIKY-IIC3a CyWRIKY-IIC3a CYWRIKY-IIC3A CY	T RS BUD I L DDGY RWRK YGB KTV KN T RS HVD I L DDGY RWRK YGB KTV KN T RS HVD I L DGGY RWRK YGB KAVKN T RS RVD I L DGY RWRK YGB KAVKN T RS RVD I L DGY RWRK YGB KAVKN T RS RVD I R RVRK YGB KAVKN T RS RVD I R RV	IN KE - PREYYRCINNGCNYKKQVGN ISK NKK - PREYYRCIYNGCNYKKQVGN ISK NKK - PREYYRCIYNGCNYKKQVGLISK NKK - PREYYRCIYNGCNYKKQVGLISK NNF - PREYYRCIYNGCNYKKQVGLISK NNF - PREYYRCIYNGCNYKQVGLISK NKF - PREYYRCIYNCINNKCUNGLISK NKF - PREYYRCIYNCINNKCUNGLISK NNM - PREYYRCIYNCINNKCUNGLISK NNM - PREYYRCIYNCINNKCUNGLISK NNM - PREYYRCINNCUNKCUNGLISK NNM - PREYYRCINNCUNKCUNGLISK	F. E. IVVTTYEG MHT HP IEKCTON	FEHIL
PtWRKY-11G3a CyWRKY-11G3a CyWRKY-11G3a PtWRKY-11G3a SbWRKY-11G3a DowrkY-11G3a PtWRKY-11G3a PtWRKY-11G3a PtWRKY-11G3a CyWRKY-11G3a CyWRKY-11G3a CyWRKY-11G3a CyWRKY-11G3a BdWRKY-11G3a BdWRKY-11G3a BdWRKY-11G3a BdWRKY-11G3a BdWRKY-11G3a BdWRKY-11G3a BdWRKY-11G3a CyWRKY-11G3a CYWRKY-11G3A CYWRK	T R S B VD I L DDGY RWRK YGBK TVKN T R S HVD I L DDGY RWRK YGBK TVKN T R S RVD I L DDGY RWRK YGBK AVKN T R S RVD I L	INKE - PRSYYRCT NHC CNYKKQ VG NSK NKK - PRSYYRCT YKO CNYKKQ VG NSK NKF - PRSYYRCT YKO CNYKKQ VG NSK NKF - PRSYYRCT YKO CNYKKQ VG NSK NKF - PRSYYRCT YGO CNYKKQ VG NSK NKF - PRSYYRCT YNG CNYKKQ VG NSK NY - PRSYYRCT YNG CNYKKQ VG NSK NY - PRSYYRCT HNT CNYKQ NSK NY - PRSYYRCT HNT CNYKQ NSK NY - PRSYYRCT NSK NY - PRSYYRCT NSK NY - PRSYNCT	F. E. IVUTTYEG MHT HP IEKCIDN	FEHIL IEDILROMOTY FEHILGOMOSY FEHILGOMOSY </td

IIC1-

IIC2-

IIC3-

IIC4-

			W	VRKY DNA-bind	ing domain							
					-							
	•							••				
	360	370	380	390	400	410	4	20	430	440	450	460
AtWRKY-IIC5a	TKS EKVALDDO	YKWRKY GKK	ITGSPF	PRHYHKCSSPI	D CN VKKK I E R		ILTTYEGR	HNHPSPSV	/výcp			
OsWRKY-IIC5a	TRSDD-EILDDG	3 YKWRKYGK <mark>K</mark>	SV <mark>KNSP</mark> N	PRNYYRCSTE	GCNV <mark>KKR</mark> VE <mark>r</mark>	DKN <mark>DP</mark> - R <mark>y</mark>	VV <mark>T</mark> MY <mark>EG</mark> I	HNHV - CPC	т v			
BdWRKY-IIC5a	TKSK · · · E · EVLDDG	3 Y KWRKYGKK	SV <mark>KNSP</mark> N	PRNYYRCSTE	G C N V KKR V E R	DR <mark>DDA</mark> - D <mark>y</mark>	V L <mark>T</mark> M <mark>Y E G</mark> I	HNHA - S <mark>P</mark> C	т v			
SbWRKY-IIC5a	V <mark>RS</mark> D DEEVLDDG	3 <mark>Y KWRK Y</mark> GK <mark>K</mark> S	SV <mark>KNSP</mark> N	PRNYYRCSTE	GCSV <mark>KKR</mark> VE <mark>R</mark>	DK <mark>DD</mark> Q - R <mark>Y</mark>	VV <mark>T</mark> MY <mark>EG</mark> V	HNHV - S <mark>P</mark> (т I			
ZmWRKY-HC5a	V <mark>RS</mark> D EDEVLDDG	3 <mark>Y KWRK Y</mark> GK <mark>K</mark> S	S V <mark>K N S P</mark> N	PRNYYRCSTE	GCNV <mark>KKR</mark> VER	DR <mark>DDP</mark> - R <mark>Y</mark>	VV <mark>TMYEG</mark> V	'HNHV - S <mark>P</mark> C	т v			
BdWRKY-IIC5b	TRSE · · · I · EILDDO	3 <mark>Y KWRK Y</mark> GK <mark>K</mark> S	SV <mark>KNSP</mark> N	PRNYYRCS TE	GCSV <mark>KKR</mark> VE <mark>R</mark>	DR <mark>DDP</mark> - S <mark>Y</mark>	VV <mark>TTY</mark> EGT	HSHV - S <mark>P</mark> S	:тv			
OsWRKY-IIC5b	TKSE···I·EILDDG	YKWRKY GKKS	SVKNSPN	PRNYYRCSTE	GCNVKKRVER	DKDDP - SY	VVTTYEGT	HNHV - SPS	STV			
ZmWRKY-IIC5b	TRSESSION	SYKWRKYGKKS	SVKNSPN	PRNYYRCSTE	GCNVKKRVER	DRDDP - SY	VVTTYEGM	HNHV - SPS				
SOWRK T-IICSD	TREE EVELLED			PRNYTRUSTE			VVTTYEGI	HNHV-SP				
OsWRKY-NC5c	TRSEVELLODG	SEKWEKYGKK		PRNYVRCSTE	GONVERNER	DREDH.RY	VITTYDGV	HNHA. SP				
ShWRKY-IIC5c		FKWRKYGKK		PRNYYRCSSE	GCGVKKRVER	DRDDP - RY	VITTYDGV	HNHA - SPO				
ZmWRKY-IIC5c	TRS ····EVEILDDO		AVKNSPN	PRNYYRCSSE	GCGVKKRVER	DRDDP - RY	VITTYDGV		AA			
BdWRKY-IIC5d	TRS · · · · AVEVMEDO	G F RWRKYGKK	AVKSSPN	LRNYYRCSAP	GCGVKKRVER	DRHDP - AY	VITTYHGV		AS			
OsWRKY-IIC5d	TRS EVEVLDDG	3 F <mark>KWRKY</mark> GK <mark>K</mark> A	AV <mark>K</mark> S <mark>SP</mark> N	PRNYYRCSAA	GCG V <mark>KKR</mark> VE R	DGDDP - RY	V V T T Y DG V	HNHA - TPO	cv			
ZmWRKY-IIC5d	TRS EVDVLDDG	3 F <mark>KWRK Y</mark> GK <mark>K</mark> A	AV <mark>K</mark> S <mark>SP</mark> N	PRNYYRCSAE	GCGV <mark>KKR</mark> VE <mark>r</mark>	DS <mark>DDP</mark> - R <mark>Y</mark>	V V <mark>T T Y</mark> D G V	HNHA - APO	AA			
SbWRKY-IIC5d	T <mark>RS</mark> EVDVLDDG	3 F <mark>KWRKY</mark> GK <mark>K</mark> A	AV <mark>K</mark> S <mark>SP</mark> N	PRNYYRC <mark>S</mark> SE	GCG <mark>V</mark> KKRVE <mark>r</mark>	DS <mark>DDP</mark> - R <mark>Y</mark>	V I <mark>T T Y</mark> DG V	HNHA - APC	AA			
GmWRKY-IIC5a	T <mark>KS</mark> · · · · EVEILDDG	3 F <mark>KWRK Y</mark> GK <mark>K</mark> M	MV <mark>KNSP</mark> N	P R N Y Y R C S V D	GCQV <mark>KKR</mark> VER	DK <mark>DDP</mark> - R <mark>Y</mark>	V I <mark>T T Y</mark> EG I	HNHQ - SYI				
CpWRKY-IIC5a	TKS · · · · EVEILDDO	3 F <mark>KWRK Y</mark> GK <mark>K</mark> M	MV <mark>KNSP</mark> N··	PRNYYKCSIE	G C P V KKR VE R	DK <mark>EDP</mark> - S <mark>Y</mark>	I I TTYEGF	HNHR · TTO	;			
VvWRKY-IIC5a	TKS · · · · EIEILDDG	3 F <mark>KWRKY</mark> GK <mark>K</mark> I	MV <mark>KNSP</mark> N	PRNYYRCSVD	G C N V K K R V E R	DREDP - KY	VI <mark>TTY</mark> EGI	HNHE - S <mark>P</mark> S	KF			
PtWRKY-IIC5a	TKS	3 YKWRKYGKKI	MVKNSPN	PRNYYRCSVE	GCPVKKRVER	DRDDP - RY	VITTYEGI	HTH				
GmWRKY-IIC5b		YRWRKYGKK	WVKNSPN	PRNYYRCSVD	GCNVKKRVER	DKDDP - RY	VITTYEGN	HTHP - 555				
GINWRK Y-IICSC		S T RWRK YGKK		PRNNYRCSVD	GCIVKKRVER	DRDDP - RY	VITTYEGN	HIH <mark>F</mark> -ISS				
ShWRKT-HLSH		S PWPKYCKKI		PPNYYPCSSE	CCPVKKPVEP	EPDDA . PE	VITTYDGY		D D			
OsWRKY-UC5e		TRWRKYGKK		PRNYYRCSSE	GCRVKKRVFR	ARDDA - RE	VYTTYDGY	HNHP - AP	HILLE			
GmWRKY-IIC5d		YKWRKYGKK		PRNYYKCSGE	GCDVKKRVER	DRDDS - NY	VLTTYDGV		TA			
GmWRKY-IIC5e		YKWRKYGKK		PRNYYKCSGE	GCNVKKRVER	DRDDS - NY	VLTTYDGV	HNHE - SP	TA			
GmWRKY-IIC5f	TKS ELEIMDDO	3 YKWRKYGKK	SVKSSPN	LRNYYKCSSG	GCSV <mark>KKR</mark> VER	DR <mark>DDY</mark> - S <mark>Y</mark>	V I TTYEGV	HNHE - SP	тт			
GmWRKY-IIC5g	TKS ELEIMDDO	3 YKWRKYGKK	SVKSNPN	LRNYYKCSSG	GCS <mark>VKKR</mark> VE <mark>R</mark>	DRDD <mark>S</mark> -SY	V I <mark>TTY</mark> EGV	HNHE - S <mark>P</mark> I	тт			
CpWRKY-IIC5b	TKS · · · · EMEVMDDO	3 F <mark>KWRK Y</mark> GK <mark>K</mark> S	SV <mark>KNS</mark> PN	PRNYYKCSSR	GCHV <mark>KKR</mark> IE <mark>r</mark>	ERDDP - R <mark>y</mark>	V I <mark>T T Y</mark> EGT	HNHE - SP	vv ····			
AtWRKY-IIC5¢	T <mark>RS</mark> K I DVMDDG	3 F <mark>KWRK Y</mark> GK <mark>K</mark>	SV <mark>KN</mark> NIN	KRNYYKCSSE	GCSV <mark>KKR</mark> VE <mark>R</mark>	D <mark>GDDA</mark> - A <mark>Y</mark>	V I <mark>T T Y E G</mark> V	HNHE - SLS	NV			
VvWRKY-IIC5b	TKS DLE IMDDG	3 F <mark>KWRKY</mark> GK <mark>K</mark> S	SV <mark>KNS</mark> PN	PRNYYKCASG	GCNV <mark>KKR</mark> VE <mark>r</mark>	DR <mark>EDS</mark> - S <mark>Y</mark>	VI <mark>TTY</mark> EGV	HNHE - SPO	:vv			
VvWRKY-IIC5¢	TKS · · · · ELDVMDDC	3 F <mark>KWRKY</mark> GK <mark>K</mark> I	MVKS <mark>S</mark> PN	PRNYYRCS SG	DCQVKKR I E R	DIED <mark>S</mark> ·S <mark>Y</mark>	V I TTYTG I	HNHP - IPC	v <mark>g</mark>			
VvWRKY-IIC5d	TRS EEDIMDDO	G F <mark>K WRK</mark> Y G K <mark>K</mark> H	KIKSNPIY-	PRNYYRCS SR	GCQVKKRVER	DRDDS - SY	VITTYEGV	HNHP - TPI	RNH			

IIC5-



PpWRKY-IIC7a	TKEPMA·DTDVLDDGYKWRKYGQKAVKNSPH··PS <mark>SYYR</mark> CATPNCPVRKRVERCIEDP·GLVATAYEGTHSHQFPSFLRCPPGYP············GGLPGLLATLSHH
PpWRKY-IIC7b	TKQPIA-HTDVMDDGYKWRKYGQKPVKS <mark>SPHPRNYYRCTTPNCPVRKRVER</mark> STEDP-DQVITTYEGRHTHQSPSFLKGSPEYPSELSQLLGALYHH
PpWRKY-IIC7¢	TRT - · · · DVD IMDDGFKWRKYGQKAVKNSPYP · GRNYYRCT PQCPVRKRVERSCEDS · GLV I TTYEGTHT HQT PGLFSHPS · · · · · · · · · · · · · · · · · · ·

PtWRKY-IIC1a	HDD	
GmWRKY-IIC1a		
GmWRKY-IIC1b	5HS Y	
AtWRKY-IIC1a		
BdWRKY-IIC1a		
SbWRKY-IIC1a	<mark>Р</mark> КК <mark>Р</mark> КL <mark>0</mark>	
CpWRKY-IIC1a	TV	
OSWRKY-IIC18		
OsWRKY-IIC2a		
ShWRKY-IIC2a		
ZmWRKY-IIC2a*	• • • • • • • • • • • • • • • • • • • •	
MWRKY-IIC2a*		
GmWRKY-IIC2a	•	II.CO
GmWRKY-IIC2b	•	IIC2-
GmWRKY-IIC2c		
OSWRKY-IIC28*		
OsWRKY-IIC2c*		
VvWRKY-IIC2a*		
PtWRKY-IIC2a*		
AtWRKY-IIC28*		
PtWRKY-IIC3a		
VvWRKY-IIC3a	ş	
CpWRKY-IIC3a		
ShWRKY-IIC3a		
ZmWRKY-IIC3a		
BdWRKY-IIC3a		
DSWRKY-IIC32 PtWRKY-IIC3c		IIC3-
PtWRKY-IIC3d		_
VvWRKY-IIC3b		
ATWRKY-IIC3a VyWRKY-IIC3a		
GmWRKY-IIC3a		
GmWRKY-IIC3b		
CpWRKY-IIC3b		
-		
AtWRKY-IIC4a	SDEL <mark>GG</mark> S <mark>G</mark> NSNNNKKK <mark>G</mark> EKKVRK <mark>P</mark> RYAFQTRSQVDIL <mark>G</mark> VVVTTYE <mark>G</mark> VHTH <mark>P</mark> IEKTTDNFEH·····	
AtWRKY-IIC4b		
ShWRKY-IIC4a		
ZmWRKY-IIC4a		
BdWRKY-IIC4b GmWRKY-IIC4a		
AtWRKY-IIC4c		
GmWRKY-IIC4b		IIC4-
GmWRKY-IIC4c		
vvWRKY-IIC4a PtWRKY-IIC4a		
OsWRKY-IIC4a		
S&WRKY-IIC4&		
ZmWRKY-IIC4b		

	470	480	490	500	510	520	530	540	550	560	570	580
AtWRKY-IIC52			 									
OsWRKY-IIC5a			 									
BdWRKY-IIC5a			 									
SBWRKY-IIC5a			 									
ZmWRKY-IIC5a			 									
BdWRKY-IIC5b			 									
OsWRKY-IIC5b			 									
ZmWRKY-IIC5b			 									
ShWRKY-IIC5b			 									
BdWRKY-IIC5c			 									
OsWRKY-IIC5c			 									
ShWRKY-IIC5c			 									
ZmWRKY-IIC5¢			 									
BdWRKY-IIC5d			 									
OsWRKY-IIC5d			 									
ZmWRKY-IIC5d			 									
SbWRKY-IIC5d			 									
GmWRKY-IIC5a			 									
CpWRKY-IIC5a			 									
VVWRKY-IIC5a			 									
PtWRKY-IIC5a			 									
GmWRKY-IIC5b			 									
GmWRKY-IIC5c			 									
AtWRKY-IIC5b			 									
ShWRKY-IIC5e			 									
OsWRKY-IIC5e			 									
GmWRKY-IIC5d			 									
GmWRKY-IIC5e			 									
GmWRKY-IIC5f			 									
GmWRKY-IIC5g			 									
CpWRKY-IIC5b			 									
AtWRKY-IIC5c			 									
VVWRKY-IIC5b			 									
VvWRKY-IIC5c			 									
VvWRKY-IIC5d			 									

	470	480	490	500	510	520	530	540	550	560	570	580	
MtWRKY-IIC62			IMSD <mark>P</mark> SG	FQEQVHQ	HHNTTS	vv <mark>a</mark> ssvl.	<mark>G</mark>	NEFAMSQSQ	YAKFADANA	NVLHHQRQ	AV <mark>P</mark> SLLYNSDYI	NNSFNIS <mark>P</mark> L	
GmWRKY-IIC6a			FMH SEA S	9 6 F	· · · · · <mark>GP</mark> TS <mark>G</mark> LG	SA		· · · · HFML <mark>P</mark> QQQ	FRDQAQA - A	MLLYNSNS	5 S L S L <mark>P</mark> L	· · · · · · NV	
GmWRKY-IIC6b			· · · FMHSEAS	G - F	· · · · · <mark>GP</mark> TS <mark>G</mark> LG	S A • • • • • •	• • • • • • • • • • • •	· · · · HFMLQQQQ	FRDQAQAQA	AMLYNSTS	SLSL <mark>P</mark> L · · · ·	· · · · · · NV	
GmWRKY-IIC6c			· · · · FVSE <mark>P</mark> SS	F	· · · · · <mark>G</mark> YA <mark>G</mark> SF <mark>G</mark>	<mark>P</mark> S - - - - - -		· · · · HFML <mark>P</mark> HQH	- HQQAQSSA	LMYNNSTT	• • <mark>P</mark> L •	· · · · · · · NV	
GmWRKY-IIC6d			• • • FVTQ <mark>P</mark> • •	· · · · · · · · · · · · · ·	AAF G	P S		· · · · HFMLPHQ ·	AQASA	LS		· · · · · · · · · V	
AtWRKY-IIC62			· · · MLTSPIL	DH <mark>G</mark>	· · · · · · · ATTAS	55		· · · · SFSIPQPR	LLTQHHQPY	'NMYNNNSL	5 M	· · · · · · · IN	
VVWRK Y-IIC62					SAAA	AS		ertropy		NN			
BUNDEY.UCCA			IFADSSC	ENL	E CAA	43 Te		SEVEROPAL	VOOK . HASN	13			
PIWRKY-IIC6b			ILSDSTG			TS		· · · · SEVIPQ - TO	DYOG HAYL	YS			
AtWRKY-IIC66			- FFGSSGAAS		S L GNG	C F <mark>G</mark> F	<mark>P</mark> IDGSTI	ISPQFQQLVQY	HQQQQQELM	15		c	
AtWRKY-IIC6c			PSNGSAS		RAH I G	L <mark>P</mark>		· · · · TLPPQLLD	NNQQQQA <mark>P</mark> S	s		F	
GmWRKY-IIC6e			•••••NNF <mark>G</mark> S		V I M S <mark>G</mark>	s		· · · · A GNYMSQY	raaahaavhi	D		A	
GmWRKY-IIC6f			NNF <mark>G</mark> -		<u>T</u> VMS <mark>G</mark>	s <u>-</u>		· · · · A <mark>g</mark> nymsqy'	rhaa avhv	N		A	
VvWRKY-IIC6b			· · · · IATTFAG		· · · · · · · · <mark>P</mark> RLQT	т <mark>р</mark>	• • • • • • • • • • • •	· · · FQH · · LHLHI	1HQQQQH <mark>P</mark> SF	ND		· · · · · · · v	
GmWRKY-IIC6y			YANN FAS		· · · · · · · VL <mark>PLG</mark>	ΝΥ		LSQYQQQHI	HHQQQKLLV	N		· · · · · · · · T	
GmWRKY-IIC6h			· · · YANNFGS		· · · · · · · VL PPG	N Y		· · · · LSQYQHYHI	100000-HLV	N		· · · · · · · · · T	
PTWRK Y-IICGC			- SISDSGFSS		· · · · · · · · · I AAFA	M <mark>P</mark> · · · · · · ·		- MQRKLSLYQQI	IQSQQQ <mark>PP</mark> F <mark>P</mark>	L		· · · · · · · · · V	
MWRKY-IIC66			ASUSTIN	PAFC		sn				N		. .	
CHWRKY-IIC65	PP		- PANSLOVSA			L <mark>P</mark>		· IET · · TLNSH	OOYHL PP FV	N		T	
BdWRKY-IIC6a			· · AHLEMPPP		QHLG	LMA		· · · · · PPLFRT · I	DLMSMMQH				
OSWRKY-IIC62	GG H		HHHHL FMPGV		· · · · · HGLPPSH	LM		· · · · · <mark>P</mark> A G F H P · I	ELM <mark>G</mark> LMHHH -				
ZmWRKY-IIC62			AHLFM <mark>P</mark> NA		· · · · · HAL <mark>PP</mark> QH	LML		· · · · · <mark>P</mark> SSFR <mark>P</mark> AI	DLM <mark>G</mark> MVH <mark>P</mark> V-				
ShWRKY-IIC62			AHL FMSNA		· · · · H <mark>gglpp</mark> H	LM		<mark>P</mark> S S F <mark>G R P</mark> I	DLMSMMH <mark>P</mark>		<u></u>	<u></u>	
ZmWRKY-IIC6b	QLH	· · · · · · HH	IHHHN I V <mark>G</mark> AF <mark>P</mark> I	P	<mark>РРРРР</mark> QТМ	С L		P F VG RT G D G G G	· V I DML <mark>g</mark> l C -		· · · · L <mark>P P</mark> R N · · H	1AAV <mark>PP</mark> AAT	
ZmWRKY-IIC6c	QLH	· · · · · · · HF	IHNHHL <mark>G</mark> - AF <mark>P</mark>	• • • • • • • • • • • •	<mark>. T</mark> PLPQT <mark>G</mark>	A	• • • • • • • • • • • •	· A <mark>P</mark> FGRAAGGG ·	· VIDML <mark>G</mark> L		· · · · L <mark>PP</mark> RN · · ·		I IICé
SHWRKY-IIC66	QLH <mark>GG</mark>	· · · · · · · HF	IHHHHL <mark>GG</mark> AF <mark>P</mark>	P	· · · · PAPLPQQM	A <mark>G</mark>		APFGRAGGGGG	VIDML <mark>G</mark> L		· · · · L <mark>P P</mark> RNNNI	HAAM <mark>PP</mark> AIG	
OSWRKY-IIC6b	HAQAAAAA	AA AHQL	ннннаннанн		· · GMAPPLPLGS	• • • • • • • • •		GAAAQFGRSSG	· · IDVLSSF ·			HGMTTMGG	
BOWRK Y-IIC60	QAA		HIGHLINGAGG						· · I DVL <mark>GG</mark> L ·		ASOOSI I	HIG AMAP	
ShWRKY-HCGe		GLOCOMHY	HP-HFRMAPPI		RPGGGANAFDA		PSSS - 00000	GHHHHHGAAAM	OLAVS GG				
ZmWRKY-IIC6d	ASAHHPMS	GL····HHF	HHLRVP	· · · AALG · · · ·	GANDAFDA	LGL	LQQ				GAGLOOV	- AAAMSGH	
AtWRKY-IIC6d	VHR GGG	GS	LLHSF	R		LM			KHS <mark>P</mark> ANYQS				
AtWRKY-IIC6e	SGTTASDY	'N	<mark>P</mark> S S S <mark>P</mark>		•••••FSDLI				1 <u></u> T <mark>P</mark> R S F S N				
AtWRKY-IIC61	S A D L M T P R	s	••••• FAHDI	M • • • • • • • • • • •	· · · · · · · FRTAA	ΥТ····			I <mark>GG</mark> S VAAALD				
GmWRKY-IIC6i	SPS-FFGS		S YMASS		L <mark>P</mark> QD -	FL		A0	QLL <mark>P</mark> SYSQN -				
GmWRKY-IIC6j	SPS-FFGS		SYMA						· · · · · · · · · · · · ·				
UPWRKY-IICSC	- PS-LLIG	SA	TVGEE		FROE				ALLPIPNSNI				
PHWRKY-IIC64	. PS. 1145					F1		T					
PtWRKY-IIC61	- PS - LLAS				F <mark>P</mark> QD.			TI	RLLPASNOQ -				
VvWRKY-IIC6d	PSVLTP	G	QM <mark>GGPG</mark>		· · · · · · · F <mark>P</mark> QEL	F F			MAS PMNNLS				
PtWRKY-IIC6g	HS · · MLT P	A <mark>P</mark>	• • • • LAT <mark>GP</mark> RI	R	· · · · · · F <mark>P</mark> AH L	Q <mark>G</mark>		• • • • • • • • • • • • • • • • • • •	. VQM <mark>P</mark> AAT				
PtWRKY-IIC6h	HS · · MLAP	AP	····MAS <mark>GP</mark> S		· · · · · · F <mark>P</mark> · HH	Q <mark>G</mark>		• • • • • • • • • • • • • • • • • • •	VQI <mark>P</mark> DAM				
GmWRKY-IIC6k	P S S L L A T P	Т	· · · <mark>P</mark> LAA <mark>G</mark> SN		· · · · · F <mark>P</mark> QDL	FL			імнн <mark>р</mark> нно́ун				
GmWRKY-IIC6I	PSSLLATP	TH	PLAAGSN		FPQDL	FL			1MH Y P				
SBWRKY-IIC68	GGAATA	RYHSAAAVA	LAEQMSSS	· S S F · · · · · · ·	- IPAR - QLYS					SEA			
210 WKK 1-11060				1351	- I PAARQULFS	L <mark>P</mark>				SET1		············	
GmWRKY-IIC6m	·····	AFASOLAPT	MS-OFYYP			L <mark>P</mark>			20NN				
GmWRKY-IIC6n	HEA	AFAGQLAPT	MS - QFYYP			L P S			RENDD				
CpWRKY-IIC6d	· · HEA	SYAGRLMP A	IS-QFCY <mark>PG</mark> -			L S			RDA <mark>P</mark>				
PtWRKY-IIC6i	HQA	T F S SHMT P F	MS-QFIY <mark>PG</mark> -		мо	L <mark>P</mark>			REN <mark>P</mark>				
PtWRKY-IIC6j	••••HEA	T F A S HMT P P	TS-QFYY <mark>PG</mark> -		ма	L <mark>P</mark> • • • • • •			REN <mark>P</mark> · · · · · ·				
VvWRKY-IIC6e	QEV	T F D <mark>G</mark> RMT <mark>P F</mark>	LTSHFYY <mark>P</mark> D-	· · <mark>·</mark> · · · · · · · ·	MS	FH • • • • • • •	· · · <mark>·</mark> · · · · · · ·	· · · <mark>·</mark> · · · · · · · · ·	REVS····				
ZmWRKY-IIC6f*	AADGRHHL	P F V E F S A A S	SAHHHYHSNP	Y S <mark>G</mark>	TASCVNDH-HL	P P L L <mark>P</mark> T T T	TLAP - NALD - ·	- NGQQELQAAM	6 M <mark>P</mark> VASV <mark>P</mark>			т	
ShWRKY-IIC61*	AAAGHHHM	PFN - FSAA	HHHHLYGT		T SGVVT DHGHL	P P L L L P · T	TPAP - NALNDS	DNNGSPLACRT	STTTSSLLR <mark>P</mark>	LDCNHQE -		· · · L L L A A A	
BdWRKY-IIC6c*	SPHHGHHL	LYNDEHPLF	PMHGCSSS		· · · · · · · S S S L F C	к <mark>Р</mark>		· · · · · · · · · · · · · · · · · · ·	555555L				
OSWRKY-IIC44*			F SAHQQL			PPPIMP · ·			ASSSEL	SET			
BdWRKY-IICan*						AP				STE			
OsWRKY-IIC61	GAAGAYQF	GAPPPPLLG	FDEALAARSQ	D I	VSQEDRLSQQQ	RSLYYR	Q S CWAAT	ORSRAHNCITKG	RKPGLGPNN	T GAT WE L G	N · · PKTHRSRI	APGRTPHQ	

PpWRKY-IIC7a	YSMASN <mark>P</mark> ALE <mark>P</mark> SFS·SHS <mark>P</mark> FSLSDLLF································	ΙГ
PpWRKY-IIC7b	YSLACNQAVRAHI <mark>P</mark> ·NF <mark>G</mark> dF <mark>G</mark> lSdllff·································	
PpWRKY-IIC7c		IL

IIC7-

Appendix 4	continued
------------	-----------



ZmWRKY-IIC4b

	570	580	590	600	610	620	630	640	650	660
AtWRKY-IIC5a										
OsWRKY-IIC5a										
BdWRKY-IIC5a										
SBWRKY-IIC5a										
ZmWRKY-IIC5a										
BdWRKY-IIC5b										
OsWRKY-IIC5b										
ZmWRKY-IIC5b										
SbWRKY-IIC5b										
BdWRKY-IIC5c										
OsWRKY-IIC5c										
SbWRKY-IIC5c										
ZmWRKY-IIC5c										
BdWRKY-IIC5d										
OsWRKY-IIC5d										
ZmWRKY-IIC5d										
SBWRKY-IIC5d										
GmWRKY-IIC5a										
CpWRKY-IIC5a										
VvWRKY-IIC5a										
PtWRKY-IIC5a										
GmWRKY-IIC5b										
GmWRKY-IIC5¢										
AtWRKY-IIC5b										
SbWRKY-IIC5e										
OsWRKY-IIC5e										
GmWRKY-IIC5d										
GmWRKY-IIC5e										
GmWRKY-IIC51										
GmWRKY-IIC5g										
CpWRKY-IIC5b										
AtWRKY-IIC5c										
VvWRKY-IIC5b										
VvWRKY-IIC5c										
VvWRKY-IIC5d										

IIC5-

11	590	600	610	620	630	640	650	660	LLxD	680
MHMBKY HCCA	ENIVNEADNI	MICOTOR	Ţ				1 1			I A PMEM
GmWRKY-UC6a	VNSASCVNNS	ENTSSIS							HAFI RONGI LOD.	
GmWRKY-IIC6b	VNSASCVNNS	SYANTSSLS						- FVPSRVVAP	HIFLEDNGLLOD -	IVPTOM
GmWRKY-IIC6c	VSSSS TSN	YMNTTSF <mark>G</mark> ··			VQDHANRHG			FGH	EALLRDNGLLQDI	IVPTQV
GmWRKY-IIC6d	VSSSSN	1 Y I N T T S F <mark>G</mark>		<mark>G</mark>	FVQDHSTRR <mark>g</mark>			F <mark>G</mark> H	EALLRDNGLLODI	Ινρτον
AtWRKY-IIC62	RRSSD <mark>G</mark> TFVN	1 <mark> </mark>		<mark>G</mark> I	GYDMSQASTS -			т	SSSIRDH <mark>g</mark> llqd -	IL <mark>P</mark> SQI
VvWRKY-IIC6a	SSLSLNFN	NHMN 5 5 5 1 N		<mark>Р</mark> ТІ	FLHERRLR			<mark>P</mark> S <mark>P</mark>	ASLLRDH <mark>G</mark> LLED -	L L <mark>P</mark> S Q M
CpWRKY-IIC62	<mark>P</mark> S <mark>P</mark> SSNVI	INIST SNST <mark>GP</mark>	v	F I D <mark>P</mark> T S I	F S S F AN V <mark>P</mark>			P A P S P S P T P T	LSLLRDH <mark>G</mark> LLQD -	IV <mark>P</mark> SQM
PtWRKY-IIC6a	SS <mark>P</mark> SLNF -	TSSTTRFS		<mark>P</mark> S I	FSLQERRL			• • <mark>G</mark> N <mark>P</mark> • • • • •	PSLLRDHGLLQD -	IVPSQM
PtWRKY-IIC6b	SS <mark>P</mark> SLNIN	NTTSNTSFS		••••••••••••••••••••••••••••••••••••••	FSFHQRRS			••• D S <mark>P</mark> •••••	ASLLRDHGLLQD -	IVPSQM
AWWRICT-IIC60	A TEXINDOEN			DU					- VLVKDNGLLQD.	VVPSHM
GmWRKY-UC6e	ISSI-GFIS-	SSSS - RNA		. .	POFT				- ALLSDYGLLOD-	VVSSHM
GmWRKY-IIC6f	LSSL-GFLS-	SSSSSRNA		T	SQET				- ALLSDYGLLQD -	VVPSHM
VVWRKY-IIC6b	PLLNYGHG	S S F <mark>G</mark> S A		AR	SVQERRSCA			<mark>Р</mark> Т	SALHRDH <mark>g</mark> llqd-	IV <mark>P</mark> SHM
GmWRKY-IIC6g	LSSL - <mark>G</mark> F <mark>P</mark> YN	NDSSS <mark>P</mark> KNA		· · · · · v	FIQERRLCSNQ -			<mark>б</mark> т	NAFLRDH <mark>G</mark> LLQD -	VV <mark>P</mark> SHM
GmWRKY-IIC6h	LSSL - GFPYN	NDSSS <mark>P</mark> KNA		· · · · · · v	FTQERRLCSNH -			<mark>G</mark> Т	NAFLRDH <mark>G</mark> LLQD -	VV <mark>P</mark> SHM
PtWRKY-IIC6¢	S L S <mark>P</mark> L <mark>G</mark> F <mark>G</mark> Y N	NGSTTNAAN		· · · · · YL0	YEKRFYTS			••••••••••••••••••••••••••••••••••••••	SALLKDH <mark>g</mark> llqd-	LVPSHM
PtWRKY-IIC6d									-	
MIWRK Y-HCGB	LSSL-GFPS-	ESTS COCY.		· · · · · · · · · · · · · · · · · · ·	HOPPPICT				- PRVQDNGLLQD-	
BdWRKY-IIC6a					SIPPP		· · · PHPTATP	- PPPLQ - OH	HETDY ALLODI	FPSTMP
OSWRKY-IIC62		ANPSMYFPG		VA	ASAPPPPAV			DHPPLQ.QH	HFTDY - ALLQDI	FPSTMP
ZmWRKY-IIC6a		AN LSMFL <mark>P</mark>			• T M <mark>P P</mark> • • • • • • • •		· · · HM <mark>P</mark> S <mark>P</mark> AS	RAHPLQQQQH	QFTDYALLQDL	F P STMP
ShWRKY-IIC62	AM <mark>G</mark> A	AN <mark>P</mark> SMFL <mark>P</mark>			- SM <mark>PPP</mark>		•••HMST <mark>P</mark> SP	- A <mark>P P</mark> L Q Q H	HFTDY ALLQDL	F <mark>P</mark> STMP
ZmWRKY-IIC6b	G L A P S P C C M S	S <mark>GP</mark> VGTVAG		••••ATT	TTA <mark>P</mark> TTTTV····		· TTTTTSS <mark>PP</mark>	PPP SLQT MQH	FMAQDDF <mark>G</mark> LLQD-	MLPSSS
ZmWRKY-IIC6c	H <mark>G</mark> MSGG	3 - PMSAVAG		· · · · · AAT1	TTAAT I T		TTSTT <mark>P</mark> S	YPPSLQ-MQH	FMAQD - F <mark>G</mark> LLQDI	MLPSFG
SBWRKY-IIC6b	LASSRGMSGG	GPMSTVAG		· · · · · · AT A/	AT AAAT T		· · · T S S S S · ·	- PPSLQ-MQH	FMAQD - FGLLQD -	MLPSFV
DSWRKT-IICED		SOSVETEVO.					· I VAASAU <mark>P</mark> S	PPSLO.MON	FMAQD - LGLLQDN	MIDSEI
OsWRKY-IIC6c	- OMPOP - DHA	GIVALIAS			ASTT PPPP		- PATOSAAA	ATTPLR MOH	FMAODYG	FIPSPE
ShWRKY-IICGc	AALPDDGDQH			AAG	TTAATT		A	ASAPLR - MQH	FMAQDYAGLLQDM	FPSFVH
ZmWRKY-IIC6d	VALAAD - DQH	I <mark>G</mark> WAAA I A <mark>G</mark> VG		AAAGG	STTATATA		• • AAAT <mark>P</mark> AAA	ASA <mark>P</mark> LR - MQH	FMAQDYA <mark>g</mark> llqdm	IF <mark>PP</mark> FVH
AtWRKY-IIC6d		VGSLSYEH		<mark>G</mark> H	1 <mark>6</mark> T S S Y N			FNN	NQ <mark>P</mark> VVDY <mark>G</mark> LLQDI	V <mark>P</mark> SMF S
AtWRKY-IIC6e	· · · · · DD	DLFRV <mark>P</mark> YAS			/NVN <mark>P</mark> SY			•••HQQQH <mark>G</mark> F	HQQ <mark>ES</mark> EFELLKEM	IF <mark>P</mark> SVFF
AtWRKY-IIC61	Y <mark>c</mark>	aygas <mark>aya</mark> s		\	/NSN <mark>P</mark> SS			· · · HQVYH · ·	Q <mark>GG</mark> EYELLRE I	FPSIFF
GmWRKY-IIC6i	DH	IQNPMENQN		LSH	1NLHPQP			00000	FQLSRDYGLLQDL	
CHWRKY-HCG		PREMEVU		L S P					FQLSHDH <mark>G</mark> LLQDL	PAFID
VyWRKY-IIC6c		PNSMYYHN			• T PHH • • • • •				FQLPD - YGLLQDI	VPSFIR
PtWRKY-IIC6e	<mark>G</mark> D	QTSIFYQ			NHDPQY			qнq	LYA <mark>P</mark> D - Y <mark>g</mark> llqDl	VPSFIN
PtWRKY-IIC61	<mark>G</mark> D	0QTSMFYH			• <u>s</u> l A <mark>P</mark> QNQQ • • • •			••••LQQHQ	LYS <mark>P</mark> D·Y <mark>G</mark> LLQDL	V <mark>P</mark> S F I H
VVWRKY-IIC6d	· · · · · · · · · · · · · · · · · · ·	AA <mark>G</mark> SFY <mark>P</mark> Q		<u>.</u> .	• <mark>G</mark> L T <mark>P</mark> • • • • • • • •			• • • • • • • • FQ	Q L Q F H D Y <mark>G</mark> L L Q D V	V PSMIH
PtWRKY-IIC6g	S	SNKNL <mark>G</mark> EY	• • • • • • • • • • • • •	P C	NV				- NQV <mark>P</mark> DY <mark>G</mark> LLQDI	VINTSF
PtWRKY-IIC68	N	INQNM <mark>G</mark> AY		<mark>P</mark> (VHQ	QYQVPDYGLLQDI	VPSIFL
GmWRKY-IIC6I									····	VPSMLH
ShWRKY-IIC6d	- VVSSAATTS	FHQHVND <mark>G</mark> D -		E I	RQAS		SSRVSMAQS	PSTPSSVPP -	AISVEKAGLLDDM	
ZmWRKY-IIC6e	P V V S S V T T S	FQR - AND <mark>G</mark> D -		E I	.WRT <mark>G</mark> T	<mark>G</mark> `	Y <mark>G</mark> S <mark>G</mark> VSVAQS	P S T <mark>P P</mark> S V <mark>P P P</mark>	V V S V Q K A <mark>g</mark> l l D D M	IV <mark>P</mark> HGVR
AtWRKY-IIC6g	• • • • • • • • • • • • • • • • •	SPRL <mark>P</mark> RPTT -		• • • • • • • E I	DT <mark>P</mark> AVS				• T <mark>P</mark> S E E <mark>G</mark> • L L <mark>G</mark> D I	V <mark>P</mark> QTMR
GmWRKY-IIC6m	NNNNS F S	SSISQ <mark>P</mark> CQAY-		DI	DAE <mark>GG</mark> LSA		MM	P L <mark>P</mark> A D V S S Q S	Q <mark>P</mark> STDE <mark>G</mark> - LL <mark>G</mark> D I	V P P G M R
GmWRKY-IIC6n	NNNNT L S	ST I SQ <mark>P</mark> CQAQ -		DI	ET <mark>G</mark> E <mark>G</mark> S S V		MM	P ADGSSPS	Q <mark>PP</mark> TDE <mark>G</mark> - LL <mark>G</mark> D I	V P P G M R
CpWRKY-IIC6d	PTQNLSF	RLHRLAGEAG -		•••••E	• HMHSE • • • • • • •			TFP	QFPIDEG LLGDI	VPPGMR
PIWER T-IICEI	···· PSTVVC	SOPI PVEAR-		· · · · · · · · E (NTV <mark>P</mark>			·····	QLAIDG ··LLGDI	VPPGMR
VWRKY-IIC6	HSGSMAC	SHOVPAGAR.		F	ARALAF					VPPGMR
ZmWRKY-IIC6f*	TSLLPPPTTA	ALAAPPVASV-		SI	ATSSSSL			PPASSSAA	PAAVDKG-LLDDM	VPPAMR
SHWRKY-IIC6f*	ASYPLSSAA	AMSSM <mark>P</mark> VPSM-		s1	TTTTSL			PPPASS	AVDKG LLDDM	V P P AMR
BdWRKY-IIC6c*	LQQLHCNF	RQELQAAAS		<u>.</u>	YTTASS			•••т <mark>v р</mark> а • • • •	<u>-</u> VDK <mark>g</mark> - L L DDM	IV <mark>PPG</mark> MR
OsWRKY-IIC6d*	Q L <mark>P</mark> L H C N -	HELQVVASC -		<mark>G</mark> G	9 Y <mark>P</mark> S S S S		• • • • • • • • • • •	· · SPPASVL ·	· · <mark>P</mark> VDK <mark>G</mark> · L L D D M	IV <mark>P</mark> RAMR
OsWRKY-IICGe*	AVSCSMPTTT	SLQELNNSE -		••••• <mark>6</mark> 1	QRPGYNN	s <mark>i</mark>	QAAVT IAQR	PPSP·SVPP·	AVSFDKG - LLDDI	VPPGVR
BOWRKY-IIC6d*	ETGVSSTTTA	ASLOQLNGSN-		· · · · · E I	KSPT		···· AAMARS	PSSS·SVPP-	SVSFDMG - LLGDI	VPPGVR
03/mnn 1-11061			👱	• vor i Hi	I LEIKING VG			N ar HUNAF	is is viriar IR <mark>u</mark>	V DUUU

PpWRKY://C7a ····SCDS-VSST<mark>P</mark>EQQQQG······FHLDYN··SLISQAAAFA<mark>P</mark>······EQ<mark>P</mark>FYPRPEYLNR··SFTAHSAALSRELSVS· PpWRKY://C7a QTKLKCNVQIQDEFSSRVQTRI<mark>P</mark>IL······WIFKTMHVPEK··SSLSVAVDFQLRCCTQHYREAHRVVPTPRF<mark>P</mark>EA···SSDFPISLTQMVKSST PpWRKY://C7c YQHQMTSS<mark>P</mark>SQCDDLVRRVLEAEEE······

∥ IIC6-





Appendix 5













ZmWRKY-IID3a	AVLLVPPCASNVTLT	· · PAKKFDRSMFLETPLLELN · · · · · · · · · · · · · · · · · · ·	SC <mark>GVPP</mark> S <mark>P</mark> SAS <mark>P</mark> AMAAVQRNSTKVAAA <mark>P</mark> A <mark>PNPCT<mark>S</mark>AP</mark> RIQFH <mark>PP</mark> H
ShWRKY-IID3a	AVLV <mark>SPPCAS</mark> NVTLT <mark>P</mark> A	PAKKFDRSMFLE <mark>T</mark> SLLELNN	SCS <mark>VPP</mark> S SS <mark>P</mark> AMA - VQKSS <mark>P</mark> KVAA <mark>P</mark> N <mark>PC</mark> T <mark>STP</mark> HIQLQ <mark>P</mark> TT
OsWRKY-IID3a	ATAAAA <mark>SS</mark> AANVT <mark>P</mark> A·····	····VVDRSLFLE <mark>T</mark> TLLDLN·····	••••• SR <mark>GAP</mark> A <mark>P</mark> •••• AASMAAAAKNSSKLA <mark>P</mark> •••••• <mark>P</mark> MVN <mark>SS</mark> S <mark>S</mark> ANH I QFQ••••

IID3-







Intel Intel <td< th=""><th></th><th></th><th></th></td<>			
нажи:1 10		WRKY DNA-binding domain	
u u	Basic-1	-	
	480 490 500 510	520 530 540 550	
	GmWRKY-IID12 ··· RKMRLK····· RVVRVPAISLKMADIPPDDYSWRKYGQKPIKGSP	H <mark>PR - GYYKCSSVRGCPARKHVER</mark> ALDDPAMLVV <mark>TY</mark> EGE	
UVWRY/1005 - WENNY 2016 - WENNY 2016 - WENNY 2016 WENNY 2016 GOVER - GOVER COUNT OF ALL DUTY AND SALE DUTY LEVEL OF ALL DUTY AND SALE DUTY AND	MtWRKY-IID1a ··· RKMRLK···· RVVRVAAISLKMADIPTDEYSWRKYGQKPIKGSP	HPR-GYYKCSSIKGCPARKHVERALDDPSMLVVTYEGD	
DJMWRY/1002 Important DJWWRY/1002 Important DJWWRY/1002	VVWRKY-IID12 ··· RKMKLK····· RVVRVPAIISMKMADIPPDDFSWRKYGQKPIKGSP	HPR - GYYKCSSVRGCPARKHVERALDDPKMLIVTYEGE	
Guiden V, HDD		HPR-GTTRCSSVRGCPARKHVERALDDPSMLIVTTEGE	
UVWWYYJD25 INKSENNEL UV UV AL TENNAD I PDDF WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF WLYDERD AMWRYJD25 INKSENNEL VI WYPALAN KAAD I PDDF WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF WL VYEED AMWRYJD25 INKSENNEL VI WYPALAN KAAD I PDEY WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF ALL VYYEED AMWRYJD25 KINSENNEL WI WYPALAN KAAD I PDEY WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF ALL VYYEED AMWRYJD25 KINSENNEL WI WYPALAN KAAD I PDEY WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF ALL VYYEED GWWYJD25 KINSENNEL WI WYPALAN KAAD I PDEY WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF ALL VYYEED GWWYJD25 KINSENNEL WI WYPALAN KAAD I PDEY WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF ALL VYYEED GWWYJD25 KINSENNEL WI WYPALAN KAAD I PDEY WWWYDGRF I GOPHPE - DWYCER VOCPARKHVE ALDDF ALL VYYEED GWWYJD25 KINSENNEL WI WYPALAN KAAD I PDEY WWWYDGRF I GOPPFE - DWYCER VOCPARKHVE AND PALL VYEED GWWYJD25 KINSENNEL HU WYPALAN KAAD I PDEY WWYYDGRF I GOPPFE - DWYCER VOCPARKHVE AND PALL VYEED GWWYJD26 KINSENNEL HU WYPALAN KAAD I PDEY WWYYDGRF I GOPPFE - DWYCER VOCPARKHVE AND PALL VYEED GWWYJD26 KINSENNEL HU WYPALAN KAAD I PDEY WWYNDGRF I GOPPFE - DWYCER VOCPARKHVE AND PALL VYEED GWWYJD26 KINSENNEL HU WYPALAN KAAD I PDEY WWYNDGRF I GOPPFE			
Adversey-Jude b	VVWRKY-IID16 ···RKLRVK····RVVRVPAISTKMADIPPDDFSWRKYGQKPIKGSP	HPR - GYYKCSSVRGCPARKHVERALDDPTVLTVTYEGD	
G#WRYY-MD15 • EKS EVM • EKS EVM <td>AtWRKY-IID16 ··· RKSRVK···· RVIRVPAVSSKMADIPSDEFSWRKYGQKPIKGSP</td> <td>H<mark>PR - GYYKCSS</mark>VRGCPA<mark>RKH</mark>VERALDDAMMLIV<mark>T</mark>YEGD</td> <td></td>	AtWRKY-IID16 ··· RKSRVK···· RVIRVPAVSSKMADIPSDEFSWRKYGQKPIKGSP	H <mark>PR - GYYKCSS</mark> VRGCPA <mark>RKH</mark> VERALDDAMMLIV <mark>T</mark> YEGD	
ANNEY 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT FOOD ARMY FALDD AMLLY VK OT GWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT FOOD ARMY FALDD AMLLY VK OT GWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT FOOD ARMY FALDD AMLLY VK OT GWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT FOOD ARMY FALDD AMLLY VK OT GWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT FOOD ARMY FALDD AMLLY VK OT GWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALDD AMLLY VK OT GWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALDD AMLLY VK OT PWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALDD AMLLY VK OT PWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALDD AMLLY VK OT PWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALDD AMLLY VK OT PWWYN 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALTONAMI VK YOF OWWRY 1022 KST MUNICATION TO VARANA AND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALTONAMI VK YOF OWWRY 1022 KST MENNANA - TT I VAA I SYND POLYMMEN YORF I GOPHER - VY KOT MOCAARMY FALTONAMI VK YOF OWWRY 1022 KST MENNANA - TT I VAA I SYND POLYMMEN YORF I GOPHER - VY KOT WOCAARMY FALTONAMI VK YOF OWWRY 1022 KST MENNANA - TT I VAA AAAAN OF ADD FWENY GOVER I GOPHER - VY KOT WOCAARMY FALTONAMI VK YOF SWWYY 1022 KST MENNANA - TT I VAA AAAAN OF ADD FWENY GOVER I GOPHER - VY KOT WOCAARMY FALTONAMI VK YOF SWWYY 1022 KST MENNANA - TT I VAA AAAANA OF ADD FWENY GOVER I GOPHER - VY KOT WOCAARMY FALTONAMI VK YOF SWWYY 1022 KST MENNANA - TT I VAA AAAANA OF ADD FWENY GOVER I GOPHER - VY KOT WOCAARMY FALTONAMI VK YOF SWWYY 1022 KST MENNANA - TT I VAA AAAANA OF ADD FWENY GOVER I GOPHER - VY KOT WOCAARMY FALTONAMI VK YOF SWWYY 1022 KST MENNANA - TT I VAA AAAANA OF ADD FWENY GOVER I GOPHER - VY KOT WOCAARMY FALTONAMI Y YOF SWWYY 1022 KST M	CpWRKY-IID16 ··· KKSRVK···· RVIRVPAISNKMADIPPDDYSWRKYGQKPIKGSP	HPR-GYYKCSSVRGCPA <mark>RKH</mark> VERALDDPMMLIV <mark>T</mark> YEGD	
ZIWWRKY-11039 KKPER EVEVPALED NAD I PADNYEWRKYOGKPI KOBPHPE OYYRCSEK KOCPARKHVER CRSDAAMLIVTYEND SBWRKY-11039 KKPER VVIVYALI DONAD I PADNYEWRKYOGKPI KOBPHPE OYYRCSEK KOCPARKHVER CRSDAAMLIVTYEND OSWERY-11039 KKPER VVIVYALI DONAD I PADNYEWRKYOR PI KOBPHPE OYYRCSEK KOCPARKHVER CRSDAAMLIVTYEND	ARWRY, HID22 K S EN N MAN C T V V V A L AK I AD I P D L Y WRW YORK I KORP I KOP ARWRY, HID28 K S EN N MAN C T V V V A V AK I AD I P D L Y WRW YORK I KOP I KOP ARWRY, HID28 K S EN N MAN T V V V A V AK I AD I P D L Y WRW YORK I KOP I KOP GWWRY, HID28 K EN N MAN T V V V A V AK I AD I P D L Y WRW YORK I KOP I KOP GWWRY, HID28 K EN N MAN T V V V V A S I AD I P D L Y WRW YORK I KOP I KOP GWWRY, HID28 K EN N MAN K V V V V A S I AD I P D L Y WRW YORK I KOP I KOP GWWRY, HID27 K EN N MAN K V V V A S I AD I P D L Y WRW YORK I KOP I KOP GWWRY, HID28 K EN N MAN K V V V A S I AD I P D L Y WRW YORK I KOP I KOP GWWRY, HID27 K EN N MAN K V V V A S I AD I P D L Y WRW YORK I KOP I KOP GWWRY, HID28 K EN N MAN K V V A S I AD I P D L Y WRW YORK I KOP I KOP JWWRY, HID27 K EN N MAN K I N V A I S K I AD I AD L Y WRW YORK I KOP JWWRY, HID28 K EN N MAN K I N V A I S K I AD I AD L Y WWR YORK I KOP JWWRY, HID28 K EN N MAN K I N V A I S K I AD I AD L Y WWR YORK I KOP JWWRY, HID28 K REN N V AN N I Y V A I S K I AD I AD L Y WWR YORK I KOP K KOP JWWRY, HID29 K KEN N V AN N I Y V A I S KI AD I AD L Y WWR YORK I KOP	HP R - G Y Y K G B T F K G G P A KKH VE K A L D D T AM L I V T Y E G E H R - G Y Y K G B T F K G G P A KKH VE K A L D D T M L I V T Y E G E H R - G Y Y K G B T F K G G P A KKH VE A A L D D T M L I V T Y E G E H R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G E Y R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G B V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L I V T Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V T Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y K G T V K G T A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G P A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G T A KKH VE A A D D T M L V Y Y E G Y R - G Y Y K G T V K G T Y K G P A KKH Y E A D D T M L V Y Y E G Y R - G Y Y K G T Y K G T Y G C A A KKH Y E A D D T M L V Y Y E G Y	IID2-
	ZIIWRKY-11023 KKP N EK B V N V A I BOKNAD I PADNYBWRKY OGKP I KGSP S9WRKY-11023 KKP NEK V V V V A I BOKNAD I ADNYBWRKY OGKP I KGSP S9WRKY-11023 KKP NEK N V V V A I BOKNAD I PADNYBWRKY OGKP I KGSP	NPR - GYYRCSSKRDCPARKHVERCRSDAAML I VTYEND HPR - GYYRCSSKRDCPARKHVERCRSDAAML I VTYEND HPR - GYYRCSSKRDCPARKHVERCRSDFAML L VTYEND	IID3-



Appendix 6

							20		40							74							400			
			۳.		ī.		Ĩ		1.		1			1		- îî					ĩ		100	1		
AtWRKY-IIE1a		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
VvWRKY-IIE1a		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
VvWRKY-IIE1b		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
OSWRKY-IIE1a		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
BdWRKY-IIE1a		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
ShWRKY-IIF1a		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 /	 	 	
ZmWRKY.IIE1a		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
GmWRKY-HE43		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
GmMDKY HEAL		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
CINVERT IN																										
MINNER J-HE 12		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
GMWRKY-IIE1¢	•••	 	 	 • • •	 	 	 	 	 	 	 	 	• • •		 	 		• •	 	 	 	 	 	 	 	
GmWRKY-IIE1d		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
MWRKY-IIE1b		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
AtWRKY-IIE1b		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 · · ·	 	
OsWRKY-IIE1b		 	 	 	 	 	 	 	 · · ·	 	 	 			 	 	· · ·		 	 	 	 	 	 	 	
ShWRKY-IIE1b		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
ZmWRKY-IIE1b		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	
ZmWRKY-IIE1c		 	 	 	 	 	 	 	 	 	 	 			 	 			 	 	 	 	 	 	 	

	matic P		
	mour E		
SbWRKY-IIE2a		AGAMAMAAA	
ZmWRKY-IIE2a	·····MDSMEDAVS·····GTAAQGDLAEVVARA	AG AMAAAAS	
AtWRKY-IIE2a	RNTLVGFAMLVQVAFSEGYCDDTDFGISCVCKW	KNKE <mark>G</mark> HSHRRE I N L H CWAL <mark>G</mark> KAVERDHT F V F F D V N MR P D T D E <mark>G</mark> N D P D I WAD L V V F E F F <mark>P</mark> V N K Q R K <mark>P</mark> L N D	sc
SbWRKY-IIE2b	· · · · RARSWFMR <mark>G</mark> VLLRMAEHFNDW <mark>DL</mark> QA <mark>VV</mark> RSC	C <mark>G</mark> SVAAH <mark>P</mark> DPAAPRAEPDAAPPEPTTTT <mark>P</mark> VAA <mark>PP</mark>	P V
OsWRKY-IIE2¢	SSLSRFH <mark>P</mark> LLKKLIMEYSNDW <mark>DL</mark> QALVRSC	C <mark>G</mark> TAVADSE <mark>P</mark> ······VFV <mark>G</mark> RAGGVP·····	
ShWRKY-IIE2c	· · · · · · · · · · · · · · · · · MDCSNDWDLQALVRSC	C <mark>GGGG</mark> TAAAAACNS <mark>G</mark> AA <mark>P</mark> TATR <mark>GG</mark> YDA <mark>P</mark> SREAADDASVVV <mark>GGGG</mark> RV	
OsWRKY-IIE2d	•••••MDGDAWWYPGGGGGGGSNNWDLGAVVRFG	GCGGGRVS <mark>P</mark> AAALLG····EAWEYDDD <mark>P</mark> FSSFLA <mark>P</mark> ·······························PMTAQQAAL <mark>P</mark> AVW····	
ZmWRKY-IIE2¢	NNWDLDAVVRL	<mark>9</mark> C - RRRLS <mark>PP</mark> R	
SbWRKY-IIE2d	····ME <mark>G</mark> ·DLRWCC <mark>G</mark> S···SSNDW <mark>DL</mark> HAVVRLA	A <mark>S</mark> · · · · · · · · · · · · CS <mark>GG</mark> SRSRVTS <mark>P</mark> S · · · · ·	
ZmWRKY-IIE2b	····MES·DLRWCSGS···SNNDWDLHAVVRF	ASCSGGSCVPSP	
ZmWRKY-IIE2d	····MEEERCF·····NNWDLDAVVRLG	<mark>G</mark> C - RRL L S <mark>P</mark> S <mark>P</mark> Q P D P F A S F L P P	
AtWRKY-IIE2b	·····MADDWDLHAVVR-G	GCSAVSSSATTT	
VvWRKY-IIE2b	· · · · · · · · · · · · · · · · · · ·	CAINEASSAAF	
VvWRKY-IIE2a		GCATATVASAAAALDLHRSSCFSP	
GmWRKY-IIE2a		GCTTTTSSAATSGFGACNPPSTSSCF	
GmWRKY-IIE2b	·····MAEDWDLHAVVR·G	GCSTVTSSSVS	
GmWRKY-IIE2c	MEEDWDLHAVVR-C		
MINNER T-HEZD			
AIWRN 1-HE2C	MUEGULEATVRO1	15	
	MCEVVCHEGWDLOALVPAC		
CPWRKY-IIE2a GmWRKY-IIE2a		GCTN····································	
CpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2e		GCTN	
CpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2e AtWRKY-IIE2d		GCTN FEACGOF KVP SSCFGLOHHEL S TGEAIT MDD NADFSYFF N TGEATT MDD NADFSYFF SSSSSYSTINS CAGHED LONCKOO OD PPPP	· · · · ·
CpWRKY-IIE2a GmWRKY-IIE2a GmWRKY-IIE2a AtWRKY-IIE2a CpWRKY-IIE2b	MGEYVCMEGWD LCA IVRA MDEFACLRDWD LEA IVRG MDEFACLRDWD LEA IVRG MSSEDWD LFA VR S MEDFWD LSA IVR S	GCTN	 S F
CpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2e AtWRKY-IIE2a CpWRKY-IIE2b GmWRKY-IIE25	MGE YUCME EWD LOAIV RAG MDE FACL ROWD LEAIV RAG MDE FACL ROWD LEAIV RAG MSSEDWD LFAVV R-S MSSEDWD LFAVV R-S MAEDWD LFAIV R-S	GCTN FEACGOF KVP 535CFGLOHHEL S TO EAIT MDDP NS TO EATT MDDPNPOFSYFF SCSSVSTTNS CAGHEDD LGNCKQQ OD PPPPPL SCRGKA.SVDK A TESDDNSLA SCQSAT TITP TTINTSSP	 S F
CpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2d AtWRKY-IIE2d CpWRKY-IIE2b GmWRKY-IIE2f GmWRKY-IIE2f	MGE YV CME EWD LOAIV RAG MDE FACL RDWOL LEAIV RG MDE FACL RDWOL LEAIV RG MSSEDWOL FAVV R-S MEDDWOL SAIV R-S MAEDWOL FAIV R-S MAEDWOL FAIV R-S	GCTN	5 F 5 S
CpWRKY-IIE2a GmWRKY-IIE2a GmWRKY-IIE2a AtWRKY-IIE2a CpWRKY-IIE2b GmWRKY-IIE2b GmWRKY-IIE2p GmWRKY-IIE2p	MGE YUCME EWU LAAIV RA MDE FACL ROWD LEAIV RG MDE FACL ROWD LEAIV RG MSE SEDWD LFAVVR S MED FOU LFAIVR S MAEDWD LFAIVR S MAEDWD LFAIVR S MTHEDWD LFAIVR S	GCTN FEACGOF KVP \$SSCFGLOHHEL S TGEATT MDD NADFSYFF SCSSSVSTTNS CAGHEDDLONCKQQ ODP PP \$LF0ASS SCRGKA SCASAT CLASLTFEEED \$P\$ SCQSAT TIPQTTTNNTSS\$ LITST SCKAATFATTN TTSCGBLOEN \$LSEN \$LSCASEN	5 F 5 S 5 S 5 S
CPWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2d AtWRKY-IIE2d CpWRKY-IIE2d GmWRKY-IIE2f GmWRKY-IIE2g GmWRKY-IIE2g GmWRKY-IIE2i	ME E YUCME EWD LOAIV RAG ME FACL RDWOL EAIV RG ME FACL RDWOL EAIV RG MS E FOUDU EAIV RG MS E FOUDU FAVV RS ME DDWOL FAIV RS ME DDWOL FAIV RS MAEDWOL FAIV RS MT HE DWOL FAIV RS OUDU FAIV RS	GCTN FFACGOF KVP \$SSCFGLOHHEL S TGEAIT MDD NADFSYFF N TGEATT MDD NADFSYFF SSSSSVSTINS CAGHEDD LONCKQQ ODPPPP [FOASS SCRGKA-SVDK A TESDDNSLA ODPPPP [FOASS SCQSAT TIP QTTINNISS LITST SCQSAT SCQSAT TIP QTTINNISS LITST SCQSAT SKAAFFATN TEP PT TTAT NTTSCS LSQENQ SF NTTSCD LSQENQ SF	SF SS SF SF
CPWRKY-IIE2a GmWRKY-IIE2a GmWRKY-IIE2b GmWRKY-IIE2b GmWRKY-IIE2b GmWRKY-IIE2b GmWRKY-IIE2g GmWRKY-IIE2i GmWRKY-IIE2i GmWRKY-IIE2j	MGE YUCMEEWD LAAIV RAA MDE FACL RDWOL LEAIV RG MDE FACL RDWOL LEAIV RG MSEEDWOL FAVVR - 5 MEDBWOL SAIVR - 5 MAEDWOL FAVVR - 5 MAEDWOL FAVVR - 5 MAEDWOL FAVVR - 5 MTHEDWOL FAVVR - 5 MTHEDWOL FAVVR - 5 MTHEDWOL FAVVR - 5	GCTN	SF SF SF SF
CPWRKY-IIE2a GmWRKY-IIE2a GmWRKY-IIE2a CpWRKY-IIE2b GmWRKY-IIE2b GmWRKY-IIE2j GmWRKY-IIE2j GmWRKY-IIE2j GmWRKY-IIE2j GmWRKY-IIE2j GmWRKY-IIE2j	ME E YU CME EWO LOA I V RAG MOE FACL ROWO LEA I V RG MOE FACL ROWO LEA I V RG MS E DOWO LEA V RG MS E DOWO LEA V R MAE DOWO LEA V R MAE DOWO LEA V R MAE DWO LEA V R MAE DWO LEA V R MT HE DWO LEA V R TOT YAC LMEWO LOA I V G ME V V CAN VG CAN VG ME V V CAN VG CAN VG CAN ME V V CAN VG CAN VG CAN VG CAN VG CAN VG ME V V V CAN VG CAN	GCTN FFACGOF KVP \$SSCFGLOHHEL S TGEAIT MDD NADFSYFF N TGEATT MDD NADFSYFF SSSSVSTTNS CAGHEDD GNCKQQ ODP PPP LFOASS SCRGKA-SVDK A TESDDNSLA SGQSAT TI PQTTINNTSS LITST SCKAATFTATTN TETP PTTTTTTS NTTSCSLSGENGSF SKKAATFTATTN TETP PTTTTTTS NTTSCLSLGENGSF SKKAATFTATT NTTSCLSGENGSF NTTSCLSGENGSF N KKAPAATV MDN MLIFSGHFGL	S F S F S F S F S F
CPWRKY-IIE2a GmWRKY-IIE2a GmWRKY-IIE2e AtWRKY-IIE2e GmWRKY-IIE2f GmWRKY-IIE2f GmWRKY-IIE2i GmWRKY-IIE2i GmWRKY-IIE2i GmWRKY-IIE2e OsWRKY-IIE2e	MEE YV CMEEWO LOA I V RAG MDE FACL RDWOL EA I V RG MDE FACL RDWOL EA I V RG MSSEDWOL EA I V RG MSSEDWOL FAU V R MEDDWOL FAI V R MEDWOL FAI V R MEDWOL FAI V R MAEDWOL FAI V R MAEDWOL FAI V R MT HEDWOL FAI V R T CT YACLMDWOL OA I V G C MMEEL CG	GCTN	S F S S S F S F
CDWRKY-IIE28 GmWRKY-IIE20 GmWRKY-IIE20 AtWRKY-IIE20 GmWRKY-IIE29 GmWRKY-IIE29 GmWRKY-IIE29 GmWRKY-IIE21 GmWRKY-IIE23 GmWRKY-IIE28 BdWRKY-IIE28	MEE EY UCMEEWO LOA IV RAG MEE EY UCMEEWO LOA IV RAG MEE FAC L RDWO LEA IV RG MES EDWO LEA IV RG MS E DWO LEA IV RG MAE DWO LEA IV RS MAE DWO LEA IV RS MAE DWO LEA IV RS MAE DWO LEA IV RS MT HE DWO LEA IV RS T GT YAC LIMOWO LAA IV RG MEE EL LG MEE EL CG MEE DWO LAAV RE	GCTN	S F S F S F S F
CpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2d CpWRKY-IIE2d GmWRKY-IIE2f GmWRKY-IIE2f GmWRKY-IIE2i GmWRKY-IIE2i GmWRKY-IIE2a StwKrY-IIE2a StwKrY-IIE2a	MEE YV CMEEWOL GA I V RAG MDE FACL RDWOL EA I V RG MDE FACL RDWOL EA I V RG MSE DDWOL FAU K RS MEDDWOL FAU K RS MAEDWOL FAI V RS MAEDWOL FAI V RS MTHEDWOL FAI V RS MTHEDWOL FAI V RS ************************************	GCTN FFACGOF KVP # SSCFGLOHHEL S TGEAIT MDD NADFSYFF N TGEATT MDD NADFSYFF SSSSVSTINS CAGHED LIGNCKQQ ODP PPP [FAASS SCRGKA-SVDK A TESDDNSLA ODP PPP [FAASS SCQSAT TT IFQTTINNTSS LITST LITST SCQSAT TT IFQTTINNTSS LITST SCQSAT SCKAAFFATN TE PTTINTSS NTTSCS LSQENQSF SCKAAFFATN TE TPTTITTTS NTTSC CD LSQENQSF SCKAAFFAT N TE TPTTITTTS NTTSC CD LSQENQSF N KAPAATV MDN HL IFSOHFG GC RRAIS GC RRAIS QQUD FASFL PR QQGVAMEVAA GC RRAISF QOD PRAFE PP RKKAFV UH L GC RRAISF QOD PRAFE PP RKKAFV UH L	S F S F S F
GpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2a BdWRKY-IIE2a BdWRKY-IIE2a SbWRKY-IIE2a SbWRKY-IIE2a	MEE YV CMEEWOL GA I V RAG MDE FACL RDWOL EA I V RG MDE FACL RDWOL EA I V RG MS SEDWOL FAV V R - 5 MAEDDWOL FAV V R - 5 MAEDDWOL FA I V R - 5 MAEDDWOL FA I V R - 5 MAEDDWOL FA I V R - 5 MTHEDWOL GA I V G CT MEEELCG NNVWOL GAV V RF MEEERCF NNVWOL DAVV V RF MEEERCF NNVWOL DAVV V RF MEEERCF NNVWOL DAVV V RF MEEERCF NNVWOL DAVV V RF	GCTN. FEACGOF KVP # SSCFGLOHHEL S. TEAT MDD NAP 5Y FF. N. TEATT MDD NAP 5Y FF. SCSSSVSTNS CAGHEDDIGNCKQQ OD PPPPL FQASS. SCRGKA.SVDK A. CLASS LTFEEED # F SQSAT TI I # QTTINNTS S# LITST SCRGKA.SVDK A. TESDDNS LA SQSAT TI I # QTTINNTS S# LITST SCRGKA.SVDK A. TESDDNS LA SQSAT TI I # QTTINNTS S# LITST SCRGKATFTATTN TE 1 # PTTITTTTS S. NTTSCS LSOENG SF. SCKAATFTATTN TE P # TTATT NTTSCS LSOENG SF. N KAP AATV MDN # LIFS AHFGH LG. GC RRR IS # AAVA QOVD # AST L. OGVAME VAA GC C RRALS # P M P AST L. OGVAME VAA GC C RRALS # P M P F S # L P M P S # M KK K V V # AF GC C RRALS # P M OF F AST L. P M P S # M KK K V V # AF GC RRALS # P M OF P AST L. P M P S # M KK K V V # AF GC RRALS # P M OF F AST L. P M C M AF GC RRALS # P M OF F AST L. P M C M AF GC RRALS # P M OF F AST L. P M AF <th>S F S F S F S F</th>	S F S F S F S F
GpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2d GpWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d SbWRKY-IIE2a SbWRKY-IIE2a SbWRKY-IIE2a ZmWRKY-IIE2a ZmWRKY-IIE2a	MEE YV CMEEWOU GAI V RAA MDE FACL RDWOL EAI V RG MDE FACL RDWOL EAI V RG MSSEDWOL FAU V RG MSSEDWOL FAU V RS MAEDDWOL FAU V RS MAEDDWOL FAU V RS MAEDDWOL FAU V RS MTHEDWOL FAU V RS ************************************	GTN. FFACGOF KVP SSCFGLOHHEL S. TGEAIT MDD NADFSYFF. N. TGEATT MDD NADFSYFF. SSSSUTTNS. CAGHEDD LONCKQQ ODPPPP [FOASS. SCGSAT TIP GTTINTSSP. CLAS.LIFEEEED F SGGSAT TIP GTTINTSSP. LITST. SKAAFFATN TEP TPT TATT NTTSCD LSOENGSF SKAAFFATT NTTSCD LSOENGSF NTTSCD LSOENGSF N KAAFATV. MDN SHLIFFOHLD GC RRAISF. NON FASFLPP OQGVAMEVAA GC RRAISF. OQ OP FASFLPP PTS SFL GC RRAISF. OQ OP FASFLPP PTS SFL GC RRAISF. OD OP FASFLPP PTS SFL GC RRAISF. OD OP FASFLPP PTS SFL GC RRAISF. OD OP FASFLPP PTS SFL GC RRAISP. OD OP FASFLPP PTS SFL	S F S F S F
CpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2d SbWRKY-IIE2a SbWRKY-IIE2a SbWRKY-IIE2a ZmWRKY-IIE2a ZmWRKY-IIE2a	MEE Y V CMEE WOL CA I V RAG MOE FACL R R WOL EA I V RG MOE FACL R R WOL EA I V RG MS SE D WOL EA I V RG MS SE D WOL FAU V R - S MAE D WOL FAI V R - S MT HE D WOL FAI V R - S MT HE D WOL FAI V R - S MEE E R G MEE A D UL I RMAEH FN D WOL CAV V R F MEE A D UL RMAEH FN D WOL CAV V R S MEE MAEH FN D WOL CAV V R S	GCTN. FFACGOF KVP PSSCFGLOHHEL S. TGEAIT MDD NADFSYFF. N. TGEATT MDD NADFSYFF. SCSSSVSTNS CAQHEDDIGNCKQQ ODPPPPILFQASS. SCRGKA.SVDK A. TESDDNSLA CLAS.LTFEEELEOPF. SQSAT TIP QTTINNTSS LITST SCRGKA.SVDK SKQAATFTATN TESDDNSLA CLAS.LTFEEELEOPF. SKQAATFTATN TIP QTTINNTSS LITST SKAATFTATN TETP PTTITTTTS NTTSCDSLSENGSF. SKAATFTATN TETP PTTTTTTS NTTSCDSLSENGSF. SCRARAS QQVD FASSL QQVD FASSL GC RRR IS FAAVA QQVD FASSL QQVM WAA GC CRR IS PAAVA QQVD FASSL QQVM AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAA	S F S F S F
CpWRKY-IIE24 GmWRKY-IIE24 GmWRKY-IIE24 GpWRKY-IIE24 GmWRKY-IIE24 GmWRKY-IIE24 GmWRKY-IIE24 GmWRKY-IIE24 GmWRKY-IIE24 GmWRKY-IIE24 BdWRKY-IIE24 SbWRKY-IIE24 SbWRKY-IIE24 ZmWRKY-IIE24 ZmWRKY-IIE24 ZmWRKY-IIE24	ME E YU CME EWO LOA I V RAG MOE FACL ROWOL EA I V RG MOE FACL ROWOL EA I V RG MS E DOWOL FA I V RG MS E DOWOL FA I V RS MAE DOWOL FA I V RS MAE DOWOL FA I V RS MAE DOWOL FA I V RS MT HE DWOL FA I V RS MT HE DWOL FA I V RS MT HE DWOL FA I V RS ME E E L CG ME E E L CG ME E E RG ME E RG	GCTN FFACGOF KVP # SSGFGLOHHEL S TGEATT MDD NADFSYFF N TGEATT MDD NADFSYFF SSSSUTTNS CAGHEDD LONCKQQ ODP PPPL FOASS SGGSAT A TESDDNSLA CLAS-LIFFEEEEDF SGGSAT TI PGTTINNTSS LITST SCKAATFTATTN TETP PTTITNTTSS LVTSTIKEE LYDAS SCKAATFTATTN TETP PTTITNTTSS NTTSCDLSGENGSF SCKAATFTATTN TETP PTTITTTS NTTSCDLSGENGSF SCKAATFTATT NTETSTER NTTSCDLSGENGSF SCKARTFTATT TETP PTTITTTS OGGENGSF SCKARTFTATT TETP PTTITTTS NTTSCDLSGENGSF SCCRRTSFAAVA QOUP FASTL QOUP FASTL GCCRRTSFAAVA QOUP FASTL QOUP FASTL GCCRRTSPA QOUP FASTL QOUP FASTL GCCRRTSPA QOUP FASTL QOUP FASTL GCCRALSP QOUP FASTL QOUP FASTL GCCRALSP QOUP FASTL QOUP FASTL GCCRALSP QOUP FASTL QOUP FASTL GCGVA-HAGPPPL <	SF SS SF
CpWRKY-IIE2a GmWRKY-IIE2d GmWRKY-IIE2d GmWRKY-IIE2b GmWRKY-IIE2b GmWRKY-IIE2b GmWRKY-IIE2C GmWRKY-IIE2C GmWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SbWRKY-IIE2C SBWRKY-IIE2C SmWRY-IIE2C SmWRY-IIE2C SmWRY-IIE2C SmWRY-IIE2C SmWRY-IIE2C	ME E Y UCME EWD LOA I V RAG MDE FACL R PWD LEA I V RG MDE FACL R PWD LEA I V RG MSE E PWD LEA I V RG MSE DWD LEA I V RS MSE DWD LFAI V RS MAE DWD LFAI V RS MAE DWD LFAI V RS MT HE DWD LFAI V RS ME E E LG ME E E LG	GCTN FFACGOF KVPPSSGFGLOHHEL S TGEAIT MDDPNDFSYFF N TGEATT MDDPNDFSYFF SSSSSTTNS CAGHEDDLGNCKQQ ODPPPPFLFAASS SGSAT TTIPQTTINNTSS CLASSITEELYDAF SCKAATFTATN TTIPQTTINNTSS LITST SKKAATFTATN TTIPQTTINTSSS LUTSTKEELYDAF SKKAATFTATN TETPTTTTTTS NTTSCDLSGLSGLNGSF SCKAATFTATN TETPTTPAT NTTSCDLSGLNGSF SCKAATFTATN TETPTTPAT NTTSCDLSGLNGSF GCCRRLSFAAVA OQUVPFASFLPP PREKKAEVUHL GCCRRLSFP OQUVPFASFLPP PREKKAEVUHL GCCRRLSPPG QPDFASFLPP PPSPFHKEKFVVFAFA AS CSGGGS-RVTTSPL SPAKEKFVVAFAFA AGSVA HERA AAFAPAFAFAFAFAFAFAFAFA GGSVA HERA AAFAPAFAFAFAFAFAFAFAFAFAFAFAFA CGSVA HERA AAFAPAFAFAFAFAFAFAFAFAFAFAFAFAFAFAFAFAF	S F S F S F
CpWRKY-IIE28 GmWRKY-IIE26 GmWRKY-IIE26 GmWRKY-IIE26 GmWRKY-IIE26 GmWRKY-IIE26 GmWRKY-IIE26 GmWRKY-IIE26 GmWRKY-IIE26 GmWRKY-IIE26 SbWRKY-IIE26 SbWRKY-IIE26 ZmWRKY-IIE26 ZmWRKY-IIE26 ZmWRKY-IIE26 ZmWRKY-IIE26 ZmWRKY-IIE27 SWRKY-IIE26 SWRKY-IIE26 SWRKY-IIE26 ZmWRKY-IIE27 SmWRKY-IIE26 SWRKY-IIE26	MEE EYUCMEEWOU GAIVRA MOE FACLROWOU EAV VRG MEE FACLROWOU EAV VRG MSE FOUDU SAV VRG MSE POWOLFAV VRG MAE DWOLFAV VRG MAE DWOLFAV VRG MAE DWOLFAV VRG MT HE DWOLFAV VRG MT HE DWOLFAV VRG MT HE DWOLFAV VRG MT HE DWOLFAV VRG MEE ELVG MEE ELVG MEE ENVG MEE ENVG MEE ENVG MEE ENVG MEE ENVG MEE EV MEE ENVG MEE EV MEE	GCTN FFACGOF KVP PSSCFGLOHHEL S TGEATT MDD NADFSYFF N TGEATT MDD NADFSYFF SCSSSVSTTNS CACHEDD LONCKQQ ODP PPPL FOR STAFS SCGSAT TT TI PGTTINNTSS LITST SCAATFTATT N TT SCS LISSEN LITST SCAATFTATT TT I PGTTINNTSS LITST SCKAATFTATT TT PT PT TTTTTS NTT SCS LISGENGSF SCKAATFTATT N TE PPT TAF NTT SCS LISGENGSF SCKAATFTATT N TE PPT TAF NTT SCS LISGENGSF CCRRATSFAAAVA GOOD PASSIL GCRRATSFAAVA GCCRRATSFAAVA GOOD PASSIL GCRRATSFAAVA GCCRRATSFAAVA GOOD PASSIL GCRRATSFAAVA GCCRATSFAAATATATAFAAP PAREFY PAREFY GCGVA HEFAA AAP RATV MDN HLIFSOHFC GCCRRATSFAAVA GOOD PASSIL GOOD PASEL PAREFY GCCRATSFAAVA GOOD PASEL PAREFY PAREFY GCCRATSFAAVA GOOD PASEL PAREFY PAREFY GCGVA HAGPP PAREFY PAREFY PAREFY GCGVA <t< th=""><th>S F S F S F S F S F</th></t<>	S F S F S F S F S F

IIE2-

IIE1-
				motif E								DPFxxxxDPF	
	, · ·	0	20	30	40	50	60	70	80	90	100	110	
GmWRKY-IIE3a GmWRKY-IIE3b	MCN-IL <mark>g</mark> ti	MENYQ <mark>G</mark> -		- DLSDIIRAS	S <mark>G</mark> ASY <mark>G</mark> ·····	scs sннннн.sssн	Т <mark>G</mark>	TSSSEAAN <mark>P</mark> F NDHWHHHOHF	SRNYHLQFSS SSSSSD <mark>P</mark> ···	D <mark>P</mark> MIFSSVLI - PITFSSSS	E <mark>G</mark> F NSNF SVL I ODTNF	GDPFSNMRD	
VVWRKY-IIE3a		MENYQG-			SGSGGA ·····	R	T	SDE <mark>P</mark> VADWQF	PSE	PMIFSPVV			
AtWRKY-IIE3b	MCS - VSELLI	MENFQG		DLTDVVRG			T	PPSNIWPLPL	SH <mark>PTPSP</mark> SDL	NI · · · · · · ·	• • • • • • • • • • • • • • • • • • •	GDP FVSMDD	
SBWRKY-IIE3a ZmWRKY-IIE3a	MCDYFLPRMI	E - GDQAGG E - GDQAGG	G	- DLTDIVRS	3G - A I PGNAAEM 3G - A I <mark>PG</mark> NAADT	5 S Т ААА <mark>Р</mark> S Т ААА	DEWQL	QGDP - ML FPP QGDATML FPP	LPSST-TSEA L <mark>P</mark> SST-TASE	VGCAA	VFGTD-	···PFSGLVD····	
ZmWRKY-IIE3b OsWRKY-IIE3a	MCDFFL <mark>P</mark> RMI <mark>P</mark> AAWA <mark>G</mark> ERMI	E - GDQAGG E - <mark>G</mark> DQA <mark>G</mark> G	<mark>G</mark> .	··· DLTDIVRSO ··· DLTDIVRSO	3 G - AVP S SAAE - 3 G G AM <mark>P G</mark> S V V V D	······ <mark>LP</mark> STA	DEWQLH AEWQL <u>-</u>	QGEPPMLFPP - <mark>P</mark> AE <mark>P</mark> MLF <mark>PP</mark>	L	GCGAGGAA	3 A N V F G A D - 3 A D I F <mark>G G G G</mark>	- DPFSGLVD GDLFSGLVD	IIES
OsWRKY-IIE3b BdWRKY-IIE3a	MCDYFLQRMI MCDYFLHRLI	E - <mark>G</mark> EQAA <mark>G</mark> EA <mark>G</mark> EQQ <mark>P</mark> -		··DLADIVLR/ Agdltdi·LR/	A <mark>GG</mark> A - AAAAVA <mark>G</mark> AGG <mark>AMPP</mark> VDAAD	<mark>GG P</mark> S L <mark>P</mark> STA	TEWQL <mark>P</mark> T <u>E</u> WRL	<mark>P</mark> AEEEEEE <mark>PG</mark> QS <mark>P</mark> TRLFA <mark>P</mark> T	L F <mark>P</mark> L <mark>P P</mark> S S S D I Q S S V A S D <mark>G</mark> -	····A	SADAF <mark>G</mark> ··· S <mark>P</mark> SAF····	- D <mark>P</mark> F A G L <mark>P</mark> D - D <mark>P</mark> F S G L Q D	
SbWRKY-IIE3b ZmWRKY-IIE3c	MCDYFLQRMI MCDYFLQRMI	E - <mark>G</mark> DQHH - E DDRHHHH	0 A	A <mark>gdl</mark> tdv-vr/ Agdltdv-vr/	A <mark>gg</mark> amhqq <mark>g</mark> aia Aggamh <mark>p</mark> - Aava	E L S S S T A E L S S T A T	T <mark>G</mark> WQL T <mark>G</mark> WQL	PAE <mark>PA</mark> PAGPG PAEPA - GPPG	L F L <mark>P P</mark> Q <mark>P</mark> S S S L F <mark>P P</mark> A R S S S S	D	3 G D <mark>G</mark> F A D <mark>G</mark>	- DAFA <mark>GLP</mark> D - <mark>G</mark> AFAGLSG	
ZmWRKY-IIE3d GmWRKY-IIE3c	* MCDYFLQRMI	OQ <mark>G</mark> DHHHH MDNYQ <mark>G</mark> -	н	A <mark>gdltdv</mark> - vr/	A <mark>GG</mark> AMHQ - AAVA SGASY <mark>G</mark>	E L <mark>P</mark> S A D - T C S	T <mark>G</mark> WQL T <mark>G</mark>	PAE <mark>P</mark> A - AGAC T <mark>P</mark> SSEAANPF:	LF <mark>PPP</mark> Q <mark>P</mark> SSS SR <mark>G</mark> - HLQFSS	D <mark>P</mark> MIFSCVLI	C - D <mark>G</mark> FA E <mark>G</mark> F NSNF	- DAFAGLPD GDPFSNMRD	

Appendix 6	continued
------------	-----------

	120	130	140	150	160	170	180	190	200	210	220	230
AtWRKY-IIE1a								MKR <mark>GL</mark> D				
VvWRKY-IIE1a								<mark>MDRRL</mark> H				
VvWRKY-IIE1b								MD <mark>G</mark> RFN				
OsWRKY-IIE1a								MD <mark>G</mark> EWSD <mark>G</mark>	AAV			
BdWRKY-IIE1a								MD <mark>G</mark> EWSEG	AVS			
SbWRKY-IIE12								MDAEWSDG	AAA			
ZmWRKY-IIE1a								MDAEWS				
GmWRKY-IIE1a						M		· · HHRRFS · ·				
GINWRKY-NETR								VKPPENT				
MINNER Y-IIE 12												
GmWRKA HEAd								MOSKERKN	NRN			
MWRKY.UE1h								MKRKFK	• • N • • • • • •			
AtWRKY-IIE1b								MHRRAA				
OsWRKY-IIE1b					MAAGEEV	MDR		<mark>S</mark> T SAE DG Y	c s A g			
SbWRKY-IIE1b								<mark>M</mark> SSAED <mark>G</mark> Y	css			
ZmWRKY-IIE1b								<mark>M</mark> SS <mark>G</mark> ED <mark>G</mark> Y	css			
ZmWRKY-HE1c								<mark>M</mark> SSAED <mark>G</mark> Y	css			
ShWRKY-IIE2a		HVMS	RAA <mark>G</mark> QI	MAIPPA - CYDE	EQELRPAAA	CGDAVMFDVP	SSMV VDPY	HOLSSSAATA	<mark>РРНG</mark>			
ZMWRKY-NE2a	THERE	·····HIVS	AARATARSQI	ML VPAAACYDE			SSSMAAAVGP Y		AP-A			
ALWRIN T-HE 23	n n n n n n n n n n n n n n n n n n n	N N N N N N N N N N N N N N N N N N N	ODADDD V	NIS <mark>P</mark> VLSLD <mark>P</mark> R		VLK····VKT	AULQEITRALF					
GeWRKT-HEZH	••••••	FEVGOR	··· VPSSAA		EYLDL				рете		<mark>.</mark>	POFOCEUE.
SAWRKY-HE2C		VATAAAGOE.		VAAWRRNI					ст <mark>р</mark> ея			ETTSOGAP.
OsWRKY-IIF2d			· · · FEGDDGD	AAWMAPLPGLO	TOGOW				CGALVVAP			
ZmWRKY-IJE2c			· · · GPEPEPY		· · · ·				LAVOPSLPOI	LPTTTPT	Г <mark>Р</mark> ТЦ	
ShWRKY-IIE2d			WASDESFS	- CLPPPPQS			к	DEVTTDAAAL	- QQPLISPA		• • • • • • • • • • • • • •	
ZmWRKY-IIE2b			RASDESFS	- CLPLPPPP-	· · · · · • • · · ·		к	DD <mark>V</mark> TDA <mark>T</mark> TAL	SQQLPIGPA			
ZmWRKY-IIE2d			КЕ <mark>Р</mark> Е <mark>Р</mark> Н	TAWR F <mark>P</mark> DLGA-	<mark>GGG</mark>			Q D <mark>G</mark> E E L <mark>L</mark> R A L	LADQPPLHQ	LPTPAATP	Г <mark>Р</mark> Т <mark>Р</mark> Т АТ <mark>Р</mark>	
AtWRKY-IIE2b			vs sht	N <mark>P</mark> I F T V <mark>G</mark> RQSN	AVSF <mark>G</mark>			- E <mark>I</mark> R D L <mark>Y</mark> T <mark>P</mark> F	т	ES		
VvWRKY-IIE2b			<mark>P</mark> L A I Q Q	DELLFSF <mark>P</mark> DFF	ETTTVL···			DE <mark>L</mark> EQL <mark>Y</mark> K <mark>P</mark> F	Y <mark>P</mark> VLQ <mark>P</mark> LS-			
VvWRKY-IIE2a			FDVVEQAHHH	D <mark>g</mark> hllcf <mark>p</mark> d <mark>p</mark> f	ETRREAF		v	E E <mark>l</mark> H D L <mark>C R P</mark> F	FPKS-QPIS	RTIPI		
GmWRKY-IIE2a		· · · · · · · · · · · · <u>·</u>	F S V Y N <mark>P</mark> A E Q <mark>G</mark>	<mark>G</mark> HVLSLSEN <mark>P</mark> F	EARSSNS		· · · · · · · · · · I	E <mark>gl</mark> hel <mark>c</mark> k <mark>p</mark> f	FLK <mark>P</mark> - Q <mark>P</mark> QTI	QTS		
GmWRKY-IIE2b		<mark>G</mark>	FNIFK <mark>g</mark> eq <mark>g</mark> i	SQVLSLSAY <mark>P</mark> F	EAR\$\$			E E L H E L <mark>C</mark> K <mark>P</mark> F	FSKS - Q <mark>P</mark> LTI	QAS		
GmWRKY-IIE2¢		<mark>G</mark>	FN I FK <mark>G</mark> EQ <mark>G</mark> I	SQALSLSAY <mark>P</mark> F	EARSS			E E L H E L C K <mark>P</mark> F	FSKS - Q <mark>P</mark> LTI	QAS		
MtWRKY-IIE2a			FSSEQK <mark>P</mark> QIL	5 L S C S T L <mark>P</mark> F	'EARSS		· · · · · · · · · · · · · · · · · · ·	E E L H E L C K P F	FSRSSQPLSI	. Q - T		
AtWRKY-IIE2¢			<mark>P</mark> S F C L <mark>P</mark>	METSSFYE <mark>P</mark> EN	IE T S <mark>G</mark> L			DE L <mark>G</mark> E L <mark>Y</mark> K <mark>P</mark> F	Y <mark>P</mark> FSTQTI-			
CpWRKY-IIE2a			····RSMDEP	EDLLISSEIFC	ESREVAS			DEL <mark>G</mark> QLYK <mark>P</mark> F	YPVLHPLS-			
GmWRKY-NE2d			SEQ	DELLDSFPEFS	ETTRVL				YPV-LHPL-			
GMWRR T-HEZe			CONTRODED	VELLUSF <mark>P</mark> EFS	TTWODDO				TPV-FHPL-			
AIWRA T-HE20	B		· SCNELQDSC	SPSMS PCCELC	NECPREI		PEALDDPLVCC	SAPASEDEUR				
GmWRKY-IIE21				OI FT PENPT	TNTTNT	SASGIN	PNS	PYFAFOFSOO	l	FHIPI		
GmWRKY-IJE2#	P	· · · · · · · N I VP	NTNEFOELH	QLETPENPT	NNT	SAPG IN	PNS	PYFAEQESQO		EHLHI		
GmWRKY-IIE2h	P	NLV0	TTNGFOFLO	HOOLLINENPT	NT T ST	· · · · STS · · T	T D T	AK F0000		PFQ		· · · · · · · · · · · · · · ·
GmWRKY-IIE2i	P	NLVQ	TTNGFQELQ	HQQLVINFNPT	NTTSTTS	· · · · STSTIT	SGLGINPNSTF	SEVAGEIGOO		PTNH····		
GmWRKY-IIE2i	. .		P EEQ	DDLLCNFQEFS	ETTTVV			DELEELYKPS	YPRHVH			· · · · · · · · · · · · · · · · · · ·
GmWRKY-IIE2k	*		<mark>P</mark> EEQ	DDLLFNFQEFS	ETTTVV			DE LEEL <mark>Y</mark> K <mark>P</mark> F	Y <mark>PP</mark> HVHVD-			
'OsWRKY-IIE2e ³	k		• • • EKE <mark>V</mark> GVE	AAWS F <mark>P</mark> ELTVF	DGGGGGCL		G	R D <mark>A</mark> D E L <mark>L</mark> K A F	CAAFPSSSS	6 K S S <mark>P</mark> L <mark>P</mark> T <mark>P</mark> I	<mark>РРРР</mark>	
BdWRKY-IIE2a	k		EKQ <mark>G</mark> AD	A <mark>G</mark> WC F <mark>P</mark> DL <mark>G</mark> AS	DRA <mark>GCG</mark>			E D <mark>A</mark> NE L <mark>L</mark> KAF	FAPPPPPTI	K P L P T Q Q Q	a <mark>a</mark>	
SbWRKY-IIE2e [*]	•		•••AKE <mark>P</mark> E <mark>P</mark> Y	AAWR F <mark>P</mark> DL <mark>G</mark> A -	AGGG			Q D <mark>G</mark> D E L <mark>L</mark> R A L	LAA <mark>PPPPP</mark> -	· - <mark>P</mark> Q <mark>P</mark> L <mark>P</mark> T <mark>P</mark>	T <mark>P</mark> L <mark>P</mark> • • • • • • • •	
SbWRKY-IIEXa	• • • • • • • • •		•••••••	· · · · · <mark>P</mark> SLN <mark>P</mark> ·				· · · · SRNSEL	AA <mark>PPP</mark> KS <mark>P</mark> -			
ZmWRKY-IIE2e	* *		RASDE <mark>P</mark> FS	SCLL <mark>PPP</mark> QSQ -			····к	DEVTDAAAP -	- QQL <mark>P</mark> TGPA	/ <mark>G</mark>	· · · · · · <mark>·</mark> · · · · ·	
ZmWRKY-IIE2f	~ *	· · · VSARRQE ·	· · · RA <mark>PPP</mark> · ·	VALLCDL	EYLDS			- DHKPFLLPV	A R <mark>P</mark> SI	R	· · · · · · <mark>P</mark> · · · · I	RVGGGGH <mark>G</mark> S -
ZmWRKY-IIE2g	*	···VPVRRQG	· · · RASPP · ·		EYLDL				APSP SPS	G	••••• <mark>P</mark> •••• I	RAGAGGR - ER
ZmWRKY-IIE2h	•••••	AAAQLE-	· · · LLGQP · ·	v <mark>u</mark> a v s s wr R <mark>G</mark> L	UNFDDLV		· · · · D	RURDPRRMPF	SVSV - T <mark>P</mark> S		•••••	CETTATCGPN
OsWRKY-IIE21	.	····		ALYDL	LTLUL		PNSP		5555SSS		AVARAI	
MALINER Y-UF 20				willior P			• • • • • • • • • • • • • • • • • • •					· · · · · · · · · · · · · · · · · · ·

IIE1-

IIE2-

	DPFxxxxI	PF											
	120	130	140	1	50	160	170	180	190	200	210	220	230
GmWRKY-IIE3a	P F L H E L D - L	PLSAYF	I-STSSSAI	IKSS <mark>g</mark> ale	EATCFGGG	vva <mark>g</mark> ssss	S-NSCVLAC	KILDEDDMR	R P C N S I L S N M	IQIS <mark>P</mark> NDK			
GmWRKY-IIE3b	PFLQELD-M	P SASAY	· - T <mark>GG</mark> ALI	LEAAAT	F G S	LEEH····		DNNNMR	R <mark>P</mark> CKN I FSNM	IQIS <mark>P</mark> NAK ·			
VvWRKY-IIE3	PLLHEL <mark>g</mark> ·I	AGSGFFC			F <mark>G</mark> QP	(ILDE		BMK	R <mark>P</mark> C - N I F S RM	LQIS <mark>P</mark> NAR ·			
AtWRKY-IIE32	P L I H · · · · ·	LPASYIS	G <mark>ag</mark> dnksi	I KS <u>F</u> AI	F <mark>P</mark> KIFEDE	нік		SQCSVF	P R I K I <mark>S</mark> QSNN	IHDASTCN			
AtWRKY-IIE3b	P L L Q E L N S I	TNSGYFS	T V <mark>G</mark> DNNN	I I HNNN <mark>g</mark> F L	VPKVFEED	н к		SQCS F	P R I R I <mark>S</mark> HSN I	IHDSS <mark>P</mark> CN			
ShWRKY-IIE3a		P F S T D Y ·	S S <mark>G</mark> A D F L I	DA - M <mark>P</mark> DAM -	AKVGF - DT	「 - A I C <mark>G G</mark> S <mark>G</mark>	SGCGGGGA	<mark>BGGG</mark> QLIDMS	RKQ - PLLPRG	- VQM <mark>P</mark> AL <mark>G</mark> - VI	. A <mark>P</mark> RMVL <mark>P</mark> S <mark>P</mark> L	s s	
ZmWRKY-IIE3a		P F S T D Y -	S S <mark>G</mark> A D F L I	DA - M <mark>P</mark> DAM -	AKVGF - DT	TAIC <mark>gg</mark> sg	SGSGCRG - C	<mark>BGGG</mark> QLIDMS	RKQQ <mark>PLL</mark> PRG	- LQM <mark>P</mark> AV <mark>GG</mark> VI	. T <mark>P</mark> R V V L <mark>P</mark> S <mark>P</mark> L	s	
ZmWRKY-IIE3b		<mark>P</mark> FCT D Y ו	SSCADFLI	DAT M <mark>P</mark> DAM -	PKVGF - DA	A - A I C <mark>GG</mark> S <mark>G</mark>	secese - ·	G Q G Q L I D V R	RKQ - <mark>P L L P</mark> R <mark>G</mark>	- VQML <mark>P</mark> VG - AI	.A		
OsWRKY-IIE3a		PFSSDY	S S G A D F L I	DA - M <mark>P</mark> DAM -	AKVGF - D1	- AV <mark>GGGC</mark> G	GGGGGGGG - C	<mark>S G G H L</mark> L D M S	RK PLLPRG	- M <mark>P</mark> MAAV <mark>GG</mark> L/	AA <mark>P</mark> R - VM <mark>P</mark> S <mark>P</mark> L	s	
OsWRKY-IIE3b	· · · · · · <u>·</u> · ·	P F G G D Y F	S S <mark>GG</mark> AAA/	ADFFDAV	AKA <mark>G</mark> FVD	GVL GGGGG	GGCDGGGVI	D <mark>GGGGG</mark> SSLL	GMSKPILPRA/	A M Q L <mark>P</mark> S V S ·			
BdWRKY-IIE3a	<mark>P</mark> F S	SCSTDY	SSS <mark>G</mark> S - A/	ADFFDALA	HDA MDA	4KV <mark>G</mark> VG	NYVE PAGA1	GGGA <mark>G</mark>	GPLDMRNH	н <mark>нмм</mark> р			
SbWRKY-HE3b		P	/RASSSS <mark>G</mark> (9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	- AADFFDI	E A PAAAVG	IGG ARF	R <mark>GGGG</mark> VLVDS	GGGGVVVERG	V <mark>PQMP</mark> ALS ·			
ZmWRKY-IIE3c		P FASDF ·	- RADSAG	/ D <mark>P</mark> S <mark>P</mark>	- A - DFFDI	E A P A V G G G	RG GM	/E <mark>GGGP</mark> · · · ·		- <mark>P</mark> QM <mark>P</mark> ALS ·			
ZmWRKY-IIE3d		P FASDF ·	RTSSSSG	/ <mark>G</mark> L C <mark>P</mark>	- ADDFFDI	E A PAAAAA	A <mark>GG</mark> <mark>G</mark> AF	R R <mark>G G G </mark> V L A <mark>G</mark> A		Q M <mark>P</mark> A L S			
GmWRKY-IIE3c	PFLHELD-L	PPSAYFN	IITTSSSAE	MSSS <mark>G</mark> ALE	EATCFGGG	V V AG S S S S	SSNSCVLAC	KILEDDDMR	R <mark>P</mark> CNSIFSNMI	1Q S <mark>P</mark> NDK			

IIE3-

							basic-2		GExxxP	
	240 250	260 27	70 280	290	300 310	320	330	¥340	350	
AtWRKY-IIE1a	TT , TT	MARSYNDHE SSQE -			ST FNGMKALISS	-HSPKR-			GSRHKG D -	
VvWRKY-IIE1a		IN <mark>P</mark> FINEQEED <mark>P</mark> Q <mark>G</mark> ··	SSETA <mark>P</mark> DS		• • • • • • <mark>P</mark> LS <mark>G</mark> DEAKLAAT •	- <mark>P</mark> SPKK-	sk <mark>rg</mark> va <mark>kr</mark> v	VSV <mark>P</mark> IGGVD	GSRSK <mark>G</mark> - · E ·	
VvWRKY-IIE1b	•••••	SNPFVSEQEEQENSP-	• • • ENS <mark>G</mark> DS <mark>PP</mark> S • • • •		····· <mark>P</mark> MFNDAKISSST ·	- SS <mark>PKR</mark> -	SR <mark>R</mark> AMQKRV	VSV <mark>P</mark> IKDVE	GSRVKGD-	
OSWRKY-IIE1a BdWRKY-IIE1a		SSPTMSGGGGREQMKG	GEDVAAADCPGSP · · · ·		· · · · VS <mark>P</mark> SPAAAQRSAA <mark>G</mark> A · · · · · · · APSTTS · · · · ·	AASPSG-	· · · · · · · · · · · · · · · · · · ·	VT VPLADVT	GPRPKGV-GE	
ShWRKY-IIE1a		ASPPTVSGGESKPGAA	GAVSSSADCPGSPP		VS <mark>P</mark> A <mark>PP</mark> STTS <mark>P</mark> AAAA	AAAGSG-	RRSANKRV	VTVPLADVS	GPRPKGV-GE	
ZmWRKY-IIE1a	•••••	· · · · · · · · GDESKAGAA	GGVSS - ADC <mark>P</mark> SS <mark>PPP</mark> -		· · · · VS <mark>P</mark> A <mark>PP</mark> STS <mark>P</mark> · · · · ·	AAT GAG -	· · · R <mark>R</mark> SAN <mark>KR</mark> V	VT V <mark>P</mark> LADVS	G <mark>P</mark> R <mark>P</mark> KGV - GE	
GmWRKY-IIE1a GmWRKY-IIE1a		KLIHVADPDEPESGP	EPASEAGPAS		·····PSSGEDTKTEA··	- PSPKK-	RREMKKRV	T I P I G D V D	GSKSKG E -	
MWRKY-IIE1a		KPLYVADREEPEPELK	· LEPILEDGPAS · · · ·		····· PSSCEDTKIEE ···	- SSPKK-	RREMKKRV	TIPIADVE	GSKSRG E -	IIE1-
GmWRKY-IIE1¢	Sү	TSEQDQDTD IAQEN	• • • • LAES <mark>PP</mark> SS • • • •		· · · · · T V F N · I D <mark>G</mark> L V <mark>P</mark> S · ·	PTSSS <mark>K</mark> -	RR <mark>R</mark> A I Q <mark>KR</mark> V	VQI <mark>P</mark> MKETE	GCRLK <mark>G</mark> ES	
GmWRKY-IIE1d		TSEQDQDTDHDVAQEN	····IAESPPSS····		····TVFN·IDGLVPS··	PTPSSK-	RRRAIQKRV		GCRLKG ES	
AtWRKY-IIE1b	·····	··IQESDDEEDETYN-	DVVPES			PTPKK-	SRRNVEKRV	VSVPIADVE	GSKSRG E -	
OsWRKY-IIE1b	т	DSPRAESVDEQ	G AAEESS <mark>P</mark> RG		GQKRELPSPSASPSSPL	PPAAKR-	SR <mark>R</mark> SVE <mark>KR</mark> V	VSV <mark>P</mark> IAECG	DR <mark>P</mark> K <mark>G</mark> A GE	
SbWRKY-IIE1b	•••••	DSPRAESPDEPLLPAA	V - ADADAE S <mark>P</mark> RAA	GSGM	- NKRERDLSDL <mark>P</mark> AS <mark>P</mark> SS <mark>P</mark> L	PPA · KR ·	SR <mark>R</mark> SVE <mark>KR</mark> V	V S V P L A E C G	DR <mark>P</mark> RGAT - GE	
ZmWRKY-IIE16 ZmWRKY-IIE16		DSPRAESPDQPQAAAA	ADADADAESPRA	GAGP · · · · · · · · · · ·	-NKRDLHAS	PPA.KP.	SRRSVEKRV	VSVPLAEDG		
		DOT KAEOT DET Q						DUL LALO		
SBWRKY-IIE2a ZmWRKY-IIE2a		PQQISQQACYGLDVAM	GGAAAADADGDE		PMMMRISPVT - PPPPSHH PVVTRISPVTTPPPPSHH			VCIPALPP -	ASSRPG GG	
AtWRKY-IIE2a		SYGLKVLAYRSLIRVS	SNGE I VMHYLLRQMGK	EILHTESKKTDK	LVDNIQSSMIATKEIEITF	SKSRRKN	NKEKRV	vevv	DR	
SBWRKY-IIE28	E · · · · · · · · · · · VM I S F <mark>P</mark> A	GAASTSGMQ	QRAS <mark>PPG</mark> RK <mark>PG</mark> ART <mark>P</mark> -			· · R <mark>PKR</mark> · ·	· · sk <mark>k</mark> s <mark>a</mark> lkkv	v R	EM <mark>P</mark> VA - D <mark>G</mark>	
OsWRKY-IIE2c	VL I S F <mark>P</mark> A	IAST SGQGRK	Q PGRKPGVRTA -		• • • • • • • • • • • • • • • • • • • •	- RPKR	SKKSQLKKV	vc	EVPVA - DG	
OsWRKY-IIE2d	·····	VLQVQQQPPPADNTQP	TTYQQGSGGDGE			RAGGSR		vv	RVAA	
ZmWRKY-IIE2¢	·····PPAAGEHR	Q S T V T A V D V P L P Q V R P	APAA RAQPSGRQVP	. <mark>a</mark>		PRSKRR-	<mark>ĸ</mark> NQVKKV	vc	<mark>P</mark> A	
SbWRKY-IIE2d		QQAFLA - AT <mark>P</mark> Q <mark>P</mark> RSEA	PPPQPPAKPRTSYRNN	D <mark>GG</mark> V <mark>G</mark> ·····		T R S <mark>K R K</mark> K	K KSQVTS <mark>K</mark> E	v т	····RV·· <mark>P</mark> V	
ZmWRKY-IIE26 ZmWRKY-IIE2d	DDSC-L	QQAFF AT PQS RNEA	PAPPLQPPAKLRI- APA RAOPSGROVP	. <mark>a</mark>		PRSKRK	KSKKE	VK	RVPV	
AtWRKY-IIE2b		SCINYPEEPRKPQNQK	R <mark>P</mark> LSLSASS <mark>G</mark> SVTSK <mark>P</mark>	s <mark>g</mark>		SRSKR	R K I QH <mark>K</mark> K	vc	HVAA	
VvWRKY-IIE2b	••••• <mark>P</mark> VRTAAA <mark>P</mark>	SISMED <mark>P</mark> KK - MK <mark>G</mark> QDH	L <mark>PGTADNGGGGP</mark> - <mark>G</mark> SH		• • • • • • • • • • • • • • • • • • •	AK <mark>PKK</mark> SR	· · · · · <mark>KNQ</mark> Q <mark>KRV</mark>	v q	н v <mark>т</mark> а	
VvWRKY-IIE2a	SSLSV	L GGG FSDQTHQ I QQQQ	KQD <mark>PPP</mark> SKH <mark>S</mark> HASSVP Fok - Osobo Fullagsa	STTHS	QS	PRSKRR-		vc	PA	
GmWRKY-IIE2b	·····sPLSS	LSSYSSA <mark>PP</mark> KSVSTQE	KQQ - QRSKQAH AV	тт		PRSKRR-	····KNQLKK	vc	QV PV	
GmWRKY-IIE2c	· · · · · · · · · · · · · · · S <mark>P</mark> L S S	L S S Y S S A <mark>P P</mark> K S V S T Q E	KQQ - QRSKQ <mark>P</mark> H AV	тт		PRSKRR-	···· <mark>K</mark> N <mark>Q</mark> LKK·	vc	QV <mark>P</mark> V	
MtWRKY-IIE2a	••••••••••••••••••••••••••••••••••••••	SLSYSSAS <mark>P</mark> RLAQTQD	KQQLQRSKH <mark>S</mark> HQ <mark>GG</mark> SL	TN		PRSKRR-		vc	····QV··PV	
CpWRKY-IIE2a		SMSLPEEFK - HPQNGQ	KKKLLQPLPSG SAH			PKAKRSK	KT Q Q KKV	vĸ	QVTE	
GmWRKY-IIE2d	s	PHT I KVAALQ <mark>g</mark> l KVPA	A <mark>P</mark> K····FQSLD···			• KS <mark>KKR</mark> D	K <mark>K</mark> SQNKSV	<mark>v</mark> к	· · · · QV · · TT	
GmWRKY-IIE2e	·····S	PHT IVTTSPPIPIEPE	QPFMYNIFFQSGD		• • • • • • • • • • • • • • • • • • • •	- VSSA - N	KKSONKSV	VK	QVTT	
CpWRKY-IIE20	TPLSSVSTSCSPSRDRRQ	NAPLLSPSQLQQFQLH	DPELPHQQQQQAHLHK	P		- RCRKR-	····KNQQKRT	vc	····HV··TA	IIE2-
GmWRKY-IIE21	HSST <mark>P</mark> SFSRFHDHQQQQQ	QINQLQALQKHEFQL	QNN - S <mark>P</mark> T VS <mark>P</mark> NAQ <mark>P</mark> QT	P		- KS <mark>RKR</mark> Y	a <mark>Ksa</mark> akkw	v c	нта	
GmWRKY-IIE2g	· · · · · · · · · · · · · · · · · · ·	QINQLQALQKHEFRLP	QNI-S <mark>P</mark> TVS <mark>P</mark> NAQ <mark>P</mark> QT	P		-KSRKR-	KSQQKKM	vc	····HV··TA	
GmWRKY-IIE2h GmWRKY-IIE2i	PTTTGFDRFOHH000010		QI SHILSPTTQPQT QT SPILSPTTOPOT	P		- RSRKR-	····KSHOKKM	vc	HVTA	
GmWRKY-IIE2j	DN	PLPVVANSPPILDEKV	KELK <mark>P</mark> SDKT <mark>A</mark> S			RCKKSK	KRQKNKRV	v	та	
GmWRKY-IIF2k ²	*NNN	PLPIVANSLPIPDEEV	KELK <mark>P</mark> SHKAAS	•••		- RCKKSK		v	· · · · · · · · · · · · · · · · · · ·	
JSWRKY-IIE2e 3	* · · · · · · · · · · · · · · · QKQKQEML	QPLMIPSEAQAPLRSP	TAAA <mark>P</mark> ARAQ <mark>P</mark> SGRPAS	. g		PRSKRR	KNOCKKV	vc	····HV··PA	
SbWRKY-IIE2e *		Q <mark>P</mark> AVAAVDV <mark>P</mark> LPQARP	A <mark>P</mark> A · · · RAQ <mark>P</mark> SGRQVP	. <mark>6</mark>		PRSKRR-	<mark>KNQVKK</mark> V	vc	нv <mark>р</mark> а	
SbWRKY-IIEXa	······································	PPPTRTANNPVARYAP	PPP · · · AQPSYAFTVL	A <mark>G</mark>	••••••••••••••••••••••••••••••••••••••		····TTKESKV	VL	VL-AE	
ZmWRKY-IIE2e	*••••VDDLC•L *••••VDDLC•L	QQAFLGLATPQPRNNE	APP - PAKRPRTPYRSN ORASSPRRKPGARTP -	N <mark>G</mark> V	GGF	I RSKRKK	KSQTSKKE	VA	RV PV FMPVA - DG	
ZmWRKY-IIE2g	*E VM S F <mark>P</mark> A	AAAA · · · · · · STSGTQ	QRASPPGRKPGARTS -			· R <mark>PKR</mark> ···	SKKSQLKKV	v <mark>R</mark> .	EMPAA - DG	
ZmWRKY-HE2h	*ARA <mark>G</mark> QQLRHDDVLFSFSA	AAAA <mark>G</mark> VVSRS <mark>G</mark> HQMVQ	PRRRQ <mark>PG</mark> RK <mark>A</mark> GGRT P -			· · R <mark>PKR</mark> · ·	· · sk <mark>kro</mark> vkkv	vc	E V <mark>P</mark> AA - <mark>GG</mark>	
DSWRKY-IIE2f *	Ε VM S F <mark>P</mark> A Ε T S T S S F N K F H N - 0 0 0 0		PR - 5 P 5 G RK PG RT P -			- RPKRRS		VY	EV <mark>P</mark> VA-DG	
alwarn 1-ne20			40.4041			1				

	RxxxTG	
	350	
Gm/WRK///////S/S/CRATES/CONTRAMAPS/R/D/INFORMATS/CONTRACTION/CONT	- AANSRQ TG - AANSRP SG	
ATWRRY 11839 ·····SPAI··T··VSSAAVAAS PWGMINVNTTNSPRNCLLVDNNNN······TSSCSQVQIS·SS·PRN··LGIKRR····KSQAKKVVCIPAPA- ATWRRY 11838 ·····SPAM··SAHVVAAAAAAS PRGINVDT-NSPRNCLLVDQ·······TTFSSQIQIS·S·PRN··LGLKRR····KSQAKKVVCIPAPA-	- AMNSRS SG - AMNSRS SG	
<i>sbwrky-liesa</i> PRE I RPY PP - L A GOMVKLG I TAGQVAG CAIDA AVVGMQMS S · PRSA - GGI KRR KNGARKVVCI PAPT . <i>zmwrky-liesa</i> PRE I RPY PA - L A GOMVKLG I TAGQVAG CAIDA CAIDA AVVDMQMS S · PRSA - GGI KRR KNGARKVVCI PAPT .	- AAGGRP TG - AAGGRP TG	73_
Zmiwrky:iE3b PRE I O PYPA - F - ASDMVKL G I TAGQ VAG CAIDA AVVGMQMPS - PRA - GG I KRR KNOAKKVVC I PAPT OswRky:iE3b PRA I R PYP - I SAGDMMKL G I TAGQ AGG CAIDA AVAGMQUSS - PRS - GG I KRR KNOAKKVVC I PAPT		
USWRYT/11239	- AAGGRT 5G AAVPGRT TG - AAGGRT TG	
ZmWRYY/JE3c PREVRPYSVMM-GGDTVKIGVPAMMPGGLPVGPP-CAFDAIA	ASAGGRT TG SSAGGRT TG	
GmWRKY-IIE3¢ L <mark>P</mark> I S . P AVDAL S RALK <mark>P</mark> S SMV I . S GDNM I D S KT S I D	- A <mark>P</mark> NSRR S <mark>G</mark>	

Appendix 6 conti	nued
------------------	------

	WRKY DNA-binding domain					
GExxxP						
			420 430		450 40	50 470 FAE
VVWRKY-IE12 AYPPSDSWAWRKYGOKPIKGS-P	Y <mark>P - RGYYRCSSSKGCPARKQVER</mark>	RVDPTVLIVTYACDHN	- HPLPPTKHHHSTAPAAVT	ATAS - IAT PSP -		KF
VVWRKY-HE10 CAPPSDSWAWRKYGQKPIKGS-P OsWRKY-HE10 GNTPTDSWAWRKYGQKPIKGS-P	YP - RGYYRCSSSKGCPARKQVER FP - RAYYRCSSSKGCPARKQVER	RVDPTMLVVTYSCEHN RNDPDTVIVTYSFEHN	- HPWPAS RNHHNT TNPPTT - HSATVPRAONROAAPOKP	AAAT <mark>P</mark> TPTSAP -		· · · · · · · · · · · · AT
BAWRKY-IIE12 GNTPTDEWAWRKYGQKPIKGS - P	F <mark>PSRAYYRCSSSKGCPARKQ</mark> VER	QADPAMVL VTYSYEHN	- <mark>HST</mark> AVANRAQQNRQAQK <mark>P</mark>	KDQ <mark>P</mark> DH <mark>P</mark>		A
ShWRKY-IIE1a GNTPTDSWAWRKYGQKPIKGS - P ZmWRKY-IIE1a GNTPTDAWAWRKYGQKPIKGS - P	F P - RAYYRCSSSKGCPARKQVER F P - RAYYRCSSSKGCPARKQVER	RAEPDKVIVTYSFEHS RAEPDKVIVTYSFEHS	- HSDAVARAQQNRQQASKP - HSEAMAAARAQNRQAPKP	KAVQRQPVPPEP	A	AES
GmWRKY-IIE1# NYPPSDSWAWRKYGQKPIKGS-P	Y <mark>P - RGYYRCSSSKGCPARKQVER</mark>	RVDPTKLIV <mark>TY</mark> AYEHN	- <mark>HSLPLPK</mark> SN <mark>SS</mark> AASA	AVSDGATSSSP -		ADSAARY
GmWRKY-IIE12 TYPPSDSWAWRKYGQKPIKGS-P MtWRKY-IIE12 TYPPSDSWAWRKYGQKPIKGS-P	YP - RGYYRCSSSKGCPARKQVER YP - RGYYRCSSSKGCPARKQVER	RVDPTKLIVTYAYEHN RVDPTNLLVTYAHEHN	- HSLPFPKSNSS AASA - HSLPLPKSHHSSTNAVTA	AVSDGA-AVLP- ATVTAAVDT <mark>P</mark> SP-		I E S A A R F V E S <mark>P</mark> A A T
GmWRKY-IIE1¢ NTPPSDSWAWRKYGQKPIKGS-P	YP - RGYYRCSSSKGCPARKQVER	C V D P T ML V V T Y S D H N	• <mark>H P WP P S RN</mark> H • • • • • A R <mark>P</mark> T	KKPEPVPDPV		
GmWRKY-IIE1# NT PPS DSWAWRKYGOKP IKGS - P IMWRKY-IIE1b NT PPS DSWAWRKYGOKP IKGS - P	YP - RGYYRCSSSRGCPARKOVERS YP - RAYYRCSSCKGCPARKOVERS	CVDPTML VVTYSSDHN RVDPTML I I TYSSDHN	- HPWPPSRNH ARPT - HAWPVS <mark>K</mark> TT TRLS	LK-KTEPDPVQ-		· · · · · · · · · · · · · · · · · · ·
AtWRKY-IIE16 VYPPSDSWAWRKYGQKPIKGS-P	YP - RGYYRCSSSKGCPARKQVER	RVDP SKLMI TYACDHN	- HPFPSSSANTKSHHRSSV	VLKTAKKEEEY-		E
SbWRKY-IIE1b GPPPSDSWAWRKYGOKPIKGS-P SbWRKY-IIE1b GPPPSDSWAWRKYGOKPIKGS-P	YP - RGYYRCSSSRGCPARKOVERS YP - RGYYRCSSSRGCPARKOVERS	RADPTVLLVTYSFEHN RADPTVLLVTYTFDHN	- HPWPQPKSSSCHAS- - HEAPQPKSSSCHQQG	KSSPRSTAPKP -		EPVADGQHPEPA EPVVEQDEL
ZmWRKY-IIE1# GPPPSDSWAWRKYGQKPIKGS-P	YP - RGYYRCSSSKGCPARKQVER	RADPAVLLVTYTFDHN	- HEAPRPKSS CCHQQG	····R <mark>PP</mark> APKP·	· · · D · · · · · · · ·	E <mark>P</mark> V AEQEDEL
ZmWRKY-IIE16 GPPPSDSWAWRKYGQKPIKGS-P	YP - RGYYRCSSSKGCPARKQVER	HADPTALLVIYIFEHN	- HEAPOPKISGGSCCHOOG	KPSAKPPAPKP -		E <mark>P</mark> SVVLVQRDEL
SbWRKY-IIE2a EVIPSDLWAWRKYGQKPIKGS-P	YPR-GYYRCSSSKGCMARKQVER	RS DPNML VITYTAEHN	- HPWPMQRNVLAGYSRPHT	HMS - NCKKK - I	NS	CRV
Atwrky-IIE2a GSRSSDLWAWRKYGOKPIKSS-P	YPR - SYYRCSSSRGCMARKOVERS	RT DPNVSVITYISEHN	- HPWPMQRNVLAGYSRPHT - HPFPTLRNTLAGSTRSSS	SKCSDVTTSASS	NS	VSQD
SbWRKY-IIE2b GSSSSDPWAWRKYGQKPIKGS-P	YPR-GYYKC <mark>SS</mark> MKGCMARKLVER	PAKPGVLIVTYMAEHC	• <mark>HPVPT</mark> QLNALA <mark>GTTR</mark> HKT	S S S <mark>G</mark> AA	s	····s <mark>P</mark> KSHEQ
SbWRKY-HE2c GVST - DLWAWRKYGQKPIKGS - P SbWRKY-HE2c GVSS - DLWAWRKYGQKPIKGS - P	YPR - GYYKCSSLKACMARKMVER YPR - GYYKCSSLKSCMARKLVER	PEKPGVLVITYIAEHC Pakpgvlvvtyiadhc	- HAVPTQLNSLAGTTRN - HAVPTMLNALAGTTRNRP	AAESPDD GDH	4	K <mark>p</mark> as <mark>p</mark> dq Hhqehhd
OSWRKY-IIE2d SGPAPDLWAWRKYGQKPIKGS-P	YPR-GYYRCSSNKNCAARKQVER	RFDPSFLLLTYTGAHS	<mark>GHDVPLHRN</mark> SLA <mark>GTTR</mark> HK <mark>P</mark>	P		
ZmWRKY-IIE2c DGSSSDMWAWRKYGQKPIKGS - P SbWRKY-IIE2d - GTSADPWAWRKYGQKPIKGS - P	YPR - GYYRCSSSKGCAARKOVER YPR - GYYRCSTDKDCRARKOVER	RADPNTFILTYTGEHN RTDASTLIV <mark>SYTGEH</mark> S	- HAAPTHRNSLAGTTRNKF - HPVPLHRNALAGTTRNKP	QPAP S		SAAPP TS <mark>P</mark> AE
ZmWRKY-IIE2b VGASPDPWAWRKYGQKPIKGS-P	YPR-GYYRC <mark>ST</mark> DKDCRARKQVER	RTDPSTVIV <mark>SYT</mark> GEHS	• HPVPLHRNALAGTTRNKP	QPASSIC		••••••••••••••••••••••••••••••••••••••
Atwrky-iie28 EALNSDVWAWRKYGOKPIKGS-P	YPR - GYYRCSSSRGCAARKOVERS YPR - GYYRCSTSKGCLARKOVERS	IRSDPKMFIVTYTAEHN	- HAAPTHENSLAGTTENKE - HPAPTHENSLAGSTEQKP	SDQQ · · · · · · · ·		·····saa <mark>p</mark> q
	YPR-SYYRCSSLKGCLARKQVER	RTDPEIFIVTYTAEHS	• HSHPTRRNSLAGITRNKF	ST <mark>P</mark> KK <mark>P</mark> SKAE <mark>PP</mark>	5	····T <mark>Р</mark> ТТ·····
GmWRKY-IIE22 ENLSSDIWAWRKYGQKPIKGS-P	YPR - GYYRCSSSKGCLARKQVER	IRSDPAMFIVTTAEHN	- HPAPTHKNSLAGSTRUKP	QTATAEDA		AT
	YPR-GYYRCSSSKGCLARKQVER	IRSDPTMFIV <mark>TYT</mark> AEHN	- HPAPTHENSLAGSTROKP	LVPQTATTTEED	5	DKSKSL
MtWRKY-IIE22 ENLSSDIWAWRKIGORFINGS-P	YPR - GYYRCSSSKGCLARKQVER	IRT DPTMFIVTYTAEHN	- HPAPTHKNSLAGSTROKP	LAPQIATTIEED		····NKDL
	YPR-SYYRCSSSKGCLARKQVER	PQNPEKFT I TYTNEHN	• HELPTRRNSLAGSTRAKT	SQ <mark>P</mark> K·····		· · · · · <mark>P</mark> TL · · · · ·
GmWRKY-IIE2d AEGLDDAWAWRKYGQKPIKGS - P	YPR - SYYRCSSSKGCLARKQVER	HLDPAVFLVTYTAEHS	- HPHPTRRNSLAGTTRKNN	ISLV <mark>PPP</mark> T		· · · · · · · · · · · · · · · · · · ·
GmWRKY-IIE20 AEGLDDAWAWRKYGOKPIKGS - P	YPR-SYYRCSSSKGCLARKOVERS	HLDPAVFLVTYTAEHS	- HPHPTR <mark>RNSLAGTTR</mark> KNN	ISLV <mark>PPP</mark> T		
CpWRKY-IIE2b ENLSADVWAWRKYGQKPIKGS - P	YPR - NYYRCSSSKGC <mark>S</mark> ARKQVER	N L <mark>D P</mark> N I F I V <mark>T Y S G D H</mark> T	- <mark>HP</mark> K <mark>PT</mark> H <mark>RN</mark> SLAGSTRNKP	LSEALKSI	5	· · · · · SADPEQS
GmWRKY-IIE21 DNLSADLWAWRKYGOKPIKGS - P	YPR-NYYRCSSSKGCMARKQVER	NTEPDMFIV <mark>TYSGDH</mark> S	- <mark>HPRPTHRNSLAGSTR</mark> NK I	P ATNPLPSP	3	SLSSFQA
GmWRKY-IIE2h QNLSSDLWAWRKYGQKPIKGS-P	YPR-NYYRCSSCKGCAARKQVER	T S E PNT F I VT YT G DHK	- HAKP VHRNSLAGNTRTKL	S · · · TTRLPETH		
GmWRKY-IIE2I DNLSSDLWAWRKYGQKPIKGS-P	YPR-NYYRCSSCKGCVARKQVER	TTEPNTFIVTYTGDHK	- HAKPVQRNSLAGSTRTKP	S TTRLSEPNI	E	SVTC <mark>P</mark> KRE
GmWRKY-HE2/ ************************************	YPR - SYYRCSSSKGCLARKHVER	QLDPGV-LIAIEDEHE	- QIKELKKEE DFVEAD	GWFPSIE		VEELK <mark>G</mark>
OSWRKY-IIE2e * DGVS AD VWAWRKYGOKP I KGS - P Rawrky-IIE2e * DGVS AD VWAWRKYGOKP I KGS - P	YPR - GYYRCSSSKGCPARKOVER	RSDPNTFILTYTGEHN	- HSAPTHENSLAGTTENKL	PSSSAASAA		· · · · · · · · · SAQPQ
S&WRKY-HE2e * DGSSSDVWAWRKYGQKPIKGS-P	YPR · GYYRCSSSKGCAARKQVER	RADPNTFILTYTGEHN	- HAAPTHENSLAGTTEHKE	P A		SAT PQ
SBWRKY-IIEXa DPTPPDSWAWRKYGOKSIKDT-P ZmWRKY-IIEXa GASPDPWAWRKYGOKPIKGS, P	YHR - SYYRCST DKKCKARKHVOR	LTQ - SFLAV <mark>SY</mark> IGEHS	- HPMPLARNGQAGTTHQK	P PRQ		<mark>P</mark> T S <mark>P</mark> V
ZmWRKY-HE21 *SSSSPDPWAWRKYGQKPIKGS-P	YPR · GYYKCSSMKGCTARKLVER	PAKPGVL I VTYMAEHC	- HP VP AQLNALAGTTRHK	SS <mark>G</mark> AATELHSAA	s	· · · · · SPSSHEQ
ZmWRKY-IIE2g GSSSSDPWAWRKYGOKPIKGS - P	YPR-GYYKCSSMKGCMARKLVER	PAKPGVLIVTYMAEHC	- HPVPTQLNALAGTTRHK	SSGAAAE - HSAA	s	····S <mark>P</mark> KSHEH
Oswrky-lie21 * GVSS-DLWAWRKYGQKPIKGS-P	YPR - GYYKCSSMKGCMARKMVER	PAKPGMLVVTYMAEHC	- HPVPTQLNALAGTTRHKS	APTG	D	DDK <mark>P</mark> TS <mark>P</mark>
MWRKY-IIE28 DNLSTDLWAWRKYGQKPIKGS-P	YPR - NYYRCSSSKGCTARKQVER	NT EADMFTVTYTGDHN	- HARPTHRNSLAGSTRTKS	8 <mark>P</mark> VTH <mark>P</mark> TTSI	s	<mark>G</mark>

310

IIE2-

IIE1-

	360	370	380	390	400	410	420	430	440	450	460	470
GmWRKY-IIE3a	EVVPSDLWAWRH	YGOKPIKGS	- PYPR - GYYRCS	SSKGCPAR	KOVERSRTDP	NMLVITYTSEHN	- HPWPTHR	NALAGSSRSQI	SKNNNIAASKN	E		
GmWRKY-IIE3b	E V V P S D L WAWRH	YG <mark>Q</mark> KP I KGS	- PYPR-GYYRCS	S S K G C S A R	KQVE R S R N D P I	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I-HPWPKQR	N A L A <mark>G S T R</mark> S Q I	SKNNNNAAAS - I	NSK		N
/vWRKY-IIE3a	E V V P S D L WAWRH	YG <mark>QKP I KG</mark> S	- PYPR-GYYRCS	S S K G C S A R	KQVE R S R T D P	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I - HPWPTQR	NALA <mark>GSTR</mark> SQI	SKNN <mark>G</mark> ASKNS <mark>G</mark>	scq		s
tWRKY-IIE3a	EVVPSDLWAWRH	YGQKP I KGS	- PYPR-GYYRCS	S S <mark>k</mark> g c s a r	KQVE RS RT DP	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I - HPWPTQR	N A L A <mark>G S T R</mark> S S S	SSSLN <mark>P</mark> S-SKS	STAA		
tWRKY-IIE3b	E V V P S D L WAWRH	YG <mark>Q</mark> KP I KGS	- PFPR-GYYRCS	S S K G C S A R	KQVERSRTDP	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	- HPWPIQRI	NALAGST RSST	'SSSSN <mark>p</mark> n <mark>p</mark> sk <mark>p</mark> :	STA		
bWRKY-IIE3a	E V V <mark>P S D</mark> L WAW <mark>R</mark> H	YG <mark>QKP I KG</mark> S	- PYPR-GYYRCS	S S K G C S A R	KQVE R S R T D P I	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I - HPWPTQR	N A L A <mark>G S T R</mark> NHI	I <mark>G</mark> KNS <mark>GG</mark> S S <mark>G</mark>	sks		
mWRKY-IIE3a	E V V P S D L WAWRH	YG <mark>QKP I KG</mark> S	- PYPR-GYYRCS	S S K G C S A R	KQVERSRNDP	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I - HPWPTQRI	NALAGST RNHH	ISKNS <mark>gg</mark> s S <mark>g</mark>	sк		
mWRKY-IIE3b	E V V <mark>P S D</mark> L WAW <mark>R</mark> H	YG <mark>QKP I KG</mark> S	- PYPRLGYYRCS	S S K G C S A R	KQVERSRNDP	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I - HPWPTQR	NALAGST RNHI	I <mark>G</mark> KNSS <mark>G</mark> SS <mark>G</mark>	sк <mark>с</mark>		
sWRKY-IIE3a	E V V P S D L WAWRH	YG <mark>QKP I KG</mark> S	- PYPR-GYYRCS	S S K G C S A R	KOVERSRTDP	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I - HPWPTQRI	NALAG <mark>STR</mark> SHH	ISKNS <mark>ggggg</mark> sg	sк <mark>с</mark>		
sWRKY-IIE3b	E V V P S D L WAWRH	YG <mark>QKP I KG</mark> S	- PYPR-GYYRCS	S S K G C S A R	KQVERSRTDP	NMLVI <mark>TYT</mark> S <mark>EHN</mark>	I - HPWPTQRI	NALA <mark>GSTR</mark> SHI	AKNSSSNSSSS	GAS		
dWRKY-IIE3a	E V V <mark>P S D</mark> L WAW <mark>R</mark> H	YG <mark>QKP I KG</mark> S	- PYPR-GYYRCS	S S K G C P A R	KQVERSRTDP	NTLVITYTSEHN	I - HPWPTQR	N V L A <mark>g s t R</mark> sh`	AKNSS NTTA	AAN		
bWRKY-IIE3b	E V V P S D L WAWRH	YG <mark>QKP I KG</mark> S	- PYPR-GYYRCS	S S K G C P A R	KOVERSRTDP	S L L V I <mark>T Y</mark> N S <mark>E H</mark> N	I - HPWPTQRI	NALAG <mark>STR</mark> SHH	IAKNSKNN <mark>P</mark> SQHI	NL		
mWRKY-IIE3c	EVVPSDLWAWRH	YGOKPIKGS	- PYPR-GYYRCS	S S K G C P A R	KOVERSRTDP	S L L V I <mark>T Y</mark> N S <mark>E HN</mark>	- HPWPTQR	NALAGSTRSH	IAKSSKNKYSSH	5 L <mark>P</mark>		
INDICY HEAD		YGOKPIKGS	- PYPR- GYYRCS	S SKGCPAR	KOVERSRTDP	S L L V I <mark>T Y</mark> N S <mark>E HN</mark>		NALAGSTRSH	AKSGKNS-SSH	s L		

IIE3-

11									WFxD		
	480	490	500	510	520	530	540	550	560	570	580
AtWRKY-IIE1a	EEDNKFMVL <mark>G</mark>		. R <mark>G</mark>		- ETT <mark>P</mark> SCVD	DE			- FAWFTEM		ETTSST
VvWRKY-IIE1a	PPE-EFAVFA		- SQ <mark>P</mark> D		LELG0	DSLLAD <mark>G</mark> H			F S WF S E V -		• • • • • • • • • • • • • • • • • • •
VvWRKY-IIE1b	PPSPKPAIFS		- SQAE <mark>PDP</mark> D		EKFTDLGE	ESLITTED			F T WF S D F -	· · <u>·</u> · · · · ·	E S <mark>P</mark> S S T
OsWRKY-IIE1a	EETHQYGVTA		- <mark>GP</mark> AT <mark>GGG</mark>		<mark>GG</mark> AAA I E	VR			DEFRWLYDVV	SV <mark>P</mark> · · · · ·	AT ST S <mark>P</mark> S D -
BdWRKY-IIE1a	KSDDAHQGTA		• ANVVN • • • • • •			EVH			DEFRWLYDVV	SV	TCCAC
ZmWRKY.IIE1a	PSSGSHDVAA		ATVCOAGAP ···		TEAG. AAAVE	VR			DEFRWLYDGY	sv	
GmWRKY-IIE1a	PPEEEMKVFA		. T D S D		· · · · · LELA	DAAVLLSHHH	ннн		· · FGWFDD ·		······································
GmWRKY-IIE1b	PPEEEMTVFA		- T D S D		LELA	DAAVLLSHHH	нн		F GWF D D		
MWRKY-IIE1a	PQPEDR <mark>P</mark> IFV		• T H <mark>P</mark> D • • • • • • •		FDLT	D HI	HAV		· · F GWF D D · ·		
GmWRKY-IIE1c	EPEEKFAD		· · · E · · · · · · · · ·		SMITTAE	E			L GWL GEM-		••••••EATSST••
GMWRKY-IIE18	EPEEKFADLG		. <mark>G</mark> DD		SMITIAE	E			L GWL GEM		DINCO
AtWRKY-IIE16	EEEEELTVTA		AEE <mark>P</mark>		- PAGLDLSH	VDSPLLLGGC	YSEI <mark>G</mark> ····		- EFGWFYDA		
OsWRKY-IIE1b	ENESSASAEL · ·		- EV <mark>PEPEP</mark> EQES	E <mark>P</mark> - VVKQEEE	QKEEQKAVVE	PAA VTTT	va	APAVEEEDEN	FDFGWIDQYH		·······
SbWRKY-IIE1b	<mark>gp</mark> ehel - Aet		- EV <mark>P</mark> EQQE <mark>P</mark> VEE	E 0	EQKVVPGLAG	BPEAEAEAE <mark>P</mark> T	· · · · · · AT	∕A <mark>P</mark> AAAEEDES	FDF <mark>GWF</mark> DQY-		<mark>P</mark> TWHRSA-
ZmWRKY-IIE1b	<mark>бр</mark> ене		· · · <mark>P</mark> EDEQE · · ·	<mark>P</mark>	EQKVVPGLAG	SPET · · · APAT	· · · · · · · AT	LA <mark>P</mark> VA · EEDES	FDLGWFDQY-		··············
ZmWRKY-NE1c	GPEHEL - VET		- EVPEQEQETEP	E · · · · EEHEP	AEEGVPCLAC	BPEA · · · · AAT	· · · · · · LT	APAA - EEDES	FOFGWEDQY		PTWHRSA-
ShWRKY-IIE2a	E <mark>P</mark> TSSWPTS			s s k N	ANYFEHNVV	SSNIECQQMT		· · · · · · · NM · M	EDNAA - GYVA		
ZmWRKY-HE2a	E <mark>P</mark> TSSW <mark>P</mark> TT	1	гт s s <mark>p</mark>	sskn	GNY	E C Q Q M T -		NMTM	IE E NAA <mark>AG</mark> YVA		
AtWRKY-IIE2a	KE <mark>gp</mark> dkshl <mark>p</mark> ss	P						A	SPPYAAMVV-		
SbWRKY-IIE2b	GQAVEKAAGRG -							· · · · · · · · A <mark>G</mark>	DREH <mark>GNNE</mark> TS		
OsWRKY-IIE2c	QQQQQPSPG								GASTDEAA	A	
OsWRKY-IIE2C		T						A	AFAATASOSP	G	
ZmWRKY-IIE2c	PPPPSVVVGGAG	AGDA							SPSPASTSTA	G	
ShWRKY-IIE2d	QP - PAASPI							· · · · · · · · · v	GVEYEEDDT -		
ZmWRKY-IIE2b	Q <mark>P - P</mark> AAS <mark>P</mark> I G							<u>.</u>	<mark>. D</mark> D	<u>.</u>	
ZmWRKY-IIE2d	PPPP SVVVGGDA							QDRQ <mark>P</mark>	S P S P T S T S T A	G	
AtWRKY-IIE26	- T SKSPT							· · · · · · · · · · · · · · · · · · ·	TIATYSSP-	 	
VVWRKY-IIF2a	VKPSAKPA								SPATSADD		
GmWRKY-IIE2a	IKPASPST							s	GMEEEVAQH -		
GmWRKY-IIE2b	TK <mark>P</mark> TS <mark>P</mark> AT · · · ·							· · · · · · · · s	GAEEEAPTP -		
GmWRKY-IIE2¢	TK <mark>P</mark> TS <mark>P</mark> AT · · · ·							· · · · · · · · s	GAEEEAPTP -		
MtWRKY-IIE2a	TKPSSPTT							· · · · · · · · · · · · · · · · · · ·	GGEEEVQT		
CHWRKY-HE2C									VVSSPISNPM	T	
GmWRKY-IIE2d	THRHQKN								- TNCS SKSI-		
GmWRKY-IIE2e	T T RHQKK • • • • •								• T T C S <mark>S</mark> S		
AtWRKY-IIE2d	KEEIHLS · · · ·						· · · <mark>·</mark> · · · · · ·		· · PTTPLKG ·		
CpWRKY-IIE2b	KPNPSCS						S <mark>P</mark> LSTNSF	^F S <mark>P</mark> A T MN <mark>G</mark> E E <mark>G</mark>	NMERNPHEGL	DSKEQK	
GmWRKY-IIE2f	TPFSSSSSSP								PHSPTSP-		
GmWRKY-IIE2h	-TVSSFSHV								GKVPSEE.		
GmWRKY-IIE2i	NACSSNSEL···						SVSDT <mark>p</mark> knee	T V S AGE P D C P	DMEIEPSEND		
GmWRKY-IIE2j	*• T R D D Q H • • • T •					.	.		PVVAEAIED-		
GmWRKY-IIE2k	LALDDRG···F·							· · · · · · · · · · L	DSTWLAIFL -	<u>.</u>	
OSWRKY-NE2e	REPERSIONAGEA	Δ							HPSPPSTSTA		
ShWRKY-IIE2	* PPPPSVVVGGAG	AGAGAAP				DA		QHQHQQ	SPSPTSTSTA	G	
ShWRKY-NEXa	EQPPQASAPAPA	P						AL	APSAASASPM		
ZmWRKY-IIE2e	[*] QPDPAASPIDDT	ΤΤ <mark>Ρ</mark> ΙΙ·····						C <mark>P</mark> S V L	GVECEEEE	т	
ZmWRKY-IIE2f	GQT VEKAAG · · ·					• • • • • • • • • • • • •		•••••LR	E R <mark>G</mark> I GNNETS	N	
ZmWRKY-IIE2g	GG ···QAAGRG ·								DREHGSSETS		
ZmWRKY-NE2h	GPAAGRAAAGE								CEDVDGNE -	AA · · · · · · ·	
MINRKY-NE25	OPNSSISSC								- SSLAPNSF -		
					_						

IIE1-

312

IIE2-

	480	490	500	510	520	530	540	550	560	570	580	
GmWRKY-IIE3a	· · · · · · EE · · · · · ·		EEEES	• NSDSN		NNNVFVRE			EEKQL - MDDVE	E		
GmWRKY-IIE3b	EGSNSQKE	s	NNYNN	S S E <mark>G</mark> N	IVS <mark>P</mark> VVA <mark>G</mark> NSS	SNSVK - EENM	E	· · · · · · D I	EKHQIE <mark>m</mark> d <mark>eg</mark> e	E		
VvWRKY-IIE3a	QKATGLKE ·····		ENKES	YNNDD	MS <mark>P</mark> IVGGSS1	TT <mark>G</mark> ASVKEEM	g	NV	EK-QLE <mark>MDD</mark> SE	E		
AtWRKY-IIE32	TTSPSSRVF		QNN\$\$	KDE <mark>P</mark> N	INSNL <mark>P</mark> SSSTI	I <mark>PP</mark> - FDAAAI	K	EE	NVEERQ - EKMI	E		
AtWRKY-IIE3b	· · · · · · · NV · · · · · · ·		• N S S S • • • • • • • •	I G SQN	IT I YL <mark>P</mark> SSTT <mark>I</mark>	P P P T L S S S A I I	к	DE	R <mark>g</mark> ddme <mark>l</mark> en vi			
ShWRKY-IIE3a	QNEKQQQQQQ	<mark>P</mark>	NNVKEE	<mark>P</mark> KD <mark>P</mark> A	ATTTTTSTI	TTTTTSTS <mark>P</mark> A	A	· · · · · · v v	KEET LA <mark>AG</mark> SSS	5	EALG	
ZmWRKY-IIE3a	- NEKQQQQS	N	TDVKEE <mark>P</mark> NKD	<mark>Р</mark> АТТТ	TTTTTTSTI	TTMTTSAS <mark>P</mark> A	A	· · · · · · vv	KEETLA <mark>AG</mark> SSI	E		
ZmWRKY-IIE3b	QNEKQEQQQ · · · · ·	<mark>P</mark>	SNVKEE	<mark>P</mark> KD <mark>P</mark> A	ATTTATSTI	TTLTTSIS <mark>P</mark> A	A	· · · · · · v v	KEETLA <mark>AG</mark> SSE	E	A . L G	IIE2
OsWRKY-IIE3a	QNDKSQQQ <mark>P</mark> · · · · ·		- S V K E E Q	· · · KDQATT	ATTTTTSTI	TTTN - SAS <mark>P</mark> V	v	v к	E E E A A L <mark>A G</mark> S S E	E	ALEL	IIE3-
OsWRKY-IIE3b	ASKNNSSHSGYHH	н	HHQK <mark>P</mark> LVK	AE <mark>P</mark> N	IDQSAAATTAA	AT V <mark>P</mark> VKEEAA	м	v <mark>g</mark>	T S S E A L <mark>A</mark> KT T C	a	KSME	L
BdWRKY-IIE3a	SKKNSSSR		- QQKP A	••••KAE <mark>P</mark> F	DH PQT AAS ST	гтт <mark>т</mark> s · · · · ·			- APPAAVKEE	4		
SBWRKY-IIE3b	QKPDLKAEP	E	HHQASAAVV	- PTGCATTA	TTAATSTTT	TATTSTTSNS	r	<mark>PP</mark>	PA-TMAVKEEA	A - M		
ZmWRKY-IIE3c	QTPNLKAEP	E	HHQAS <mark>G</mark> AV	<mark>P</mark> SATA	ATAATATSTI	TTTSTTSNS	r	<mark>PP</mark>	<mark>P P P</mark> T MA <mark>V</mark> K <mark>E</mark> E A	A I M		
ZmWRKY-IIE3d	.QK <mark>P</mark> NLKAE <mark>P</mark>	E	HHQASEAV	- PSYATTAT	AATTTSTTT	TATTSTTSNS	r	<mark>PP</mark>	ET - VAAVKEEA	A - M		
GmWRKY-IIE3c ^{**}	· EE		EEEES	N S G S N	NVP	NNSAFVRE			EEKQLE <mark>M</mark> D <mark>DG</mark> E			
				_								

ppenaix		-(1								
	LESP	A			acidic-4					
	C 00			640			676	680	CD 0	700
AMARKY 1154-	LIES DIESS			••••	EDESIEADIGELE	ECSVVE.	, , , , , , , , , , , , , , , , , , ,	680	650	
WWRKY-IIF1a		YSAEDDV			EEESLFADLGELP	ECSLVF				
VvWRKY-IIE1b	· · · · · MLESPIFAEI	NNVRDADV	AMFF <mark>P</mark> MRE		EDESLFADLGELP	ECSLVF				
DsWRKY-IIE1a	IDAADEMQLYDQ	<mark>P</mark> · · · · · MF F <mark>GG</mark> AV <mark>V</mark>	GTAALL <mark>P</mark> DEF - <mark>G</mark> I	DVGGLGGEGLG	E E E A L F E <mark>G L G E L P</mark>	ECAMVFR-				· · · · · · · · RR <mark>AG</mark> D
BdWRKY-IIE1a	VEAADDLLLY <mark>GP</mark>	· · · · · · · MFF <mark>G</mark> KAAV	DT LL <mark>P</mark> DEF - <mark>G</mark> I	F <mark>G G G D A A G G E R S</mark> E	E D D A M F A G L G E L P	ECAMVFR-				· · · · · · · · · RH <mark>AG</mark> D
SbWRKY-IIE1a	VEAADEM-LYGA	MEEGA	PPAAPLPDEFVG		EEDAMFAGLGELP	ECAMVER-				RHAGD
GmWRKY-IIE1a	·····VLESPICGE		AV - PMRE		EDESLFADLGELP	ECSAVF				···· RRMN I
GmWRKY-IIE1b	····VLES <mark>P</mark> ICGE	voov	AV·· <mark>P</mark> MRE····		EDESLFADLGELP	ECSAVF				
MWRKY-IIE1a	····VLVS <mark>P</mark> IC <mark>GG</mark>	· · · · · · · · · · · · VED <mark>V</mark>	TL • • TMRE • • • • •	• • • • • • • • • • • • •	EDESLFADLGELP	E C S T V F				•••••RQRN I
GmWRKY-IIE1¢	·····VLESPFMAT	Manual Manu Manual Manual Manu	AL - I PMRE	••••••	EDESLFADLGELF	ECSVVF				·····RQGLL
SINWRKY-IIE18	ILESPIMAT		SVEL PMEE		EDELLEADIGELE	ECSTVF				RUGLL
AtWRKY-IIE19	GSSNFLD	TLERGF	S VGQE		EDESLFGDLGDLF	DCASVF				
OsWRKY-IIE1b	· · · · · YA <mark>P</mark> LL <mark>PP</mark>	EEW	ER··ELQ <mark>G</mark> ····		- DDALFAGLGELP	ECAVVFG-				
SbWRKY-IIE1b	····· LYA <mark>P</mark> LL <mark>PP</mark>	· · · · · · · · · · · · · · EEW	ER••ELQ <mark>G</mark> ••••		· E DALFAGLGELP	ECAVVF <mark>G</mark> -				RRRE L
ZmWRKY-IIE16	····· LYAPLLPP	· · · · · · · · · · · · · · · EEW	ER··ELQ <mark>G</mark> ····	••••••	· EDALFAGLGELP	ECAVVFG				· · · · · · · · · · · RRREL
ZmWRKY-HE1¢	······································		EK ELQ <mark>G</mark>		· EDALFAGLGELF	ECAVVIG				
				L						
SBWRKY-IIE2a ZmWRKY-IIE2a AtWRKY-IIE2a		····YAIDTLDEEGVA	AMHQPINRNSI AMHQ <mark>P</mark> INC <mark>G</mark> SV OVDVEEDTFI <mark>P</mark> EL	F <mark>P</mark>	S - DEVFAELEELE SADEVFAELEELE ED <mark>T</mark> FADMDKLE	PSN PTT ENSQTMF-				
ShWRK Y-IIE2b		- SSMAGEFGGEETAV	A - IDDDEFWPAG	M						FDFEH
ShWRKY-IIE2c		- STMTVEENDAWLVE	DHMALEDDVVDGD	C	PFDDFLWPFDDD-					LDQ
OsWRKY-IIE2d		LS <mark>P</mark> TT <mark>P</mark> LRASSME	ELH <mark>G</mark> EDDA · · · ·	E	- AELQVEEDDMAI	DDE	· · DDD · · · · ·		DV	ADETISTV <mark>PWG</mark> T
ZmWRKY-IIE2c		· · LS <mark>P</mark> TT <mark>P</mark> LRT <mark>P</mark> SME	EEEMDDD····							
PANDICV HEAD					· DELLVEDMEMVG	EDELLFLN	- IDADAEADA	GAPL · · · · · · ·	- EPMPSLFDV	VDE <mark>P</mark> FLNS <mark>PW</mark> LT
7mMBKV. HEAL		VAASVLLEDAE <mark>T</mark> E	GEEDVLSLF		- DELLVEDMEMVG - LELAPS <mark>PS</mark> NG <mark>S</mark> G	EDELLFLN SKDVMVST	- IDADAEADA ELHN - <mark>G</mark> R <mark>G</mark> SQI	GA <mark>P</mark> L KVVALSKLHEF	- EPMPSLFDV RH <mark>P</mark> ATTSSSR	VDEPFLNSPWLT
ZmWRKY-IIE20 ZmWRKY-IIE20 ZmWRKY-IIE2d		· · VAASVLLEDAE TE · · LSPTTPLRTPSME	GEEDVLSLF	EEEE	- DELLVEDMEMVG - LELA <mark>P</mark> S <mark>PS</mark> NGSG E DELLVEDMEMAG	EDELLFLN SKDVMVST EDELLFLS	- I DADAEADA E LHN - <mark>G</mark> R <mark>G</mark> SQI - T DADA <mark>G</mark> A	GAPL KVVALSKLHEF	- EPMPSLFDV RHPATTSSSR PTSSLFDV	VDEPFLNSPWLT TSDGLGAAPAAM ····LGAAPAAI VGEPFLSPPWVT
ZmWRKY-IIE20 ZmWRKY-IIE20 ZmWRKY-IIE20 AtWRKY-IIE20		- VAASVLLEDAE <mark>T</mark> - LS <mark>P</mark> TT <mark>P</mark> LRT <mark>P</mark> SME TSADEFVL <mark>P</mark> VEDHL	GEEDVLSLF EEEEEEE AVGDLD <mark>G</mark> E	EEEE	- DELLVEDMEMVG - LELAPSP <mark>SNGSG</mark> EDELLVEDMEMAG EDLLSL <mark>S</mark> DTV	EDELLFLN SKDVMVST EDELLFLS SDDFFD <mark>G</mark> L	- I DADAEADA E LHN - <mark>G RG</mark> SQI - T DADA <mark>G</mark> A E E F AA <mark>G</mark> D:	GAPL KVVALSKLHEF SF	- EPMPSLFDV RHPATTSSSR PTSSLFDV S	VVDEPFLNSPWLT TSDGLGAAPAAM ····LGAAPAAI VVGEPFLSPPWVT GNSAPASFDLS-
ZmWRKY-IIE20 ZmWRKY-IIE20 AtWRKY-IIE20 VvWRKY-IIE20	v	 VAASVLLEDAET LSPTTPLRTPSME SADEFVLPVEDHLA SAME - DELQPYVS 	GEEDVLSLF EEEEEE AVGDLDGE SQQEIRGE	EEEE	- DE LLVE DMEMVG - LE LAPSPSNGSG E DE LLVE DMEMAG E DLLSLSDTVV - E DQKLNERNNTN	EDELLFLN SKDVMVST EDELLFLS SDDFFD <mark>G</mark> L EVHISD <mark>E</mark> I	- T DADAEADA E LHN - G RG SQI - T DADAGA E E F AAGD LNND F F LG LE	GAPL	- EPMPSLFDV RHPATTSSSR PTSSLFDV S	VVDEPFLNSPWLT RTSDGLGAAPAAM ····LGAAPAAI /VGEPFLSPPWVT GNSAPASFDLS· ··DMDFSVQFPS
ZmWRKY-HE20 ZmWRKY-HE20 ZmWRKY-HE20 AtWRKY-HE20 VvWRKY-HE20 VvWRKY-HE20		VAASVLLEDAET LS <mark>P</mark> TT <mark>P</mark> LRT <mark>P</mark> SME TSADEFVL <mark>P</mark> VEDHLA SAME - DELQPYVS LVQQSTSVQSKD L	GEEDVLSLF EEEEEEE AVGDLDGE QQEIRGE VODEE	EEEE	- DELLVEDMEMVG - LELAPSPSNGSG EDELLVEDMEMAG - EDLLSLSDTVV - EDQKLNERNNTN DELGFSDTAV	EDELLFLN SKDVMVST EDELLFLS SDDFFDGL EVHISDEI DDDFFMGL	- TDADAEADA ELHN - GROSQI - TDADAGA - EEFAAGD LNNDFFLGLE EELAEQ-	GAPL	- EPMPSLFDV RHPATTSSSR PTSSLFDV S 	VDEPFLNSPWLT (TSDGLGAAPAAM LGAAPAA VGEPFLSPPWVT GNSAPASFDLS- DMDFSVQFPS DHFPP-SFPIP-
ZmWRK(Y-IIE28 ZmWRKY-IIE28 AtWRKY-IIE28 VvWRKY-IIE28 VvWRKY-IIE28 GmWRKY-IIE28 GmWRKY-IIE28		· VAASVLLEDAET · LSPTTPLRTPSME TSADEFVLPVEDHLA · SAME · DELQPYVS LVQQSTSVQSKD · L AKSESTEEEDMEDLA GEKSEFFFF0.	GEEDVLSLF EEEEEEE QGLLGGE QGEIRGE VEDEE-D IKNDEEL <mark>P</mark>	EEEE	- DE LLVE DMEMVG - LELAP S PSNGSG - E LLVE DMEMAG - E DLLSLSDTVV - E DLLSLSDTVV - E DLLGFSDTAV - DELGFSDTAV 	EDELLFLN SKDVMVST EDELLFLS SDDFFDGL EVHISDEI DDDFFMGL SDDFFEGL SDDFFESI	- I DADAEADA E LHN - G RG SQI - T DADAGA - EE FAAGD LNNDFFLGLE EE LAEQ EE LT	GAPL KVVALSKLHEF SF	- EPMPSLFDV RHPATTSSSR PTSSLFDV 	VVDEPFLNSPWUT (TSDGLGAAPAAM ····LGAPAAI VGEPFLSPPWVT GNSAPASFDLS- ··DMDFSVQFPS DHFPP-SFPIP- DPFS.·AIAIPS
ZmWR(Y-HE2b ZmWR(Y-HE2b ZmWR(Y-HE2b AtWR(Y-HE2b VyWR(Y-HE2b VyWR(Y-HE2a GmWR(Y-HE2a GmWR(Y-HE2c		- VAASVLLEDAET - LSPTTPLRTPSME TSADEFVLPVEDHLA - SAME - DELOPYUS LVQQSTSVQSKD - L AKSESTEEEDMEDLA QGEKSESREEKED - V	GEEDVLSLF GEEEEEE GUDGE GQEIRGE VGDLGE VGDEE VGDEE VGDEE VGDEE VGDEE VGDEE VGDEE VGDEE		- DE LLVE DMEMVG - LE LAPS PSNGSG EDE LLVE DMEMAG EDE LLS SDTVV - EDILLS SDTVV - DE LG FSDTAV - NE FG LTE TVV - NE FG LSDMVL - DE FG LSDMVL	EDELLFLN SKDVMVST EDELLFLS SDDFFDGL EVHISDEI DDDFFMGL SDDFFEGL SDDFFESL	- I DADAE ADA E LHN - GRGSQI - T DADAGA EE FAAGD LNNDFFLGLE EE LAEQ - EE LSQLT DE LSQLT	GAPL KVVALSKLHEF SF	- E P M A S L F D V RH P AT T S S S R P T S S L F D V S D C F A G S A T D C F S	V DE P F L N S P W T (T S D G L G A A P A A M L G A A P A A M V G E P F L S P P W V T G N S A P A S F D L S DM D F S V G F P S D F F A I S S N I D R :D P F S A I A I P S D P F S A I A I P S
SIMARA 1-11220 ZmWRKY-11220 ZmWRKY-11220 AtWRKY-11220 VVWRKY-11220 GmWRKY-11220 GmWRKY-11220 GmWRKY-11220 MWRKY-11220		. VAASVLLEDAE SPTTPLRTPSM TSADEFVLPVEDHL SAME - DELOPYVS LVQQSTSVQSKD - L AKSESTEEEDMEDLA QGEKSESREEKED - V GGEKSESREEKED - V GGDKSESREEKED - V	GEEDVLSLF EEEEEEE AVGDLDGE GQEIRGE VEDEE VEDEE MNDEE MDDEE MDDEE		- DE LL VE DME MVG - LE LARS PSNOS 26 - ED LL VE DME MAG - ED LL SL SD TVV - ED GKLNE RNNTN - DE LG FSD TAV - NE FG LSD VVL - DE FG LSD VVL - DE FG LSD VVL	EDELLFLN SKDVMVST EDELLFLS SDDFFDGL EVHISDEI DDDFFMGL SDDFFESL SDDFFESL SDDFFESL	- I DADAE ADAA E LHN - GROSQI - T DADAGA EE FAAG UND F F LG LE EE LAEQ EE LAEQ EE LT DE LSQLT DE LSQLT DE F G	GAPL KVVALSKLHEF SF VTG APSVVTG SVVTG	- EPMPSLFDV RHPATTSSSR - PTSSLFDV - DCFP - CSAT - DCFS - DCFS - FP	VVDEPFLNSPWLT (TSDGLGAAPAAN (TSDGLGAAPAAN (TSDFFLSPPWVT GNSAPASFDLS- - DMDFSVCFPS DHFPP-SFPP- DPFFS-AIAIPS DPFS-AISIPS DPFSS-AISIPS
SUMKKY-IIE20 ZmWRKY-IIE20 ZmWRKY-IIE20 VWWRKY-IIE20 VWWRKY-IIE20 GmWRKY-IIE20 GmWRKY-IIE20 GmWRKY-IIE20 AtWRKY-IIE20	E S S S	OKASVLIEDAE SPITPINT SME SAME OF VLEVEN UN VOLT V V V V V V V V V V V V V V V V V V V	GEEDVLSLF EEEEEEE AVGDLDGE SQEIRGE VEDEE MDDEE MDDEE MDDEE MDDEE MDDEE MDDEE SVAETSH	EEEE	- DE L LVEDMENNG E DE L LVEDMENAG E DE LLVEDMENAG - E DLS DTVV - E DQKLNERNNTN - DE LG FS DTAV - NEFGLS DVVL - DEFGLS DVVL - DEFGLS DVVL QAAGAIE GRRLSN	EDELLFLN SKDVMVST SDDFFDQL EVHISDEI DDDFFMQL SDDFFESL SDDFFESL SDDFFESL SDDFFESL	- TDADAE ADA E LHN - GRGSQI - TDADAGA - EE FAAGD LNNDFFLGLE - EE LAEQ - EE LAEQ - EE LAEQ - DE LSQLT - DE LSQLT - DE LSQLT SGSGTFPSFT	GAP L KVVALSKLHEF SF VTG APSVVTG G	EPMPSLFDV RHPATTSSSR - PTSSLFDV - DCFP - CSAT - DCFS - DCFS - FP	VVDEFFLNSPWLT ITS DGLGAAPAAM VVGEFLSPWVT
JMWRKY-IIE28 ZmWRKY-IIE28 ZmWRKY-IIE28 WWRKY-IIE28 WWRKY-IIE28 GmWRKY-IIE28 GmWRKY-IIE26 GMWRKY-IIE26 GMWRKY-IIE26 CPWRKY-IIE26 CPWRKY-IIE26	E S S P	- VAASVLLEDAE - LSPTTPLRTPSME TSADEFVLPVEDHL - SAME - DELOPYV LVQQSTSVQSKD - I AKSESTEEEDMEDL QEEKSESREEKED - I QECKSESREEKED - I QECKSESREEKED - I - SADESVAVQEMB - KMTEYDRLLQYV	GEEDVLSLF	EEEE	- DE LL VE DMEMAG E DE LL VE DMEMAG 	EDELLFLN SKDVMVST EDELLFLS SDDFFDGL DDDFFMGL SDDFFEG SDDFFESL SDDFFESL GL <mark>P</mark> S-DLM DISMPDII	- TDADAEADAA ELHN- GRGSQI - TDADAGA - EEFAAGD LNNDFFLGLE - EELAEQ - EELT - DELSQLT - DELSQLT - DELSQLT SGSGTFPSFT FSDELFPNLE	GAPL KVVALSKLHEF SF	- EPMPSLFDV RHPATTSSSR - PTSSLFDV - SS - CFS - CFS - CFS - CFS	VVDE FLNSPWLT ISDGLGAARAAM LGAARAAM VGE FLSPWVT WGE FLSPWVT WNSA PASFDLS. DMDFSVGFFS DHFFF.SFIP. DFFTAISSNIDR FS.AIAIPS DFFS.AIAIPS DFFS.AISIP DFDELLNSGE
SUMAKI YILE20 ZmWRKY-IIE20 AtWRKY-IIE20 VVWRKY-IIE20 GmWRKY-IIE20 GmWRKY-IIE20 GmWRKY-IIE20 MWRKY-IIE20 GWWRKY-IIE20 GmWRKY-IIE20 GmWRKY-IIE20 GmWRKY-IIE20	E E S S S S S 		GEEDVLSLF		- DE L VEDMENAG E DE L LVEDMENAG - E DLLS LEDTVV - E DALLS LEDTVV - DEL GEDTVV - DEL GEDTAV - DEL GEDTAV - DEFGLSDMVL - DEFGLSDMVL - DEFGLSDMVL - DEFGLSDMVL - REVFCCOHDL - KEEDFTDVL 5D0	EDELLFLN SKDVMVST EDELLFLS SDDFFFDGL SDDFFMGL SDDFFEG SDDFFESL SDDFFESL GLPS-DLM DISMPDII SAQFVDGW	- IDADAEADA - IDADAEA- - TDADAEA- - EEFAAGD LNNDFFLGLE - EELAEQ - EELAEQ - EELSQLT - DELSQLT - DELSQLT SQSQTFFSFT FSDELFPNLE ISOLEKLI	GAPLLSKLHEF SF VTG APSVVTG ASSVVTG	EPMPSLEDV RHTSSLEDV DCFP GSAT DCFS DCFS FP	VUDE FLNSPWLT ISOCLGAARAAM LGAARAAM LGAARAA OVEE FLSPPWVT ONSARASFDLS. DMDFSVOFFS DHFFPSFFPIF DFFTAISSNIDR DFFS-AIAIR DFFS-AIAIR DFFSS-AIAIR FYDELLNE OFFSSLSIFN GLECOH
SJIVKKY-IIE28 ZmWRKY-IIE28 ZWWRKY-IIE28 VVWRKY-IIE28 WWRKY-IIE28 GmWRKY-IIE28 GmWRKY-IIE28 MWRKY-IIE28 AWWRKY-IIE28 GmWRKY-IIE28 GmWRKY-IIE28 GmWRKY-IIE28		- VAAS VLLEDAEN - LSM TTULRT SM - SADEFVL VEDHL - SAME - DELON - VAOST SVORSKO - I - AKSESTEEEDMEDL - OEKSESREEKED - OEKSESREEKED - OEKSESREEKED - OOEKSESREEKED - KTEYDRLLOYV - KTEYDRLLOYV - V	EEEEEEE VOOLDOE VOOLDOE VOOLDOE VEDEE VEDEE VMDDEE VMDEE VMDEE VMDEE VAGETSTH VVKQEEER VAGETSTH VKAGEEER INLK DLK OG QOUNGEEEEEEE	EEEE	- DE L L VE DME MAG - DE L L VE DME MAG - E DL S L S DT VV - E DGK LNE RNN N - DE L G F S DT AV - DE F G L S DWL - R AVFKG C DHD L - K E DF T DWL BDD - K E D F L DWL	EDELLFLN SKDVMVST SDDFFDGL EVHISDEI DDDFFDGL SDDFFFGL SDDFFESL SDDFFESL SDDFFESL SDDFFESL GLFS-DLM DISMPDII SAQFVDGW	- IDADAEADA - IDADAGA - TDADAGA - EEFAAGD INNDFIGLE - EEIAEQ - EEIAEQ - DEISQLT - DEISQLT - DESQLEKLI - SSDLEKLI - CSDLEKLI - CSDLEKLI	AALSKLHEF SF	EPMPSLEDV PTSSLEDV S OCFP GSAT DCFS DCFS DCFS FF	VVDE FLNS PWLT VVDE FLNS PWLT VVDE FLS PPWT ONSAFASFDLS - DMDFSVQF PS DHF FFSFIF DFFTAISSNIDR FFS-AIAIPS DFFS-AIAIPS DFFS-AIAIPS - DFDELLNS QE - NL - YOMMEF - UECQH - GLECQH
300 WRKY-11E2b 2m WRKY-11E2b 2m WRKY-11E2b 40 WRKY-11E2b WWRKY-11E2b Gm WRKY-11E2b Gm WRKY-11E2c MWRKY-11E2c GM WRKY-11E2c Gm WRKY-11E2c Gm WRKY-11E2c AtWRKY-11E2b AtWRKY-11E2b	E	- VAASVLLEDAEM - LSMITHLENTSMI - SAME-VELON - SAME-DELON VWEDHL - SAME-DELON VWEDHL - SAME-DELON - SAME-SEREKED - SAMESSEREKED - SAMESSEREKED - SAMESSWAVEN - KUTEVDRLLOVV - V - OOD - NDDVQETNOEDMI - SAMESSAVO	EEEEEEE VODLDGE VODLDGE VEDEE VEDEE MDDEE MDDEE MDDEE MDDEE VACISE NMDEE VACISE NUKCEEE NUKCEEE NUK LNA LL VACISE LL VKQEEE NL LL VACISE LL VINDEE LL LL LL LNDHDN LNDHDN LNDHON LNDHON	T	DELLVEDMENAG DELLSLENT DELLSLENT DELLSLENT DELLSLENT DELGSENT NEFGLSDWL DEFGLSDWL DEFGLSDWL KEDFTDUS	EDELLFLN SKDVMVST SDDFFDGL EVHISDEI DDDFFDGL SDDFFESL SDDFFESL SDDFFESL SDDFFESL GLPS-DLM SAQFVDGW	- I DADAEADA - I DADAEAA - E LHN - GIRGSQI UND FI C LE - E E LAC - DE LS QLT - DE LS QL	SF	- EPMIPSLEDV RHPATTSSSR DCFP GSAT DCFS - DCFS - FP	VVDE FLNSPWLT VVDE FLNSP ALAA VVCE FLSPWV GNSAFASFDLS- - DMDFSVGFS DHFFASSNIDR DFFFASSNIDR DFFFASSNIDR DFFS-ALAIPS DFFS-ALAIPS DFFS-ALSID DFFS-ALSID CHAPF - GLECAH - VKDRDD - VKDRDD
3JJWR(Y-11628 ZIWWR(Y-11628 ZIWWR(Y-11628 VVWR(Y-11628 GIWWR(Y-11629 GIWWR(Y-11629 GIWWR(Y-11620 GIWWR(Y-11620 GIWWR(Y-11620 GIWWR(Y-11620 GIWWR(Y-11620 ALWWR(Y-11620 ALWR(Y-11620 GIWWR(Y-11620 GIWWR(Y-11620) GIWWR(Y-11620) GIWWR(Y-11620) GIWWR(Y-11620) GIWWR(Y-11620) GIWWR(Y-11620) GIWWR(Y-11620)	E E S S S S 	- VAAS VLLEDAE TSADE FVL VEDHL - SAME FVL VEDHL - SAME DELOFYS LVQQSTSVQSKD - I AKSESTEE DMEDL QGEKSSSREKED - V QGDKSSSREKED - V QGDKSSSREKED - QQD - SADESSVAVQEM - V - QQD - NDDVQETNGDEDM - SDDLLIM - LF	EEEEEE EEEEEEE NOD LD E SQE IR E VEDEE - D MD DEE - E MD DEE - E MD DEE - E ND DEE - E VAETSTH VKGEEER NN LK CQE VMMEEEEEE CN DEN EEEEEEE NN LK CQE VMMEEEEEE CN DEN D LK CQE VMMEEEEEEE CN DEN CO D D D D D D D D AF L	EEEE	DELLVEDMENNG EDELLVEDMENNG EDELLSEDTVV DEGLSEDTVV DEGLSENTV DEGLSEDTVV DEGLSEDVL DEGLSDNV AGAGAIEGRRE KAVKGCDHDL KEEDELD EEVEDDEEEDD EEVEDDEEEGDO SKFCRVSSK	EDELLFLN SKDVMVST SDDFFDGL EVHISDEI DDDFFMGL SDDFFEGL SDDFFESL SDDFFESL GLPS-DLM DISMPDII DISMPDII	- IDADAEADA E LHN GRGSQI - TDADAGA - E E FAAGD INNDF I GLE - E E LAC -	SF. -VTG- -AFSVVTG- AFSVVTG- ASVVTG- ASVVTG- ASVVTG- ASVVTG- ASVTG-	E FMM SLE DU MEATTSSE DCFP GSAT DCFP CSAT	VVDE FLNS PWLT VVDE FLNS PWLT VVGE FLS PWVT ONSAPASFDLS - DMD FSVGFFS DH FFM - SF PIP DF FLAISSNID FS - AIAIFS DF FS - AIAIFS DF FS - AIAIFS - OFDELLNS OE - NL - YOMAMEF - OLE COH - VKDRDD - PVDRDD
22mWR(Y-11622) 22mWR(Y-11622) 22mWR(Y-11622) VVWR(Y-11622) GmWR(Y-11622) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620) GmWR(Y-11620)		VAASVLEPAEN SADEFVLEVEPHL SADEFVLEVEPHL SADEFVLEVEPHL VAGSTSVEVEPHL AKSESTEEPMED GEKSESREEKED GEKSESREEKED GEKSESREEKED GEKSESREEKED GEKSESREEKED VO SADESVAVOEN NDVGETKOEPM OG NDVGETKOEPD NDVGETKOEPD SEEPPETGUEPD EEDTK CEDVL	GEEDVISLF. VIGDL DG E VIGDL DG E VIGDL DG E VIGDL DG E MNDDEE - D MNDDEE - E MDDEE - G NDDEE - G NDGE - G VIGESER VIGESER ONET DG DDAFL P AMS - DS IFLG LT	- EEEE 	- DE LL VEDMENAG - EDLLSLSDTVV - EDKLMERNN DE LGFSDTAV - DE LGFSDTAV - DE GLSSDWL - DEFGLSDWL - DEFGLSDWL - DEFGLSDWL - DEFGLSDWL - DEFGLSDWL - DEFGLSDVL - DEFGLSDVL - DEFGLSV - SVFGVSK - VFFDGSSGV - VFFDGVSK - VFFDGVSK	EDELLFLN SKDVMVST EDELLFLS SDDFFDGL DDFFMGL SDDFFEGL SDDFFESL SDDFFESL GLPS-DLM DISMPDII SAQFVDGW	- IDADÄEADA - IDADÄEAA - ELHN GRÖSQI LNNDFFLGLE - EELAEQ - EELAEQ - EELSGLT - DELSGLT - DELSGLT - DELSGLT - DELSGLT - SSDLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLI	SF	E PMP SLEDU M ATTSSER DCF DCF GSAT DCF DCF DCF DCF F	V DE F LNS P WLT IS DE LGAAP AAM
3 JAWRY, 11823 ZIWWRY, 11823 ZIWWRY, 11823 AUWRY, 11823 WWRY, 11823 GIWWRY, 11823		- VAAS VLLEDAEW - LSE TELRTESM - SAME - DELOB VY - SAME - DELOB VY - VASSTEVENED - SAME - DELOB VY - VASSTEVENED - SAME - SAME - DEMEDIA - OCKSESSREEKED - OCKSESSREEKED - VADESSVAVOEM - VADESSVA - VADESSVAVOEM - VADESSVAVOEM - VADESSVAVOEM - VADESSVAVOEM - VADESSVA - VADESSVA	GEEDVISLF. VODLDGE. VODLDGE. VEDEF.D. VEDEF.D. MDDEE.E. MDDEE.E. MDDEE.E. NDDEE.E. VEGEF.N. MDEE.E. UNDEE.T. VKOEER. UNDHONIFSRIQ DNETDGDDDAFL LNDHONIFSRIQ DMETDGDDAFL I. AMS.DSIFLEIT	T S N Z	DELLVEDMENAG DELLSLENT DELLSLENT DELLSLENT DELLSLENT DELGFSDTAV DEFGLSDML DEFGLSDML DEFGLSDML NEFGLSDML NEFGLSDML NEFGLSDML NEFGLSDML NEFGLSDML DEFGLSDML SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFG SMFC SMFG SMFC	EDELLFIN SKDVMVST SDFFDGL EVHISDEI DDFFHGL SDDFFEGL SDDFFESL SDDFFESL SDDFFESL SDDFFESL SDFFESL SDFFESL SDFFESL SDFFESL SDFFESL SDFFESL	- IDADAEADA - IDADAEAA - ELHN GRGSQI UNNDFFLGLE - EELAEQ - EELSQLT - DELSQLT - DE	SF	- E MMPSLE DU RH ATTSSER DCF GSAT DCF GSAT DCF CF DCF F	VVDE F FLNS P WLT VVDE FLNS P AAM ····LGAAP AAM ····UGE FLS P PWVT GNS AP AS FD LS ····DMD FS VG F DH FP - SF P I P ·····DMD FS ···································
32HWR(Y-11623) 22HWR(Y-11623) 22HWR(Y-11623) 44WR(Y-11623) 44WR(Y-11623) 63HWR(Y-	S S S S		EEEEEEE VODLDGE VODLDGE VEDEE VEDEE MDDEE MDDEE MDDEE WAETSTH VKAEESTH VKAEEER NULK GOEVMMEEEEEE NNLK MDLE SAESTH VKAEEEER NNLK GOEVMMEEEEEEE NDH ANS - DSIFLOEIT AMS - DSIFLOEIT AMS - DASY ELK	EEEE T	DELLVEDMENAG EDELLSLBDTVV EDGKLNERNNEN EDGKLGFEDTAV NEFGLSDTVV NEFGLSDNV DEFGLSDNV DEFGLSDNV GAGGAIEGRRLBN KEODFDU EVEDDEGV KEOFD EVEDDESGG SKFCRVSK FDROGVSTL FSRFGLSS	EDELLFIN SKOVMVST EDELLFIS SDDFFDGL EVHISDE DDFFMG SDDFFEG SDDFFES SDDFFES SDDFFES SDDFFES SDFFES DFFFES DFFFFS DFFFS DFFS DFFS	- IDADAEADA - IDADAEAA - TDADAEAA - ELHNO EFLEFAACOU LINND FFLGLE - EELTA - EELTA - EELSU - EE	SF	- E MMP SLE DOV MH ATTSSER DC E M - C SAT - DC F - C SAT -	VVDE FLNSP WLT VVDE FLNSP AM
30-WWRY-11823 2110/WRY-11823 2110/WRY-11823 2110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823 3110/WRY-11823			EEEEVLSLF. VEDEE-D VEDEE-D VEDEE-D MDDEE-E MDDEE-E MDDEE-E NDDEE-G VKAETSTH- VKQEERN NKK ELNDHONIFSRLO METDODDAFL MS-DSIFLOIT AMS-DSIFLOIT AMS-DSIFLOIT AMS-DSIFLOIT AMS-DSIFLOIT ANS-DSIFLOIT ANS-DSIFLOIT	- EEEE EEEE 	- DE L L VE DME MAG - E DL S L S DT VV - E DAK IN ERNMEN - E DL S L S DT VV - E DAK IN ERNMEN - DE G L S DWL - S DV S T - S KF G C VS S K - F DR DG VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - HE D L IN - S KF G C VS S K - S KF G C VS S K	EDELLFIN SKOVMVST EDELLFIS SDDFFDEL EVHISDE DDFFME SDDFFES SDDFFES SDDFFES SDDFFES SDDFFES SDDFFES SDDFFES SDDFFES SDFFES DFFS DFF	- IDADAEADA - IDADAEAA - ELHN GRGSQI - TDADAEA - EEFAACO - EEFAACO - EELAEQ - EELSQLT - DELSQLT - DELSQLT - DELSQLT - SSSCFFSSL - SSSLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLI - SSDLEKLA - SSSLEKLA - SSSSLEKLA - SSSLEKLA - SSSSLEKLA - SSSSLEKLA - SSSLE	SF	- E MINSLE DV - MINSSER - MINSSER - DCF - C - C - C - C - C - C - C - C	VVDE FLNS PWLT VVDE FLNS PWLT VGE FLS PPWT ON 5 A AS FDLS - DMD FS VOF S DM FFY SFTI DF FTA ISSN IDR DF FS - AIAIRS DF FS - AIAIRS DF FS - AIAIRS DF FS - AIAIRS - OFDELLNS CE COM - OLE COM - OLE COM - VECOM - VEC
3 JAWRY, 11823 ZIWWRY, 11823 ZIWWRY, 11823 ZIWWRY, 11823 VVWRY, 11823 GIWWRY, 11823 GIWWRY, 11824 GIWWRY, 11824 GIWWRY			GEEDVISLF. FEEEEFF. VODLDGE. VODLDGE. VEDEF. MDDEE. MDDEE. MDDEE. MDDEE. MDDEE. VEOBER. NMDEE. NMDEE. LINDER. QEVARTSTH. VK0EER. NILK. LINDHON IFSRLQ DMET DE DDDAFL I. AMS. DSIFLE. AMS. VAT. EDEEEEE	T S N E E	D E L VE DME MAG T E LA E SP NG 3 G D E L S L S DT VE D C C S DT A D E G L S DME D E G L S DME S DT D D E G L S DME S D D S D D D D	EDELLFIN SDDFFDGL SDDFFNG SDDFFNG SDDFFSS SDDFFSS SDDFFSS SDDFFSS GL SDFFSSS SDFFSSS SDFFSSS SDFFSSS SDFFSSS SDFFSSS SDFFSSS SDFFSSS SDFFSSS SDFFSSS SDFSSSS SDFSSSS SDFSSS SDFSSS SDFSSSS SDFSSSS SDFSSSS SDFSSSS SDFSSSS SDFSSSS SDFSSSS SDFSSS SDFSSSS SDFSSS SDFS	T DADAEADA ELHN GRGSQI INNDFFLGLE EELAEQ EELAEQ EELT DELSQLT DELSQLT SGSTFPSTI SGSTPSLEKLI DVDDLLIPN CSDEKKLI DVDDLLIPN SCDGVGRS SCDGVGRS	SF	T PMSSLFDU	VVDE FLNSPWLT VVDE FLNSPADA OVGE FLSPWVT ONSAFASFDLS - DMDFSVGFPS DHFFAISSNIDR DFFFAISSNIDR DFFFAISSNIDR DFFFAISSNIDR DFFFAISSNIDR - OFDELLNS GLECOH - VKDRDD - KSSDD - KSSDDD - KSSDD -
32HWRY(X-11822) 22HWRY(X-1822) 22HWRY(X-1822) 44WRY(X-1822) 44WRY(X-1822) 63HWRY(X-1822) 63HWRY(X-1822) 63HWRY(X-1822) 63HWRY(X-1822) 63HWRY(X-1823) 63HWRY(- VAAS VLLEDAE H - LSTTTLRTSM - SAME - DELOG YVE - SAME - DELOG - SAME - SAME - SAME - SAME - SAME - SAME - SAME - SAM	EEEEEEE VODLDGE VODLDGE VEDEE VEDEE MDDEE MDDEE MDDEE VEDEE MDDEE VEDEE MDDEE VACTSTH VKAETSTH VKAESUN VKAETSTH VKAESUN OMENDE SUNTATION MADEE SUNTATION ONENDE VKAETSTH VKAESUN OUENDE SUNTATION OUENDE AMS AMS AMS VANS VANS VANS CELK VAT EDEEEEE EDEEEEE	т	D E L L VE DME MAG E DE L L VE DME MAG E DE L L VE DME MAG E DE L L S DT VV P E DGKI. ME RNNTN - DE C F DT AV NE F G L S DMV L DE F G L S DMV L DE F G L S DMV L DE F G L S DMV L ME F G L S DMV L C E D F D VL DD D E V E D D E E D VL F R VF NG S GG - S K F C RVS SK - F D R DG V S T L - F S R C L VS C - F D L VE DME MAG - L VE DME MAG	EDELLFIN SODFFDGL EVHISDEL SDDFFEGL SDDFFEGL SDDFFESL SDDFFESL SDFFESL SDFFESL SDFFESL FBOFFSL SDFFESL	- I DADAEADA - I DADAEAA - T DADAEAA - T DADAEAA - EE F AAGO - EE LAEQ - E LA	GARL L VVALSKLHEF SF. -VTG	T PMSSLFDV MSSLFDV PMSSLFDV MSSLFDV MSSLFDV MSSLFDV MSSLFDV	VUDE F LNS PWLT VUDE F LNS PWLT VUCE F LS PWVT ONSAPASFDLS - DMD FSVQFFS DHF FF - SF PIP DF FTAISSN IDR DF FS - AIAIS DF FS - AIAIS DF FS - AIAIS - OFDELLNS CE CQH - OLE FSN - SF VE - SF VE - SF VE - SF VE - OLE FLS PWWE - OUE FLS PWWE
32HWRY:11823 22WWRY:1823 22WWRY:1823 22WWRY:1823 22WWRY:1823 22WWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1824 SWWRY S			EEEEDVISLF. FEEEEFE VGDLDGE CQEIRCE CQEIRCE CALERCE MDDEE MDDEE MDDEE MDDEE MDDEE MDDEE MDDEE MDDEE MNDKEEER NIK DIK GOEVMMEEEEEEE NIK DIK GOEVMMEEEEEEE AMSOBIFER AMSOSIFER AMSOSIFER EEEEEE EDDEEEEEE EDDEEEEEEE EDDEEEEEEE VYREGDD	T	D E L VE DME MAG T E LAR PROS E D E L VE DME MAG T E D L S L S DT VE D G K L S L S T VE D G K L S L S T VE T E G L S DME D E L VE DME MAG D E L VE DME ME ME D E L VE DME ME ME ME ME ME ME ME ME ME	EDELLFIN SDDFFDGL SDDFFGG SDDFFGS SDDF	IDADAEADA IDADAEAA INNDFFIGLE EELFAGOU INNDFFIGLE EELFG EELFG EELSGLT DELSGLT DELSGLT DELSGLT SOLFFISTI FSDLFKLI SOLFKLI SOLFKLI SOLFKLI ALDDRG ALDDRG ATSEDAA	SF	T PMSSLFDV T PMSSLFDV PMSSLFDV PMSSLFDV	VVDE FLNS PWLT VVDE FLNS PWLT VGE FLS PPWT ON SA PAS PDLS - DMD FSVO FS DM FFY SFTI DFFTAISSN DDF FFS - AIAIS DFFS - AIAIS DFFS - AIAIS DFFS - AIAIS DFFS - AIAIS - OFDELLNS OE - OLE COH - OLE COH
328WRY,11822 228WRY,11822 228WRY,11822 228WRY,11822 328WRY,11823 328WRY,11824 348WR			GEEDVISLF. FEEEEEEE. VODLDGE. VODLDGE. VODLGE. VEDEE. MDDEE.E. MDDEE.E. MDDEE.E. MDDEE.G. VAETSTH. VKGEER. MLK. UNDHONIFSRLQ DMETDGDDDAF. I. AMS.DSIFLGIT AMS.DSIFLGIT EDDEEEEE. EDDDEEE. EDDDEEE. EDDDEEE.	т	D E L VE DME MAG T E LAR SPNOS C D E L VE DME MAG C D E L VE DME MAG C D E L S DT VV P D G L S D T VV P D G L S D T VV P D G L S DMV L D E F G L S DMV L D E F G L S DMV L C D E F G L S DMV L C D E G L VE D E S G G C S C D F L VE D ME MAG D E L VE D ME MAG C D L L VE D ME MAG S R M L N D MOMAGEN	EDELLFIN SDDFFDG EVHISDE UDDFFMG SDDFFGS SDDFG	- I DADAEADA - I DADAEAA - ELHN GRGSQI INNDFFLGLE - EELAEQ - EELSQI - DELSQIT - DELSQIT - DELSQIT - DELSQIT - DELSQIT - SDEKLI - CSDIEKLI DVDDLIJFN - CSDEKLI - C	6 A L L	TPMSSLFDV	VVDE F FLNS PWLT VVDE F FLNS PWLT ON SAPASFDLS
32HWRY,111122 22HWRY,11122 22HWRY,11122 22HWRY,11122 GHWRY,11122 SAWRY,11122 ZHWRY,11122		- VAAS VILLEPAE - LSTTPLRTSM - SAME - DELO - SAME - SAME - SAME - SAME - SAME - SAME - SAME	EEEEDVISLF. FEEEEEE OQEINGE OQEINGE MDDEE MODEE MDDEE MDDE MDIK DLK MDENDODAFL MS <dsiflet< td=""> AMS<dsiflet< td=""> QELK QE VYERED VYEREE DDEEE QUEDEEE QUEDEE QUEDEE QUEDEE QUEDEE QUEDEE QUEDEE QUEDE QUEDE QUEDE QUEDE</dsiflet<></dsiflet<>	т	DE L L VE DME MAG E DE L S DT VV E DG KL NE RNNTN - DE G L S DMV L DE G L S DMV L ME G L S DMV L DE G L S DMV L C DE G L S DMV L C DE G L S DMV L C DE G L S DMV L C S D S S C DH D L - VE DD S S G C - VE DD S S G C - VE F D S S G C - VE F D S S G C - VE F D S S G C - VE S S S S S S S S S S S S S S S S S S	EDELLFIN SDDFFDGL SDDFFDGL SDDFFESL SDDFFESL SDDFFESL SDDFFESL SDFFFESL SDFFFESL SDFFFESL SDFFFESL SDFFFESL SDFFFESL S	- I DADAEADAE - I DADAEAA - T DADAEAA - EE HARG - EE LAEG - E	G AN L	T P MSSLFDV	VUDE FLUS PWLT VUDE FLUS PWLT GUSAPASPOUS OVER FLUS PWVT GUSAPASFDLS - DMDFSVOF S DHFFF-SFFID - FTAISSNIDR DFFF-SFFID - FTAISSNIDR DFFS-AIAIRS - OFDELLNS - OFFELSIN - VEDROD - VED
32HWRY:11823 22WWRY:1823 22WWRY:1823 22WWRY:1823 22WWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1823 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 GWWRY:1824 SbWRY		VAAS VLLEAAE LSTTPLATESH SADEFVLYVEOHLA SADEFVLYVEOHLA AKSESTE EEDHEDLA GEKSESREEEDHEDLA GEKSESREEKED GEKSESREEKED GOKSESREEKED SADESVAVGEH NDVGETROEDEH SEUT EU SEUT SE	E E E D V I S L F F E E E E E E E VE D L D G E . E MU D L E . E MD D E E . E MD D E E . E MD D E E . G VE D E . G MD D E . G MD D E . G MD D E . G MD D E MD D E .	т. 	- DE L L VE DME MAG - E DE L L VE DME MAG - E DL I S L B DT VE D QK I.M E RNNTN - DE L G F DT AV - DE G L S DWL - DE F G L S DWL - NEF G L S DWL - DE L L S DWL - DE L VE DME MAG - DE L L AP VO DD - DE L AP VO DD - DE L L AP VO DD - DC AP VO DD - DC DD - DC AP (AP (AP (AP (AP (AP (AP (AP (AP (AP	E DE LLFIN E DE LLFIN S DD FF DG E UHISDEI S DD FF FS S DD FF SS S DD FF SS S DD FF SS S DF FS S	T DADAEADA ELHN GRGSQI INNDFF LGLE - EELFAQOU EELT- DELSQLT. DELSQLT. DESQLFRST FSDELFRNLE SSDLEKLI CSDLEKLI CSDLEKLI SSDLEKLI CSDLEKLI ALDDRG FRKTWLF GODANALDG FPKFWLF	SF	T PMSSLFDV	V DE F FLNS P WLT V DE F FLNS P WLT VGE F FLS P FVVT GNS AP AS F D LS - DMD FS VGF P S DH F P - S F P I P DF F A I S N I DR DF F S - AI AI T DF F S - AI AI T DF F S - AI AI T OF FS - AI S I DF F S - AI S I DF F S - AI S I DF F S - AI S I OF FS - AI S
328WRY-14E28 228WRY-14E29 228WRY-14E29 24WRY-14E29 44WRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 GWWRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 228WRY-14E29 238WR		- VAAS VLLEDAEM - LSE TELEVEDHL - SAME - DELO VY - SAME - DELO VY - VAS STEVENER - SAME - DELO VY - SAMA DE SAME - DELO - YY - SADE SY DEMEDIA - SADE SY AVOEMS - KET YORLLOYV - VOOL - DEDN - VOOL - DEDN - VOOL - SADE SY - KET YORLLOYV - VOOL - SADE SY - SE FET DELO - SE FET DELO - SE STALEN - SE STALEN - SE STALEN - SE STALEN - SE STALEN - SE STALEN - STALEN	GEEDVISLF. FEEEEFF. VODLDGE. VODLDGE. VEDEF. MDDEE. MODEE. MDDEE. MDDEE. MDDEE. MODEE. MODEE. MODE. VODE. VODE. VODE. </td <td>с с с с с с с с с с с с с с с с с с с</td> <td>D E L VE DME MAG T E L AR SP NO 3 G T E L S E S T V T E G L S DT V T E G L S DT V T E G L S DM L T E G S G L S DM L T E G S G L S DM L T E G S G S G S G S G S S G S S S S S S S</td> <td>E DE LL FI N S ND YM VST E DE LL FI S S DD FF FO S DD FF F S S DD FF S L S DD FF S L S DD FF S L S DD FF S L S D FF S S D S DF FS S S D FF S S S D S S S D FF S S S D S S S S S S S S S S S S S S</td> <td>- I DADAEADA - I DADAEAA - EE HAN GRA - EE FAAGO I NND FF LG LE - EE LAEO - EE LAEO - EE LAEO - EE LSOLT. - DE LSOLT. - DE LSOLT. - DE LSOLT. - TF SNELF NLE - SODEKL DVDDL LI FN. - CSDLEKL DVDDL LI FN. - CSDEKL - CSDEK</td> <td>6 A L L</td> <td>T PMSSLFDU PMSSLFDU PMSSLFDU PMSSLFDU PMSSLFDU</td> <td>VVDE FLNSP WLT VVDE FLNSP WLT VVGE FLSP PWVT ONSAPASFDLS- - DMDFSVGFS OHF FP - SFP I DF FTAISSN IDR DF FTAISSN IDR DF FS - AIAIPS DF FS - AIAIPS - CLE COH - CLE COH - VNS IMR - SF - SF VNT - SF VNS - DF FLS FVNT - DF FLS FVNT</td>	с с с с с с с с с с с с с с с с с с с	D E L VE DME MAG T E L AR SP NO 3 G T E L S E S T V T E G L S DT V T E G L S DT V T E G L S DM L T E G S G L S DM L T E G S G L S DM L T E G S G S G S G S G S S G S S S S S S S	E DE LL FI N S ND YM VST E DE LL FI S S DD FF FO S DD FF F S S DD FF S L S DD FF S L S DD FF S L S DD FF S L S D FF S S D S DF FS S S D FF S S S D S S S D FF S S S D S S S S S S S S S S S S S S	- I DADAEADA - I DADAEAA - EE HAN GRA - EE FAAGO I NND FF LG LE - EE LAEO - EE LAEO - EE LAEO - EE LSOLT. - DE LSOLT. - DE LSOLT. - DE LSOLT. - TF SNELF NLE - SODEKL DVDDL LI FN. - CSDLEKL DVDDL LI FN. - CSDEKL - CSDEK	6 A L L	T PMSSLFDU PMSSLFDU PMSSLFDU PMSSLFDU PMSSLFDU	VVDE FLNSP WLT VVDE FLNSP WLT VVGE FLSP PWVT ONSAPASFDLS- - DMDFSVGFS OHF FP - SFP I DF FTAISSN IDR DF FTAISSN IDR DF FS - AIAIPS DF FS - AIAIPS - CLE COH - CLE COH - VNS IMR - SF - SF VNT - SF VNS - DF FLS FVNT - DF FLS FVNT

1	Т	7	1	
	I	٦,	1	-

IIE2-

314

					acidic-5			
	600	610	620 6	30 640	650	G 60	670 680	690 700
GmWRKY-IIE3a		· FC - D <mark>g</mark> v <mark>p</mark> yyr	PISKMNNY	LQ	L DQ <mark>GLFAEFGE</mark>	I EKDQ	SVNLLS <mark>PP</mark> -Q	GFDDQQRESNA LDSF
GmWRKY-IIE3b		• • • F S • D <mark>G</mark> I <mark>P</mark> • • Y K	PSLMDHQS-TN	NQ	Q S <mark>E D F F A E L G</mark> E	EADP	- L D L L F T Q <mark>G</mark> F <mark>G (</mark>	GAD - DQRESKA LD <mark>P</mark> F
VVWRKY-IIE3a		• • • F S • E <mark>G</mark> F <mark>P</mark> R S Y K	PAMPD · · · · · ·	• • • • • • • • • • • • • • • • • • •	Q S E D F F A D L G E	IEAD <mark>P</mark> · · · · · · · ·	- LNLLFA Q <mark>G</mark> FS!	S <mark>G</mark> DHEERESKA LD <mark>P</mark> F
AtWRKY-IIE3a		FDYNDVENTYR	PELLQEF	QH	Q <mark>P E</mark> D F F A D L D E	L E G D S	- L T M L L S H S S G (<mark>gg</mark> nmenktt i <mark>p</mark> D <mark>v</mark> fs
AtWRKY-IIE3b		· · · DDDDNQ I A <mark>p</mark> yr	PEL - HDH	QH	Q <mark>P D D F F A D L E E</mark>	LEGDS	- L S M L L S H <mark>G</mark> C <mark>G (</mark>	GD <mark>G</mark> - KDKTTAS D <mark>G</mark> IS
SbWRKY-IIE3a	QQVMDTTALAVVDHN	I E L M D Q V F <mark>G</mark> - E S Y K	PM-IPEAG		H S D D F F S D L A E	L E S D P	- MSLIFSKEYMEAK <mark>P</mark> SS <mark>G</mark>	<mark>g</mark> dr <mark>g</mark> hhqekam - Skdld <mark>p</mark> lf
ZmWRKY-IIE3a	Q - VMDTAA - AAVDHS I	ELMDQVFSCESYK	PMMIPEAGGH -		S S D D F F S D L A E	LESD <mark>P</mark> · · · · · · · ·	- MSLIFSREYMEAK <mark>P</mark> S <mark>GG</mark> I	D <mark>g</mark> dhr <mark>g</mark> hqekamt - Skdld <mark>p</mark> lf
ZmWRKY-IIE3b	Q - VMDT - · · AVADHN	ELMDQVFS-ESYR	PM-IPEAG		H <mark>P D</mark> D F F S D L A E	LESD <mark>P</mark> · · · · · · · ·	- MSLIFSKEYMEAK <mark>P</mark> S <mark>GG</mark> I	D R <mark>g</mark> hheka I Skolo <mark>p</mark> lf
OsWRKY-IIE3a	ERVMDTTAAGVVDHS -	ELMDHVFS-ESYK	РМ- I <mark>Р</mark> ЕТ <mark>G</mark>		Q <mark>P D D F F A D L A E</mark>	LESD <mark>P</mark> · · · · · · · ·	- MSLIFSKEYMEAK <mark>P</mark> S <mark>GG</mark> I	D HAQEKAM AKELD <mark>P</mark> - F
OsWRKY-IIE3b	DAAAAASATAAAVEHS	DLMQQMFS-QSYR	PM-IPEAAAGG		HHD D F F A D L A E	L E S D P	- MSLIFSKEYMATNYK <mark>P</mark> A	- <mark>G D P A G</mark> K E M N A V D K <mark>G L D P</mark> A Y
BdWRKY-IIE3a	- AGDDTSAATLLDDH -	- L L Q Q M F S Q S C Y R	PM-IPDQSQAG	<mark>G</mark> • • • • • • • • • • • • • • • • • • •	HH <mark>D</mark> DFFADLTE	L D S D P	- VSLIFSAEYMETCRQ <mark>GG</mark> I	E <mark>P</mark> DKEKQAVDNK DLQLD <mark>P</mark> F -
SbWRKY-IIE3b	EKGMDH - DASVLLDHG	DLMQQMFSQSYYR	PM-IPEAGGGG	<mark>G</mark> · · · · · · · · · · · GHI	HADDFFADLAE	L E S D P	• MSLIF • • • • • • • • <mark>PGGG</mark> I	D <mark>PG</mark> KEKEMM <mark>P</mark> NK SL <mark>G</mark> AD <mark>P</mark> LF
ZmWRKY-IIE3c	EKVIDHHDASGWLDHG	DLMQQMFSQSYYR	PM-IPEAAGGG	LI	HADDFFADLAE	L E S D P	• MSLIF••••••• <mark>P</mark> AGGI	E <mark>PG</mark> KENK S L D A D <mark>P</mark> L F
ZmWRKY-IIE3d	EK <mark>g</mark> tdh - Dasvlldh <mark>g</mark>	DLMQQMFSQSYYR	PK - IPEASGG -	HI	HADDFFADLAE	LESD <mark>P</mark> · · · · · · · ·	- MSLIF <mark>PGG</mark> - I	D <mark>PG</mark> KK <mark>G</mark> RMSSESL - <mark>G</mark> M <mark>G</mark> TD <mark>P</mark> LF
GmWRKY-IIE3c	·	· · · FS · D <mark>G</mark> <mark>P</mark> Y C K	PISKMKNS		L D <mark>QGLLAELGE</mark>	IEFAD	5 VN L L S · · · · · · · · P P · Q	GFDNQKRESNT LDSF

IIE3-







OsWRKY-IIIA1a SbWRKY-IIIA1a ZmWRKY-IIIA1a BdWRKY-IIIA1a	120 130 A L V	140 TRSFSSAS <mark>GGG</mark> A- RASPA-DLAPSP- RSSP-DLASPL- TAAAMT-	150	160 170	180 	190 200 LTSCT <mark>PSP</mark> LSDGSDH 5DATSDHHH DASDHPFF VSHQC	TARWT 210 9 F FRTT NAK <mark>RKTTA</mark> H <mark>P</mark> FRAAGAS PK KRKATA AAASAS - AS PK KRKATA QP FRQP NEK K <u>RKATA</u>	220 230 IRWT S - QVRV SAAGG G IRWT SQQVRV SAAGG G IRWT SQQARV SAAGG G IRWT S - QVRV SAAGG	IIIA1-
BdWRKY-IIIA2a OsWRKY-IIIA2a ZmWRKY-IIIA2a SbWRKY-IIIA2a ZmWRKY-IIIA2b SbWRKY-IIIA2b ZmWRKY-IIIA2c	 	KSGHFG RSGHFDG TSSNLDGA- TSSNLDGA- TSSSSSGLDGA- ITTS			KRSAG HKKRSAAAVA RRQKR <mark>P</mark> AGDAG ARRKRSASDAA RRKRSASDAA RR KRSAGDAG SGLAG SSS <mark>P</mark> H	LDS <mark>P</mark> TPSPLK AGDLDSATPSPLS LASPLSATPTS LASPLSATPTS LASPLSATPTS ARKRSAVP GGASDAPP	QVSDMTFK <mark>P</mark> NKKRKT-b DVSDL <mark>P</mark> FKATKKRKTST DVTDGPFQSTKKRKKN DVTDGPFKNNTKKRKVb GVTDGPFKNNTKRRVb FVKG-TKKRKTB FVKG-TKKRKTB	IERRKQQVRVI PSGEG-A EKKRHQIRVSSTG-G-V IEQRRQRVSSAGG GQRRERVSSAGG IDKKRHEVRVSSAGG IDKKRHEVRVSSAGG-G IDKKRHEVRVSSAGG-G	IIIA2-
OsWRKY-IIIA3a SbWRKY-IIIA3a OsWRKY-IIIA3b ZmWRKY-IIIA3a ZmWRKY-IIIA3a BdWRKY-IIIA3c BdWRKY-IIIA3c ZmWRKY-IIIA3c ZmWRKY-IIIA3c OsWRKY-IIIA3c	A	ASCFPPPE. RSCCVDSPASCS. RSCCTEAAAACR. VSCCAEGRPG. VSCRECARP. VSCRECARP. TSNSLDCSPAAC. GTGTPRCCG. RSCCADLPD.	rqqa T		HPPPAAGN. GSPRSDGGN. GAPESPPS.ADGS. GAPESPPS.GGDGS. PSPPSGGDGS. PSPPSGGDGS. APESPPSGGDGS. APESPPSGGDGS. APESPPSGGDGS. APESPPSGGDGS. APESPPSGGDGS. APESPSGGDGS. APESPGGGGS. APESPGGGGS. APESPGGGGGS. APESGGGGGGS. APESGGGGGGS. APESGGGGGGS. APESGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	AG AG S - P R S G S D H AG - E L S - R S G S D A G G E L R S G S D Q A G G E V R S G S D D H A D V S S S R G A V L D S D Q A G G G L NV D V V V K G V H S C A	RDAAF KKGG AAQ3 - KRKKG RCR - ANAAG CHKKKT RGRGN - AAGGCHKKKT RGRGN - AAGGCHKKKT RGRGN - AAGGCHKKKT TGNAF GQGKORKRKT TGNAF GQGKORKRKT - R KRGT KRGT GVHSDSQRSSAA <u>SSKS</u>	RVSAVQD IAKVRRQVRVTSVQD FCVRRQLRAASVQD PKWSSQVRVSAVQD PKWSQVRVSAVQD PKWSQVRVSAVQD PKWSSQVRVSAVQD PKSSIQVRVSAVQD PKSSQVRVSAVQD SKWSTQVRVSAVQD SKWSTQVRVSAVQD SKWSTQVRVSAVQD SKWSTQVRVSAVQD SKWSTQVRVSVAQD SKWSTQVRVSVAQD SKWSTQVRVSSVAD FCVRRQLRAASLQD CVRRQLRAASLQD LRE	IIIA3-
AtWRKY-IIIB1a VVWRKY-IIIB1a Gii WRKY-IIIB1 AtWRKY-IIIB1b AtWRKY-IIIB1b AtWRKY-IIIB1b Gii WRKY-IIIB1b Gii WRKY-IIIB1d Gii WRKY-IIIB1d Gii WRKY-IIIB1d Gii WRKY-IIIB1d Gii WRKY-IIIB1d Gii WRKY-IIIB1d Gii WRKY-IIIB1d Gii WRKY-IIIB1d Gii WRKY-IIIB1d	· · · · · · · · · · · · · · · · · · ·	K C G G F V S E	L P H S	AL AL LA LA VA	L P G S S IR		E I D G KD S K N V F K K R K V A G K · · · D V F K K R K I L D P N · · · A F K K R N T L I D P N · · · A F K K R N T L I D N · · · A F K K R N T L I D N · · · A F K K R N T L O N · · · · A F K K R K M L D H O P I F N S K K K M L D H O P I F N S K K K M L D H O P I F N S K K K M I D H O P I K N S K K K M M D H O P V K D S K K K K M D O P O P C N C S S K R K T L A · · · · · A L K K R K M A · · · · · A L K K R K M D D S D O E P U V I K S S K K M	EKNTEKVKVFVATE Q - RWTKQI KVT IQTV	IIIB1-





Appendix 7 continued LETP 220 230 160 170 BdWRKY-IIID1a - AGGSLFGAEPAVDDKS----LVRKKMLNRAISPGEIRMKEEEQTRPVRS-----VA--TKRERK----DGKRSRSLVTNV RdWRKY-IIID16 -QKRRRK----GSKQSRSLVTNV ZmWRKY-IIID1a -ShWRKY-IIID1a IIID1-. . . KTRRRK----QQATTSTMVTTV - RR<mark>KNK</mark>K - - - - KRQQSKCLVTTV . . . -KTSRRK----LQTTTSTMVTTV-HxGS OSWRKY-UD2a -VSKRRK - - - - NAEHTGSTVAQA GNKRRK - - - - NAQHT GSIVTQA GNKRRK - - - - NAQHT GSVMTQA IIID2-GNKRRK NANH ISSVVTOT KNKRRK----NAGHTGSVVTAT PIKEILs IIID3-JARAP ARAP ALAP ARABA BERGED ALAPA ARABA BERGED ALAPA ARABA BERGED ALAPA ARABA BERGED ALAPA ARABA BERG IIID4-IIID5-SWAT IIID6-*ZmWRKY-IIID6*b · · · · · · · P AT RGAKRRSSVP AQGKK · · · · AA<mark>SSSWATLTAV</mark> · · · · · VT P ISV I KAE PHQLSPP SP ASAD · · · · · · P EAT RGAKRRSSVP AQGKK · · · · · AA<u>SSSWATLTAV</u> · IIID7-

IIID8-	

\$bWRKY-111D82	LDVSDAA <mark>G</mark>	· · · · · · SDD <mark>G</mark> AR <mark>P</mark> R <mark>P</mark> QS <mark>G</mark> · · · · · ·	SSC <mark>G</mark> NK <mark>RKQ</mark> SSSRRSQR <mark>P</mark> SDKKIT
BdWRKY-IIID8b · · · · · · · · ·	AK <mark>PVG</mark> S		- KRKAAAAAARRDTSRR RQSSGFTKVV
BdWRKY-IIID8c	AKAGAG	EWSSDVQSEVTY <mark>GG</mark> S <mark>GG</mark> -GK	RKSGAGGDSSRGACRRRT · OOSSVVTKTM · · ·
SbWRKY-111086		- FEFFASEVOSDVTCAGTAG-SS	KRKAAGGGDKRASCRKRGQQGSSVVTKNI
7m MPKV-1110#2			KPKAAGGAGKPAAACPPPAOPSSVVTKSV
USWRKY-IIID89	LDKAAVS	· · AAGGEGEVQSEVICGGGAS · AG	GRRRAP AAD KRAAN CRRRI QQSSGNSVVV



	WRKY DNA-binding domain	
240 250 260 270 Atwrky-suic1a P 1 Y DGYL WKK Y GGK SIKK SNHOR S Atwrky-suic1a P 1 Y DGYL WKK Y GGK SIKK SNHOR S Atwrky-suic1a P 1 Y DG FL WKK Y GGK SIKK SNHOR S Atwrky-suic1a P P L DD G FT WKK Y GGK TIKT S L Y OR C Atwrky-suic1a P R P DD G FT WRK Y GGK TIKT S A HK R C	280 290 300 310 320 330 340 350 YYRCSYNKDHNCEARKHEOKI KONPPVYRTYFGHHTCKTEHNLDA KNPPVYRTYFGQHICQLHQAYAT YYRCAYKDANCYATKRVOMI QDSPPVYRTYFGQHICQLHQAYAT YYRCAYKDANCYATKRVOMI QDSPPVYRTYFGQHICKA YYRCAYKAGNCYATKRVOMI QDSPPVYRTYFGQHICKA	IIIC1-
A tWRKY-HIC 24 - ALVD DGFAWRKYGOKTIKTSPHORW CPWRKY-HIC 22 - ALVD DGFAWRKYGO KDILKA NHPRS GmWRKY-HIC 22 - ILIED GYTWRKYGO KMTSO SKYLRS GmWRKY-HIC 22 - APID DGHWRKYGO KEILNA KFPR GmWRKY-HIC 22 - APID DGHQWRKYGO KEILSA KFPRN	YYRCAYAKDONCDATKRVOKI	
G MWRKY-IIIC29 - T PKLDCHQ WRK YG QKC I LNAKY 5 RN VYWRKY-IIIC29 - N FHDDQYA WRK YG QKC I LNAKY 5 RN AtWRKY-IIIC29 - K 55 ED RYA WRK YG QKC I LNTT FPRS AtWRKY-IIIC29 - K DDN HA WRK YG QKC I LNAK FPRS G MWRKY-IIIC27 - Q T DDN HA WRK YG QKC I LNSG FPRS G MWRKY-IIIC29 - Q T T DDN HA WRK YG QKC I LNSG FPRS G MWRKY-IIIC29 - Q T T DDN HA WRK YG QKC I LNSG FPRS G MWRKY-IIIC29 - Q T T DDN HA WRK YG QKC I LNSG FPRS G MWRKY-IIIC29 - Q T T DDN HA WRK YG QKC I LNSG FPRS G MWRKY-IIIC29 - Q T T DDN HA	YYRCTHKYDDINCQAIKQVQRI	IIIC2-

Appendix 7 continued	WRKY DNA-binding domain	
240 250 260 270 BdWRKY-IIID1a PHYDGHQ WRKYGQKNINGMQH3RS BdWRKY-IIID1a PHYDGHQ WRKYGQKNINGRQH3RS BdWRKY-IIID1a PHYDGHQ WRKYGQKNINGRQHPRS SDWRKY-IIID1a PHYDGHQ WRKYGQKNINGRQHPRS SDWRKY-IIID1a PHYDGHQ WRKYGQKHINNGRQHPRS SDWRKY-IIID1a PHYDGHQ WRKYGQKHINNSKHPRS OSWRKY-IIID1a PHYDGHQ WRKYGQKHINNSKHPRS SDWRKY-IIID1a PHYDGRQ WRKYGQKHINNSKHPRS SDWRKY-IIID1a PHYDGRQ WRKYGQKHINNSKHPRS SDWRKY-IIID1a PHYDGRQ WRKYGQKHINNSKHPRS SDWRKY-IIID1a PHYDGRQ WRKYGQKINNNSHPRS SDWRKY-IINNSHNSHS SSDWRKY-IINNSHNSHS SDWRKY-IIID1a PHYDGRQ WRKYGQKINNNNNNNSHS SDWRKY-IINNAKHPRS SDWRKY-IINNAKHPRS SDWRKY-IIID1a PHYDGRQ WRKYGQKI I NNNNNNNSS SSDWRKY-IINNAKHPRS SDWRKY GQKI I NNNNNNNSS SDWRKY-IIID1a PHYDGRQ WRKYGQKI I E GAMYRS SDWRKY IINNAKHPRS SDWRKY III01a PHYDGRQ WRKYGQKI I E GAMYRS SDWRKY-IIID1a PHYDGRQ WRKYGQKI I E GAMYRS SDWRKY III01a <td< th=""><th>280 290 300 310 320 330 340 350 YYRCTYK - ERNCSATKT VOEQD- HNRSSFSYGDETVKYTVYYGHTCNGE HNRSSFSYGDEAVNTVWYGHTCNGE HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSSSTGTGEANF HNRSSFSSSSTGTGEAS HNRSSFSSSSTGTGEANF HNRSSFSSSSSTGTGEANF<th>IIID1-</th></th></td<>	280 290 300 310 320 330 340 350 YYRCTYK - ERNCSATKT VOEQD- HNRSSFSYGDETVKYTVYYGHTCNGE HNRSSFSYGDEAVNTVWYGHTCNGE HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSYGDEAVNTVWYGHTCNGEA HNRSSFSSSTGTGEANF HNRSSFSSSSTGTGEAS HNRSSFSSSSTGTGEANF HNRSSFSSSSSTGTGEANF <th>IIID1-</th>	IIID1-
OsWRKY-IIID28 · PHNDGHQ · · · · WRKYGGKWI SRAKHSRS ZmWRKY-IIID28 · PHFDGYK · · · · · WRKYGGKWI SKAKHSRS SBWRKY-IIID28 · PHFDGYQ · · · · · WRKYGGKWI SKAKHSRS SBWRKY-IIID28 · · PYFDGCH · · · · · · WRKYGGKWI SRAKHSRS	YYRCANSKVQGCPATKTVQQMDSSGN.GTSKLFNVDYYGQHTCRG. YYRCANSKDQGCLATKTVQQKESDGS.GTVSLFDVEYYGHICKK YYRCANSKDQGCLATKTVQQKESDGST.GTVRLFNVEYYGQHICKK. YYRCAYSKEQGCPATKTVQQKENDGN.GTVRLFNVNYYGQHICKS.	IIID2-
BelWRKY-IIID3a ··PENDOEFH······WRKYGEKKILNAAFPRL SolwRKY-IIID3a ··PENDHEH······WRKYGEKNILYAEVPRL ZmWRKY-IIID3a ··PENDREH······WRKYGEKNILYSEVPRL	- <mark>YYRCGY</mark> SDEHK <mark>CPAKK</mark> YVQQQDN <mark>G</mark> DPP	IIID3-
OSWRKY-IIID4a · · <mark>PYDDGY</mark> Q · · · · · · · WRKYGQKKINNTNFPRS ZmWRKY-IIID4a · · PHSDGYQ · · · · · · WRKYGQKRITKSPFPRC SbWRKY-IIID4a · · PHCDGYQ · · · · · · · WRKYGQKRITKTQFPRC	- <mark>YYRCSY</mark> HRERR <mark>CPAQKHVQ</mark> - QRD <mark>GDD</mark> VPALHVVVYTHEHTCLQGAPAEL <mark>P</mark> DAATNGGAAA - <mark>YYKCS</mark> FHRERSCRATKQVQ QCSDGDQ HPPQYVVMYFNEHTCD TAAWEPAAA - YFKCSFHRERNCRATKQVQ QCSNDD <mark>PP</mark> QYVVIYFNEHTCDD TAAWDP	IIID4-
BelWRKY-IIID5a · PYKDGSQ · · · · · WRKYGGKN I RNR I FARY ZmWRKY-IIID5a · PYKDGYV · · · · · WRKYGGKN I QNCNFVRY	- <mark>YYKCMY</mark> SHER <mark>GCRAKKGVO</mark> QQDNSSDHRPMFLITYVNEHTCQQLH <mark>P</mark> TENIS	IIID5-
OSWRKY-HIDG& - PHEDGFQWRKYGEKKIQGTHFTRS OSWRKY-HIDG& - PHEDGYQWRKYGEKKIQGTHFTRS Sowrky-HIDG& - PYDDGYEWRKYGEKKINGTLFTRS ZmWRKY-HIDG& - PYDDGYEWRKYGEKKINGTLFTRS ZmWRKY-HIDG& - PYDDGYEWRKYGEKKINGTLFTRS	- Y FRCTYRDDRGCQATKQIDQKDKNDPP MFQVTYSNEHTCTTTRLIN. Y FRCTYRDDRGCQATKQIQQEDKNDPP	IIID6-
OSWRKY-IIID7a - AT MDDKFLWRKYGGKEIKNSKHPR BawRKY-IIID7a - GT TEDGFAWRKYGGKDINGCRHPRL OSWRKY-IIID7b - GT TDGFIWRKYGGKDINGCKHPRL SbWRKY-IIID7a - GT TVDGFIWRKYGGKDINGKHPRL ZmWRKY-IIID7a - GT AVDGFIWRKYGGKDINGHKHPRL ZmWRKY-IIID7b - ET MEDGFIWRKYGGKDINGHKHPRL OSWRKY-IIID7b - ET MEDGFI	YYRCEYKDDHGCTATKQVQQSETA	IIID7-
BdWRKY-HIDBa - DTLVDGHVWRKYGGKE I ONSPHPRS SbWRKY-HIDBa - ANLEDGHVWRKYGGKE I ODSPYPRS BdWRKY-HIDBb - KNVEDGOSWRKYGGKD I ONSEHPKS BdWRKY-HIDBb - NDLDDGOAWRKYGGKY I HNSKHPRA SbWRKY-HIDBb - KDLEDGHSWRKYGGKE I ONSKYPKA ZmWRKY-HIDBa - RDLEDGHAWRKYGGKE I ONSKYPKA OSWRKY-HIDBa - KNLDDGOAWRKYGGKE I ONSKHPKA	- YYRCTHKSDOGCNAK RQVQACEADPSKYAVTYYGEHTCSDYPSSTAPMIIVAAAADEDDDN	IIID8-

10 20 20 20 20 44 40<	PHAQ			
BUMPRY/MATE PHAOTLOSIS AAAD TVNT DTNT/TT ENR AAAD TVNT AAAAL TVNT DTNT/TT ENR AAAD TVNT AAAAL TVNT AAAAL TVNT AAAAL TVNT AAAAAL TVNT AAAAAAL TVNT AAAAAL TVNT AAAAAL TVNT AAAAAL TVNT AAAAAL TVNT AAAAAL TVNT AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	360 370 380 390 400 410	420 430 4	40 450 460	470
Demonstry maday LTV WT E LAAMAG OSWET - TTTT F ST ST AR DEF OR VY KAR ST F T DOUCH REVEWERS F T DOUCH REVEWERS F T T T T T T T T T T T T T T T T T T	OSWRKY-IIIA1a	PENR AP PESCPVRGAS SPWS PVG PESCPVRGAS SPWSLASSV PP SMA	<mark>PP</mark> L Q Q Q Q Q Q Q V S A S P V A S D S Y <mark>G L G G</mark> S D S N G C L Q H Q G V S P C P V P S D S N G C M - H Q D V S P C P V P	GYGDWR GYGDWA GYREWT AFTEWP
downry/mage et pont tarkellog		_	ENWG	-
zummery/mass IFFENDED ADDIGULES LG ADLIVITYTETEG LE ADATUCLSSTILLS ADLIVITYTETEG LE ADATUCLSSTILLS ADLIVITYTETEG LE ADATUCLSSTILLS	<i>BelWRKY-IIIA2a</i>	FSS <mark>PAP</mark> - · · · · <mark>PPP</mark> QERYSKFSA <mark>P</mark> ST LSSPVSG - · · · LAPPDQHN <mark>P</mark> FSAPST	PDDCCFGRGVSVELSPATSG G PENR - LAAAASSAASPATSDSM A	SSHLAMA
SAMERY JIIA22, AQ. OAAAD AGAA E YNDYKET NI TVYKTEOTT VE A EGOORGIVD A AT F CFSST AT TATASMEL VE LED FFSA F STSNNWG. VS AT SDSNNVVS	ZmWRKY-111A22 •••••• PPEHRPDAHGHLRSLG••••AGLTVKTETEEGLP••••••••AAAPLC SowRKY-111A22 •••••• PAEHNADAHAHLGTLS••••AGLTVKTEGLPT••••••••••AATPLLY ZmWRKY-111A22 ••••••• AAEHNPDARGHLGTLS••••AGLTVKTEGLPDFP•••••••••••••••••••••••••••••••••••	LSAST <mark>PP</mark> VSG <mark>GGCPCP</mark> CLA <mark>P</mark> ST LSASTPLAPASTA LSASSTPPST	PENWG - VSPASSNSNHAAS	ylpfe ylpfd frpfe IIIA2-
OSWRYYHIA32 AFEHAQGE OQRQSSLAADTEGIHQQVAEMAAFILFTTAQGVDD GYFSFISFANSDCQFSS OSWRYYHIA32 ALDOTHQQT DO <aselaadteglaaallemntysfatvagaaa< td=""> SFFISFANSDCQFSS OSWRYYHIA32 AASACKONTER OCANGCATAVLFSSCAFAALLEMNTYSFATAAAAAACA SFFISFANSDCQFSS OSWRYYHIA32 AASACKONTER OCANGCATAVLFSSCAFAALLEFNTYSFATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</aselaadteglaaallemntysfatvagaaa<>	SBWRKY-IIIA2B ·······AG ··QAAADALAPEYNGKPGTNLTVKTEGPTVEPAEQQVQGWD·APTPFC ZIIIWRKY-IIIA2c ······TAGRQAAADAQAPEYD····ANLMVKTEGPAEQPVGGWDAPTPFL	FSST <mark>P</mark> ATATASWCLV <mark>P</mark> ELS <mark>PP</mark> FSA <mark>P</mark> ST FSST <mark>P</mark> ACLV <mark>P</mark> ERS <mark>P</mark> -FSA <mark>P</mark> ST	SNNW <mark>G</mark> - VS <mark>P</mark> ATSDSNHVVS SENW <mark>G</mark> - VS <mark>P</mark> ATSDSNHVVS	· · F <mark>PP</mark> FE · · F <mark>PP</mark> FE
Severy: Jinase Lo Lo Churche and Ender and the full stand of the analysis of t	05WRKY.#1433		GYESELSPANSDCOESS	
ShwRkY-JIIA3b OP AVGEOTTS PS of FAATAVLE PS LR AS. NKT T: TG ADAATS SRF JTT GC VS. TAS FLS FATE ES LVS SSSS. YAVGGGGG Zim WRKY-JIIA3b Oq QOTO TGT SP G. FAAFAVLE PS LR AASNKT TG AADAATS SRF LAT GC T.V. TAS FLS FATE ES LVS.GGGAGG AVGGGGG Zim WRKY-JIIA3b HAF AT DOATS PG. CVVADE AALE PS LR AASNKT TG AADACATS SRF LAT GC T.V. TAS FLS FATE ES LVS SSSS GGGAGG AVGGGGG ShwRKY-JIIA3b LAE INTO FAAGA. Dea SS VGAGG COME FOR TE EAR APOG LLE PAR POG PLAP POG PL	SBWRKY-IIIA33LQLQ <mark>PG</mark> HGAQEDQ.AS <mark>P</mark> LALETEGL-RAALLE <mark>P</mark> HT <mark>P</mark> YSFA OSWRKY-IIIA35AASAEQKHQ <mark>PTPP</mark> QEQNAVS-VAFTSMAVVNAS	TVAGAGAG	ASAGADFAGWC TSS <mark>P</mark> FVS <mark>P</mark> AMSDCQIS	· · YEL <mark>GG</mark>
Zmwrky-Jillaja Inter to do a coss to covar on or end and build as sket at score of the sket of available as the sket of avail	SBWRKY-IIIA3B ····································	R F V T T <mark>G</mark> C V S V	TASPFLSPATPESQLVSSSSSG - YA	
Bewrey-Midse Lonlorh Dorgtaaaaagge Qgggdecle Veagg Mowe Zmwrey-Midse Lollorensagge Stoldgalrs Stoldgalrs Gwrey-Midse DVE IRFOSE ADPHLF Stolgalrs Gwrey-Midse DVE IRFOSE ADPHLF DDTTVSTRHDSNTDASS	2011 WRKT-1111A39		FPSNKPEPAVVAHAASN	
OSWRKY-HIA32 ⁶ DVEPIRPOSILDDDMCT DDTTTVSTRHDSNTDASS AtWRKY-HIA32 ⁶ FEQNRVHVTEQSEDMKETKSEEVMISLEDLENKKNIFRTFSFSNHEIENGVWKSNLFLON FVEDLSTATSOSAITSEVLS VWWRKY-HIB13 FEQNQQQQQEAEQIVSGFOTG IKVKTEDLDSFEQ-FLVF SSFSFFLSD FAESMIENSLMOGFSP.SSLSATS ESNYFPVS GMWRKY-HIB13 HQAQAQAQALLSLRKG IKVKTEDLDSFEQ-FLVF FFFFLSTN IK NESHVFSSVUENFTS.FSVNSPAAS GIGHFSVS GMWRKY-HIB13 -QQANILANLEQQASLLISIKKG IRVQTENLDSFEQ-FLVF FFFFLSTN IK NESHVFSSVUENFTS.FSVNSPAAS GIGHFSVS GMWRKY-HIB14 -QQANILANLEQQASLLISIKKG IRVQTENLDSFEQ-FLVF FFFFLSTN IK NESHVFSSVUENFTSSPVNSPAAS GIGHFSVS GMWRKY-HIB15 -QQANILANLEQQASLLISIKKG IRVQTENLDSFEQ-FLVF FFFFLSTN IK NESHVFSSVUENFTSSPVNSAAS GISHFSVS MWRKY-HIB15 -QGANILANLEQQASLLINKG IRVQTENDSFEG-FLSTN IK NESAVFSSVUENFTSSPVNSAAS GISHFSVS AtWRKY-HIB15 -QUARKY INTATA SSFFFTSSSCRNEN- NNSSP IFFENN-IVSCFSS SSFSSSS ISHFSSS GWWRKY-HIB16 -AITAK INTKTEGL ITVTTOGL GKASST SSFFTSSSSCRNEN- NNSSP FFENN-IVSCFSS SSISATS SSNSSSS	B4WRKY-MIA38*	V <mark>G</mark> AGG		
AtwRKY-IIIB1a IFEGNRVH VTEQSEDMKETKSEEVMISLEDLENKKN IFRTFSFSNHE IENGVWKSNLFLGN. FVEDLS ATSGSA ITSEVLS. FSNFF LSD. SETVDKI. FASS SA ITSEVLS. FSNFF LSD. SETVDKI. FSNFF LSD. SETVDKI. FASS SA ITSEVLS. FSNFF LSD. SETVDKI. FSNFF LSD. SETVDKI. FSNFF SNVLENFTS. FSVMSAAS. GISHFSVS MWRKY.IIIB1a OGNILGNLEGGSDILLSLRKG. IRVGTENLDS POQSLVF FRFF LSTN. IK. HESTYF SNVLENFTS. FSVMSAAS. GISHFSVS MWRKY.IIIB1a OGNILGNLEGGSDILLSLRKG. IRVGTENLDS POQSLVF FRFF LSTN. IK. HESTYF SNVLENFTS. FSVMSSAAS. GISHFSVS MWRKY.IIIB1a OSLOGAS DULLSLRKG. IRVGTENLDS POQSLVF FRFF LSTN. IK. HESTYF SNVLENFTS. FSVMSSAAS. GISHFSVS MWRKY.IIIB1a OSLOGAS DULLSLRKG. ILVGTENLDS POQSLVF FRFF LSTN. IK. HESTYF SNVLENFTS. FSVMSSAAS. GISHFSVS MWRKY.IIB1a OSLOGAS DULLSLRKG. ILVGTENLDS POQSLVF FRFF STSFS STREM. HNSFF. STYFS IN GGT STYHVHVGSS SPF OF MAKKY.IIB1a OSLOGAS DULLSLKSN. ITVRT GGL. SFVFF FFF OSSFF DTFP. FYNYGT ING. FFF HVESS SA ISSENT SNSFF STREM. HNSFF. STYFS STREM. HNSFF. SFYFY SNSFF STSFS STREM. HNSFF. SFYFY SNSFF STSFS STREM. HNSFF. SFYFY SNSFF SNSFF SSTRS SNSFF SSTRS SNSFF SSTRS SNSFF SST. SFYFY SNSFF SST. SFIFT SNSFF SST. SFYFY SNSFF SST. SFIFT SNSFF SST. SFYFY SNSFF SST. SFYFY SNSFF SST. SFYFY SNSFF SST. SFYFY SST. SFYFY SST. SFYFY SST. SFYFY SST. SFYFY SST. SFYFY SST. SST. SST. SNSFF SST. SST. SNSFF SST. SST. SST. SST. SST. SST. SST. S	OSWRKY-IIIA3¢ [↑] ·····DVE <mark>P</mark> IR <mark>P</mark> QS	<mark>g</mark> ldddmct	DDTTT <u>V</u> STR <u>H</u> DSNTDASS	
GNWRKY-JIB11 GWRKY-JIB11	Atwr.Ky.IIIB1a	VWKSNLFL <mark>G</mark> N·····F	VEDLSPATSGSAITSEVLS	
MWWRKY-MIB1a QSLDQQKPPNEQLLNNLRTG LRVQTENLDFQNQ.SFV FG MWWRRY-MIB1a TVAVN YONLRAS LTVRTGL GSEAFSFPVTSF LYVTPSINGGFTYHHVSSGP AUWRRY-MIB1b TVAVN YONLRAS LTVRTGL GSEAFSFPVTSF LYVTPSINGGFTYHHVSSGP CJWRRY-MIB1a HQAKE QEILLNLRQD LTVRTGLDD GSEAFSFPVTSF FF FF CJWRRY-MIB1a HQQKE QEILLNLRQD LTIKTEGLNK.NDMPPSFSFFSTSFSCRNREN-HNFSP PTFENN-IVSCFFSP SFISPATP ESNYRSAL CJWRRY-MIB1a HQQAE GSLAFFRSI LSVNTDNLDN.GOMAYAFTFFSTSFCRNREN-HNFSP FTFSNLSULYQT FLISPATS ESNYFFSF GMWRKY-MIB1a HQQAE SLAFFRSI LSVNTDNLDN.GOMAYAFTFFSTSFCRCMRQDN-HSLIP AFNKONJONLLG(YSP.TFISPATS ESNYFFSF GMWRKY-MIB1a YTQDGS LTVKTDNI COMARYFFTFSTSFCRCMAQDNYHDLFPS LVLENDPFSTLSQLYQT.HLLSPTTP ESNYFFSF GMWRKY-MIB1a YTQDGS LTVKTDNI SAFGCMAQDNYHDLFPS LVLENDP FSTLSQLYQT.HLLSPTTP ESNYFFSF GMWRKY-MIB1a YTQDGS LTVKTDNI SAFGCMAQDNYHDLFPS LVLENDP FSTLSQLYQT.HLLSPTTP ESNYFFSF GMWRKY-MIB1a HRAQA QOESLAFFRSI LSVNTDNLNN.GDMAYP FTF STSFGCMKQDN.HSLIP WALENDSFLSQLYQT.HLLS	GmWRKY/IIIB1aH-QHQQNILQSHEQQQNELLLSLRKG····LRVQTENLDSPEQ·PLVPFRFPLSTN·· GmWRKY/IIIB1aH-QHQQNILQSHEQQQSLLLSLRKG····LRVQTENLDSPQQQSLVPFRFPLSTN··		SPVLENFTS - PSYVSPAAS G SPVLENFISSPSYMSSAAS G	IGHFSVS
Attrack Attra Attrack Attrack	MWWRKY-IIIB1a ·····QSLDQQK <mark>PP</mark> NEQLLNNLRT <mark>G</mark> ····LRVQTENLDFQNQ-SFV <mark>P</mark> FG······ AWWRKY-IIIB1b ·····TVAVN····YQNLRAS····LTVRT <mark>GG</mark> L····· <mark>G</mark> SEAFSF <mark>P</mark> VTS <mark>P</mark> ··	NHVLE	SSFAENFNS - <mark>P</mark> SYVS <mark>P</mark> ANS G YESIN <mark>GGG</mark> TFYHHV <mark>G</mark> SS <mark>GP</mark>	I SH F SMS
VVWRKY/HB16 HQQAH IENPLNYS SMASTSYGCMKSET.SLFQGS AFNKONYDNLLGCYSP.TFISATS DSNCFSLS GmWRKY/HB16 HLAQA GGESLAKFRSILSVNTONLON.GOMAYAFTFFSTSFGCMKQDN.HSLIP LALENDSFLSDLYQT.HLLSPTTP ESNYFPSP GmWRKY/HB16 YTQGG LTVKTDNI TAPSASFGCMAQDNYHDLFPS LVLENDPFSTLSQLYQT.HLLSPTTP ESNYFPSP GmWRKY/HB16 HRAQA GGESLKFRNI LSVNTONLNN.GOMAYFTF STSFGCMKQDN.HSLIP WALENDSFLSDLYQT.HLLSPTTP ESNYFPSP GmWRKY/HB16 HRAQA GOMQAQADMLLS.FGSG. INVTY ESNYFPSP WALENDSFLSDLYQT.HLLSPTTP GmWRKY/HB16 LC GQAQADMLLS.FGSG. INVTY ESNYFPSP WALENSFLSDLYQT.HLLSPTTP GmWRKY/HB16 LC GQAQADMLLS.FGSG. IRVYFG SSSTSSS. IGSENEDNM. ESNYFPSP GmWRKY/HB17 KNQP LLEKTQQQNADMLLS.FGSG. IRVYFG SSSTFALENNAMLANDANI LANPART ESNYFPSP GmWRKY/HB16 KNQP LLEKTQQQNADMLLS.FGSG. IRVYFG SSTFALENNAMLANDANI LANPART ESNYFPSP GmWRKY/HB18 KNGP LLEKTQQQNADMLLS ISTSFALENNAMANANANANAN INTERNAMANSSFSP.VFISPATS ESNPFCLS GmWRKY-HB16 KNGP LLEKTQQNP EGIFF ISTSFALENNAMANANANANANANANANANANANANANANANANANA	AtWRKY-IIIB1c ·······AITQKPKDILESLKSN····LTVRTDGLDD····GKDVFSFPDTPP·· CpWRKY-IIIB1a ·········HQQKE <mark>S</mark> QEILLNLRQD····LTIKTEGLNK··NDMPPSFSFPSTSFSC	······································	IY <mark>GT ING EFGHVESS</mark> I - IVSCFFS <mark>P</mark> SFIS <mark>P</mark> AT <mark>P</mark> E:	SNYRSAL
GmWRKY-HIB19	VWWRKY-111816	MKSET SLFPQ <mark>P</mark> S AFNKDNY MKQDN - HSLIP MAODNYHDIFPS LALEND	'DNLL <mark>G</mark> CYS <mark>P</mark> TFISPATS D SFLSDLYQT HLLSPTTP E BFESTISOT SSISANTP E	SNCFSLS SNYF <mark>P</mark> SP IIIB1-
GWWRKY-HIB1Y ···KNQPLLEKTQQQNPEGIFT-FETE····LKVKTEELET-KEDIFPWFSFSSSS···IGSENEDNM······LPDTMIENHFMESFSP··VFISPATS·····ESNPFCLS GWWRKY-HIB1Y···KNQPLLEKIQQQTEVIFT-FETE····LKVKTEELET·KEDIFPWFSFSSS···IGSENEDNM·······LPETMFENHFMESFSP··VFISPATS·····ESNPFCLS AWRKY-HIB1Y GWWRKY-HIB1YQIKNQPRQEKIEQQPPETFFT-FGSSS··GLEVKIEDIDH.HHVLHYPSFSSSFNLENNNAYMLQMKRY-HIB1YQIKTON-STSFSSDLGTSINYHFPASG GWWRKY-HIB1YQIKNQPRQEKIEQQPPETFFT-FGSSS··GLEVKIEDIDH.KEGIFPSFCFSSPSI···VGSENGDDNNSSI·····FSYTMIENNLMESFSP··AFISPTTSDI···ESNTFC-·	GmWRKV-IIIB1e············HRAQA_QESLTKFRNI·····LSVNTDNLNN··GDMAYPFTFPSTSFGC VWWRKV-IIIB1e······LGPQQNQQQADDMLLS·FQSG····LRVVTEGLDTPSDQAFPPFCFTSTSN··	MKQDN - HSLI <mark>P</mark> WALENE IKVEE <mark>PG</mark> FS <mark>P</mark> SMMD	SFLSDLYQT - HLLSTTIP E NNIVGNFPT SFVSPAAS R	SNYF <mark>P</mark> SP SNYFSMS
AtWRKY-IIIB18 ········LEPNOTQEHGNLDMVKESVDNYNHQAHLHHNLHYPLSSTPNLENNNAYMLQMRDQNIEYFG······STSFSSDLGTSINYNFPASG······ESNTFC·· GWWRKY-IIIB18 JIKNOPROEKIEQQPPETFFT-FGSSS··GLEVKIEDIDH-KEGIFPSFCFSSPSI··VGSENGDDNSSI······FSYTMIENNLMESFSP··AFISPTTSDI····ESNTFC··	GmWRKY-HIBH1 ··KNQPLLEKTQQQNPEGIFT·FETE···LKVKTEELET·KEDIFPWFSFSSSS··· GmWRKY-HIBHg··KNQPLLEKIQQQTPEVIFT·FETE···LKVKTEELET·KEDIFPWFSF <mark>PSP</mark> S···	IGSENEDNM ····································	NHFMESFS <mark>P</mark> VF I S <mark>P</mark> ATS E NHFMESFS <mark>P</mark> VF LS <mark>P</mark> ATS E	SN <mark>P</mark> FCLS SN <mark>P</mark> FCLS
	AtWRKY-IIIB14LEPNQTQEHGNLDMVKESVDNYNHQAHLHHNLHYPLSSTPNLENNNA GmWRKY-IIIB14 QIKNOPRQEKIEQQPPETFFT-FGSSS-GLEVKIEDIDH-KEGIFPSFCFSSPSI- GmWRKY-IIIB14 QIKNOQOFIED PREFEETFFT-GSS-GLEVKIEDIDH-KEGIFPSFCFSSPSI-	YMLQMRDQN EYF <mark>G</mark>	TSFSSDLGTSINYNF <mark>PASG</mark> NNLMESFS <mark>P</mark> AFISPTTSDIE NNLMENES	SNT F C

	360	370	380	390	400	410	420	430	440	450	460	470	
AtWRKY-IIIC1a AtWRKY-IIIC1b AtWRKY-IIIC1c AtWRKY-IIIC1d AtWRKY-IIIC1f AtWRKY-IIIC1f		IFIAGQI FPIDTSI FGVHDN VAVHDD FAVHDD FAVHDD	D <mark>P</mark> LDDFKS1 D-FEEHEG TYG TYG TYS 3TYG	⁻ QMIRF <mark>G</mark> K SHMIRF <mark>Q</mark> H SEMINFDQ SEMIKFDQ STMIRFDQ SKMIKFDY	D D D D D D D D D D D D D D D D D D D	QDQEKESRSN <mark>G</mark> F3 NISFSSSTSN	SLSVKHEED		IKEQAIDO LRQHO TIQ TIDE TIDE SIDS	YRE I TSNDQ NHQDR I KDE QAVLME - DE QA I TME - DE IQV I TVE - EN QE I TMEDKD	DCQDVIEEYLS YMK <mark>P</mark> VIAEDWS AN-HIMNQEYC AIDHIMNQECC SAEHIMNQECC TDDHILNY	SPSGSYP SQWMS INDYLVD INDFSVD INDFLVD INDYLVD	IIIC1-
CpWRKY-IIIC22 GmWRKY-IIIC2 GmWRKY-IIIC2 GmWRKY-IIIC2 GmWRKY-IIIC2 GmWRKY-IIIC2 GmWRKY-IIIC2 GmWRKY-IIIC2 GmWRKY-IIIC2 CpWRKY-IIIC2 CpWRKY-IIIC2 CpWRKY-IIIC2	9	L VNE S I VEF FS MME F AF I LD PMS I LD PMS I LD SNNI I MVENS' KT CDHI MVT HSE' VVT HSE' MATE FF I MSD SD F I GSSY MLEF AS	QLGFS 	ITNDSIKR SILLSFDN SKFLSFDN SKFLSFDN SILLSFSNNAD ADSEKTLAASTA OPDAND GPDAND GPESN VVSNSSD NMVGSE	DHT P FS I S 	F SSSLKN LGSKQENF LTSKQECF FLI FFT SKQECF FLI FFT FKQEC FFT FKQEC FFT KQECE I PNI QEQENNTSSVTA QEQENNTSSVTA NEHDST IQSQSL NEHYST IQSQSL NEHYST IQSQSL AIDHHISSPNL TPSSNQHB IPEEVQEMKE	S		E FSS FSSF FFSSF VT VT WKK MKK VT VT VT VT VT VT VK VT VK VT VK SSSF VV V VV SSSF VV SSSF VV SSSF VV SSSF VV V V V	SKQEVLTTNN SSILASSTKHE SSLSSSTKLE VKRECKEEV - KREDTEKI NNHTNNNK ECONPADES EC	GLSSSSNYLI PGEVINNEHS, PMEVIQDDHI PFSISSNYLI ISMICDIVLI STSSNHYLI ISMICDIVLI KDYVEGSSTG SDLTDANFWSI SDLTDANFWSI SDLTDANFWSI DEGVIWK PMEVINEDHI	SPDITTF AQNKLSTF AQNQLSSS ISSDLTFD ISSDLTFD ISSDLTFD ISLUWQ IDLSLVWQ IDLSLVWQ IDLSLVWQ IDLSLVWQ IDLSLVWQ IDLSLVWQ IDLSLVWQ IDLKDFELS INFLNSA IEALDSF /HNQLPSS	IIIC2-

	360	370	380	390	400	410	420	430	440	450	460	470	
BdWRKY-IIID1a	N I SN	G - NVDL PQLV				SMDLDQ	TVEEMARMTTY-		QAQEFD	DEGD	L		
ZmWRKY-IIID1a	· · · · · · · · I I HA	V ST VQL PQL V	A			SMDLDF SVDL <mark>g</mark> Q	GTEMDQTTATR-		AGGVQESD	DEADH	L		
ShWRKY-IIID1a	HS	ISIVQL <mark>P</mark> QLV	s	<u>.</u>		CMDL - G	NME I AQTSS		DVQDP -	EAD	_ L		
OsWRKY-IIID1a	· · · · · · · <mark>P</mark> AAL	ADDH <mark>V</mark> VVEAS	Q I STDS	HCQS <mark>P</mark> SSSSDL	QAAE VHA <mark>G</mark> NSS	QCSNISVTO	S <mark>P</mark> SVVVEDCN		к <mark>г</mark> г	. DMM <mark>P</mark> AADE L	т		
OsWRKY-IIID1b	<mark>P</mark> AAS	ADDHVVVEAS	Q I STDS	H C Q S <mark>P G</mark> S S S S E	LQAAAHA <mark>G</mark> DSS		SSSVVVEDCN		· · · · · · · · · KLL	. DML <mark>P</mark> AADE L	т		
ShWRKY-IIID1h	· · · · · · · · NN · ·			 	SETNSQSSIST	ISTD <mark>P</mark> YG	RET PSLDGNKL -		LDKSADLI				IIIID1- II
BdWRKY-IIID1c		DTDGTVVDS -		RCSS		N S V T C	TSVAVIDHHRQ-		· · · · SSLESSLL	DMAEEDLAT	E		
OsWRKY-IIID1c	NMSN	APLHVVETST	TQ S ST	TCCSDDL <mark>G</mark> DYS	QKMENMHT <mark>P</mark> EL	AEVCSDEL	SYHAIIGAEH		SAL	. <mark>G</mark> LED	_ E • • • • • • •		
SbWRKY-IIID1c *	• • • • • • • • DS L	EA <mark>P - V</mark> ILETT	T T V V A <mark>P</mark> S N S	AATANTTTYTD	S I VV <mark>P</mark> TMSDH <mark>G</mark>	SCSTITIT	GTES <mark>P</mark> AIS <mark>G</mark> DD -		ITCWSSTS	G G A S S S D Y N	Y		
OsWRKY-IIID1d*	E <mark>P</mark>	S L <mark>P P V</mark> I L D T T	VRTTN	NHQQ <mark>P</mark> AAAE S P	AAT S S	SSSNMVMTS	SET GN		WS <mark>G</mark> QH	IGAYACR - QM	1		
ZmWRKY-IIID1c*	DDAL	EAPPVILETT'	TSVVVRA <mark>P</mark> A	AHT DP VVVVVP	ATATTSAAAAA	5 A C 5 T T V T T 8 8 8 0 0 0 0 0 T T	GTESPAISGDD-			G-SSSSGYS	Y		
03144(7-440)70							<u>o</u> oroozr <mark>o</mark> nee						
OsWRKY-IIID2a	D <mark>G</mark> I A	D <mark>P</mark> YV <mark>V</mark> DTAH -			H S M	E <mark>P</mark> I NQNE CN	IS <mark>P</mark> TLEHEAH -		EVQDERFE	NLCMVQNM	E		
ZmWRKY-IIID2a	DVVN	H <mark>P</mark> CV <mark>V</mark> DTAHY			YSV	<mark>P</mark> I ANQNQS S	SS <mark>P</mark> TFVHNDVY -		<mark>G</mark> I Q D E S F E	NLFMV <mark>P</mark> SI <mark>P</mark>	E		
SbWRKY-IIID2a	DD	HPYVVETTD-			· · · · · Y S A	P I ANYNQSS	SSSMFVHNDVL -	• • • • • • • •	<mark>G</mark> I H D E S F E	N F F M V <mark>P G</mark> M P	E		IIID2-
ShWRKY-IIID2h	D <mark>G</mark> IV	HPHVVGATQ-			DSM	PIVSQNQNS	SS-VFVNTDVH-			SLFMVPDMP	E • • • • • • • •		
USWKK I-IIID2W		C <mark>P</mark> DIVEID3			E CI	551ND <u>R</u> 1A3				NEFAV <mark>F</mark> DM3			
BdWRKY-IIID3a		ANSQVLDFT -			к	A S I S S <mark>P</mark> T M A	AASAAV <mark>P</mark> RLKK-		EEDVVAG -	.			
ShWRKY-IIID3a		GS-Q <mark>V</mark> LDFT-			к	ASLS <mark>P</mark> EEDS	SM <mark>P</mark> VSMHRYS		. .				
ZmWRKY-IIID3a	I S S S S S	GSSQ <mark>V</mark> LDFT -			····к	ASLSS	•••• <mark>P</mark> LH••••						mbs
		-											
OsWRKY-IIID4a	AASPDYFP	A GGETPSSLR	RLRGVGGGG	LQ <mark>P</mark> QFVDHR	AAM	EERERQVLV	SSLARVLQGR		QCYDD-	DDDDD	T		
ShWRKY-IIID4a		LDDLSKSKLQ.	- LLVAROAG	SLLLDER		EEHERRLLV	SSLACVLOOQ			A0501	T		
BdWRKY-IIID5a	<u></u> T T N	TTARNDHFD <mark>P</mark>	<u></u> .		PQHVG · · · · TS	TELENKIMA	KCLANVVIG		RAAA <mark>P</mark> S	swss <mark>p</mark> as		s	
ZmWRKY-IIID5a	•••AAS <mark>PP</mark> TTN	RRWY <mark>Y</mark> R DD.	AARDDD <mark>gg</mark> -	LQFDLSS - F	P RMS GP GG AGG	SAQENQTIV	SCLADVIR <mark>g</mark>		AA <mark>P</mark> S <mark>P</mark> -	- WS <mark>P</mark> VAADA	S D A <mark>G</mark> A A S Y	<mark>G P</mark> A <mark>P</mark> MQ MQ M	IIID5-
			N								u		
OSWRKY-IIIDG	· · · · · · · · NTNN	NPAALHSLTA	N		PNGHPDDDSDD'	TILTKMIKG	EQQAAWLPSPP -		P-DLTTIS	NNFDETPGL	H		
ShWRKY-IIIDGa	I ANN <mark>G</mark> S	S SNN LAAL LA	G C C S N <mark>G</mark> S <mark>G</mark> S	<mark>g</mark> s <mark>g</mark> k <mark>g</mark> lttmtt	TNAR <mark>P</mark> TEHAAA	AAAMNMMKG	EP-PLLLPALI-		DLQQ <mark>P</mark> SAC	F <mark>P</mark> N EQI <mark>P</mark>			
ZmWRKY-IIID6a	T <mark>G</mark> S S S N	NNSH <mark>L</mark> AALLA	DCCD <mark>G</mark> S	GCNKGLTT-TT	ING HAAAA	AAAMNMNMK	QE <mark>P</mark> SLL <mark>PPPPP</mark> -		LAE <mark>PP</mark> SAC	F <mark>P</mark> Y DQM <mark>P</mark>	Q		IIID0-
ZmWRKY-IIID6b	AA <mark>P</mark> NT <mark>G</mark>	G S S N L A A L L A	A <mark>G</mark> CCD <mark>G</mark> S <mark>G</mark> R	GTGKGLT TS	ISAR <mark>P</mark> T <mark>G</mark> H - QH	AAAMNMKQE	E <mark>PRP</mark> LLL <mark>P</mark> L <mark>PP</mark> -		L VE <mark>PPP</mark> AS	F <mark>P</mark> Y <mark>G</mark> HQTM <mark>P</mark>	1		
6-W910/WE-	·								N a a N a N a		.		
USWRKY-IIID7a	DD	AAAMVVDGGE	EEDQLS <mark>P</mark> AQ	MVISFASSN <mark>GG</mark>		DAQNNSETS	HESSP			EKLR <mark>P</mark> CTAA		IEST <mark>PP</mark> A <mark>P</mark> EL	
OsWRKY-IIID78			CROGELMPP	AV INSGASSEA	AAWNMASRE PA	SSLAVERRS	CDGDA		PSFTSOGV	NS. <mark>P</mark> SFSSFV		AGADSSAS	
ShWRKY-IIID7a		T AWQQLG	AA P	AVVDFGSNSWG	SADANNNG		- GSPA		ASMSQGGV	NS - PSASSEV	GFDF··EA		
ZmWRKY-IIID7a		<mark>A</mark> WQQ <mark>GG</mark>	AAAAAVA <mark>PP</mark>	A V V D F <mark>G</mark> S S S W <mark>G</mark>	SAADT SR				s s <mark>P</mark> s q <mark>g g</mark> v	NS S <mark>P</mark> SASSEA	<mark>g vg f g</mark> t q l	AHEWHDTAA	
ZmWRKY-IIID7b	E A	P A P F V I D F <mark>G L</mark>	SAAACD - D <mark>g</mark>	LQL <mark>P</mark> QY <mark>G</mark>	S P WP S C D D D G P	<mark>g p</mark> v e l q t <mark>p f</mark>	PRTSA		D L C S S <mark>P P</mark> E	E E E L RA <mark>G</mark> ACD	VAEFVAEG	STTVTAELM	
0sWRKY-IIID7c	EN	SQH F 🔽 I N F <mark>G P</mark> .	ATASRS - <mark>G</mark> S	P P L L Y D D <mark>G</mark> D D <mark>G</mark>	DVWRETAAT PP	S S R Q S R C S <mark>F</mark>	EGDGE		E S <mark>G</mark> V K M S H	KEE <mark>P</mark> - VDSC <mark>P</mark>	G P S A V S S <mark>P</mark>	ADVVSCSS <mark>P</mark>	
			-		-								
BOWRKY-IIID82				LS	PHQLAKEEGVV	AA5 RMS 5 SV	чсі Т 5 А чете		DDVFSSS	ADP FVQLAA		GSAGRTSDA	
BdWRKY-IIID85	VH	AGSHLICFAP	NAGPASTTT	S VTTNOTO	GMDGAAPGSAS	GLPILKVF	G		DOEEVRSCI	TPGSSAVHS		AAAAGAG	[]
BdWRKY-IIID8c	т	AGCHLISFG-	<mark>Р</mark> Т <mark>Р</mark> ТТТ	TSTTTTAQVGS	GLQSLKRESGG	DQEEIQCRO	AGCAR		····KVEKAMVS	WSFGAVLSS	· · · · · ST I	KNSVDNEDV	
SbWRKY-IIID8b	н	A <mark>G</mark> SR <mark>L</mark> I SFAA	ANNNASAAT	T S T T T T <mark>G</mark> N T T N	QQL · · · · AVL ·	Q <mark>P</mark> LKLEC	. <mark>6666</mark>		EQEEVLSSI	L T <mark>P</mark> A G S S A A A	E AMRN <mark>G</mark> NA	AAAAATTT G	
ZmWRKY-IIID8a		A <mark>G</mark> SR <mark>L</mark> I SFAA.	AN ATTAS	T - TTTT <mark>G</mark> NASN	RQA <mark>G</mark> HK <mark>G</mark> AVL <mark>P</mark>	AE <mark>PP</mark> RQLE <mark>G</mark>	GGGGGGGEREQEL -		ELEEVLSSI	L T <mark>P AG</mark> SSAAT		- AEAVRTAT	
OsWRKY-IIID82	LQ	A	- AA <mark>P</mark> AA <mark>P</mark> VD	А ААА <mark>Р</mark> ТТ S Т	ITTVTA <mark>PG</mark> ··P	LLQ <mark>P</mark> LKVE <mark>G</mark>	GVGSS		DQEEVLSSI	L T <mark>P G</mark> S S A A R <mark>G</mark>	· · · · · · · · ·	GGGGGGVAGP	

430 440 450 460 470 480 490 500 510 520 530 OSWRKY-IIIA12	IIIA1-
BdWRKY-IIIA22 - PPPQERYSKFSAPSTPDDCCFGRQVSVELSPATSQGSSHLAMAAPFRAQSEMDKMVSALALVASAFPAAAFSIDEFDGFQLDDFDVSSFFA OSWRRY-IIIA22 - LAPPDQHNFFSAPSTPENR LAAAASSAASPATSDSMAAAPFHQAAAGGGDEAWRDAELQEVVSALVAATTTTATAQAPATAMVDADLSALDAFEFDPGFTIDITSFFA ZmWRKY-IIIA22 - GGCPCPCLAPSTPENR VSPASSNSNHAASYLPFEDA - EWRGHAGLOEVVSALVAATTTTATAQAPATAMVDADLSALDAFEFDPGFTIDITSFFA SDWRKY-IIIA22 - STORTAGENWG VSPATSDSNHVAASYLPFEDA - EWRGHAGLOEVVSALVAATTTTATAQAPATAMVDADLSALDAFEFDPGLLLVDIDDIASFFA SDWRKY-IIIA22 - STORTAGENWG VSPATSDSNHVAASYLPFEDA - EWRGHAGLOEVVSALVAASAPPPP - AVDSLDDLLFDIDIASFFA SDWRKY-IIIA22 - STORTAGENWG VSPATSDSNHVASSFRPFEAA - EWRGQAENGVVSALVAASAPPPPP - AVDSLDDLLFDIDIASFFA SWRKY-IIIA22 - STORTAGENWG VSPATSDSNHVSFPFFEVAG - DDVQFGRFEEVMS	IIIA2-
OsWRKY-IIIA3a OYFSFISFANSDCQFSS - DFSAGSVQVDMDHEARFED - LFSSTLEFFQSEIQNL SWRKY-IIIA3a - ASAGADFAGWC - FLLSP OsWRKY-IIIA3b - TSSFFVSFAMSDCQIS - YELGOGSMAGVRNVP OsWRKY-IIIA3b - TSSFFVSFATESCRIVS - GYALG SWRKY-IIIA3b - TASFFLSFATESCRIVS - GYALG SWRKY-IIIA3b - TASFFVSFATESCRIVS - GYALG SWRKY-IIIA3b - TASFFVSFATSCRIVS - GYALG SWRKY-IIIA3b - TASFFVSFATSCRIVS - GYALG SWRKY-IIIA3b - TASFFVSFATSCRIVS - GYALG ZmWRKY-IIIA3a - TASFFVSFATSCRIVS - GYALG ZmWRKY-IIIA3b - TASFFVY ATPOCQLVSGGGGGYANGGGGGV - GYNNV POVELASTTNSPAMGGCMDFMFPLDAASFLELESPASYC ZmWRKY-IIIA3a - TASFFVY ATPOCQLVSGGGGGYANGGGGGV - GYNNV POVELASTTNSPAMGCMDFMFPLDAASFLELESPASYC ZmWRKY-IIIA3a - FFSNKPEFAVVAHAASN - OFFEWTECHVSGAKNVPDVELASTTNSPAMGCMDFMFPLDAASFLELESPASYC SWRRY-IIIA3a - FFSNKPEFAVVAHAASN - OFFEWTECHVSGAKNVPDVELASTTNSPAMGCMDFMFPLDAASFLELESPASYC SWRRY-IIIA3a - FFSNKPEFAVVAHAASN - OFFEWTECHVSGAKNVPDVELASTNSPAMGCMDFMFPLDAASFLELSPASYC SWRRY-IIIA3a - FFSNKPEFAVVAHAASN - OFFEWTECHVSGAKNVPDVELASTNS SWRRY-IIIIA3a	IIIA3-
AtWRKY-JIB12 VWRKY-JIB12 AtWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS FAESMIENSLMGGFS FAESMIENSLMGGFS FAESMIENSLMGGFS FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS GUWRKY-JIB12 FAESMIENSLMGGFS FAESMIENSLMGGS FAESMIENSLMGGFS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLMGGS FAESMIENSLM	IIIB1-

1	430	440	450	460 47	480	490	500	510	520	530	
AtWRKY-IIIC1a EDI		IKEQAIDQY	REITSNDQDCQ	D V I E E Y L S S <mark>P</mark> S <mark>G</mark> S	5 Y <mark>P P</mark> S S S S <mark>G</mark> S E S A D F	N	• • • • • • • • s [ILLFDN <mark>P</mark> DSWD	RYDQFYF		
AtWRKY-IIIC1b · · ·		· · · LRQHQN	HQDRIKDEYMK	PVIAEDWS-PSQV	WM SSEVALAVEAFEF	•N•••••	<mark>P</mark> F	WTSHDLSS			
AtWRKY-IIIC1c			AVLME - DEAN -	HIMNQEYDINDYL	. V D D E V F W <mark>G</mark> N E F <mark>P</mark> L F	s	· · · · · · SE	DLML - F			
AtWRKY-IIIC1d		TIDEQ	AITME - DEAID	HIMNQECDINDFS	5 V D D D <mark>P</mark> F WA S Q F <mark>P P</mark> F	s	· · · · · · · · SE	DIMF - FDNIA	NLD		
AtWRKY-IIIC1e · · ·		T I DHQ	VITVE - ENSAE	HIMNQECDINDYL	. V D D D <mark>P</mark> F WA S Q F <mark>P P</mark> F	• • • • • • • • • • • • • • •		DTMF - LENIS	AFD		
AtWRKY-IIIC11 · · ·		S I D S Q	EITMEDKDTDD	HILNYINEHL	MEDEAY DVF <mark>P</mark> DV	/L	· · · · · · · · · · · · · · · · · · ·	RCCF <mark>G</mark> LE <mark>P</mark> F <mark>P</mark>	GLNINKS		
a-110101 1100 -								wa w			
GWWRKT-IIIC23 ····		Feee	UEVLIINNULS LLASTKUE <mark>D</mark> OE	VIUNEURAANKIR	TEENLIEVOVDI <mark>D</mark> E	M	· · · · · · · · · · · · · · · · · · ·	MOKSISV	OF DE VIER FER	vor	
GmWRKT-IIIC24 · · ·			ICASIKHERQE	VIODDULVONOLS	eenvillenvalne	M	· · · · · · · · · · · · · · · · · · ·	KNAILLJJIE	AVGFENVIEG	1 <mark>0</mark> 7	
GmWRKY-INC20			K DECKEEV DD S	ISSNDVI ISSDIT	EDSSBDUUVTI	S	ет	IDSEVKSVD.	ISDVI VDSAOI		
GmWRKY-IIIC2d			KRECKEEVPPS	TSENUVIISEDIT	FDSS PPUUVTI	s		IDSEVKSVD.	ISDVLYDSAG		
GmWRKY.IIIC2a			KRGDTEKIIIS	MICDIVII						EDDATE <mark>T</mark> TEETK	
VvWRKY-IIIC2a ····		VT YN		SSSSDYLLSLELT	T FESNL <mark>G</mark> SDH <mark>G</mark> DVL	s		NSSCTOSTHS		VLIGFEC······	
AtWRKY-IIIC2a		MVKEE	ONNNGDOSKDY	YEGSSTGEDLSLV	WOETMMFDD HON	HYYC <mark>g</mark>	ET	STTSHOFG	F DNDDQ	FSSFFDSYCADYERT	IIIC2-
AtWRKY-IIIC2b · · ·		MKEE	DNPHR	HHGSSTENDLSLV	WPEMVFEEDYHHQA	SYVNG	KT	STSIDVLGSQ	DLMVFGGGGD	FEFSENEHFSIFSSC	
GmWRKY-IIIC2f		IVKQE	Y PNDET DP SDL	TDANFWSDFKDFE	LSNDKPAGLKIASE	N	ADT	VYSCTGSRSL	DMDFGIFSSH	FCSTEDFHFDESQLL	
GmWRKY-IIIC2g		IVIQE	Y <mark>P</mark> NDETD <mark>P</mark> SDL	TDANLWSDLKDFE	L SNDK <mark>P</mark> AGLK I AP	N	ADS	VYSCT <mark>g</mark> srsl	DMDF <mark>G</mark> IFS <mark>P</mark> HI	FCSTEDFHFDESQLL	
GmWRKY-IIIC2h		TIKQE	F <mark>P</mark> KEDTH	S D V A D H K	(YSD <mark>P</mark> NL			.	· · · · · · · · · · · · · · · · · · ·		
CPWRKY-IIIC2b · · ·		F S L S G	LRQENKASDEQ	V IWKDLM <mark>P</mark> LN	ISAE <mark>P</mark> AVYS		- c s	EATSTSTHCL	<mark>P</mark> MDFVVD	F DTAFCFDHS <mark>G</mark> - D	
VvWRKY-IIIC2b*		TVVSD	L T S D N V S S M D S	IN LWSGLEALD) S F M <mark>P</mark> A M V T Q R <mark>G G</mark> S D	DH <mark>G</mark> NQDQNVDS1	MYSRNATT	TTATTASHNT	DMD F V L <mark>G</mark> C L D I	F E <mark>G</mark> D D Q F Q F D E S <mark>P</mark> F Q	
GmWRKY-IIIC2i		F S S S	L Y S S MKQ E <mark>P</mark> ME	VIHEDHIVHNQL	SSDYHMLCDYDLDF	N	YS	RY <mark>G</mark> TMLSSTE	SVQFDEVCRSI	DQEFD <mark>G</mark> ·····	

	430	440	450	460 470	480	490	500	510	520	530	
BdWRKY-IIID1a MT	тү	QAQEFD	E <mark>G</mark> D L					DV <mark>P</mark> AL	LEVLDN <mark>P</mark> LLI	WDMW	
BdWRKY-IIID1b SA	τ	· · · EVQEFD	E <mark>G</mark> E • • • • • L					NMSAL	LEVLDD <mark>P</mark> LLI	1WEIIC	
ZmWRKY-IIID1a TA	TR	- AGGVQESD	EADH L			• • • • • • • • • • • • • • • •		DL <mark>P</mark> AL	LEVFDSSLVI	OWEALYHS <mark>P</mark> NVATHS-	
SEWRKY-IIID1a SS		DVQDP -	E A D L					DL <mark>P</mark> AL		DWEDIWKI	
OSWRKY-IIID12 DC	N	···· KLL	DMMPAADELT					ADVLL	F DMI AYAPLI		
SHWRKY.IIID16 GN	KI	LINKSADLL	TRN					SMYE		D. SWALDAFLER	
ZmWRKY-IIID1b QA	FL	- PHDSKPIV	DDDESAEEEL					DMREP	FDVAAFAPMI	DFD - SWELDALLRFGA	IIID1-
BdWRKY-IIID1c HH	IRQ	SSLESSLL	DMAEEDLATE					EYDQL	F D V A A Y E <mark>P</mark> L I	SDPAWEMDDAHGHGH	
OsWRKY-IIID1¢ AE	н	SAL	G L E D E					HMHKL	LDTFAC <mark>G</mark> ALI	DLD-SWEIDAIVRSGF	
SbWRKY-IIID1c [*] S <mark>G</mark>	DD	ITCWSSTS	<mark>gg</mark> asssdyny	• • • • • • • • • • • • • • • •				ADDDY	YYDC <mark>G</mark> - <mark>G</mark> L F <mark>(</mark>	AAVHG GGWAT GP	
OsWRKY-IIID1d [®] T <mark>G</mark>	N	••••ws <mark>g</mark> qн	GAYACR-QMI					AADEE	Y C C WD T <mark>P</mark> A T 1	TTTS <mark>G</mark> SN <mark>GG</mark> NS-TCA	
ZmWRKY-IIID1c ^{**} S <mark>G</mark>	DD	 VACCWSSS 	<mark>G</mark> - SSSS <mark>G</mark> YSY					ADDSC	CCCCCD <mark>G</mark> LL <i>I</i>	AVVH <mark>GG</mark> C <mark>GG</mark> SWA <mark>PGP</mark>	
0sWRKY-IIID1e [®] T <mark>G</mark>	N	WS <mark>G</mark> QH	GAYACRRQMI					AADEE'	Y C C WD T P A T 1	TTS- <mark>G</mark> SD <mark>GG</mark> NSSTCA	
OSWRKY-IIID2a HE	AH	- EVQDERFE	N L CMVQNM <mark>P</mark> E					· YL I DF	· ELERAFEF	VN S <mark>P</mark> L <mark>G</mark> S E HWT F	
ZmWRKY-IIID2a ND	νγ	GIQDESFE	NLFMVPSIPE					YLTDF1	TDIEMA <mark>G</mark> ALI	VT SEMISENIWA	
SBWRKY-IIID22 ND	WL	GIHDESFE	NFFMVPGMPE					YLTDF1	TDFETAEALI	VTSMIISEDIWA	IIID2-
SBWRKY-IIID28 TD	VH	GVQDETFE	SLFMVPDMPE					YLTEF	VDVEMARAFI		
USWRKY-IIID20 DD	но	- PKNDMKPE	NLFAV <mark>P</mark> DMSL		- -			FSENM	MUTTFEDVIN	INSTISLEGEARDSWI	
							-				
SAWRKY-IIID32 KL	. K.K. • • • • • • • • • • • • • • • • •	· EEDVVA <mark>G</mark> ·			- MSVII <mark>P</mark> SVI	YDELSSSSL <mark>P</mark> LMS	*	MQ WE I	MEMEMASLFI	(RQDASS <mark>G</mark> S	
ZmWRKV-III032											IIID5-
					11						
OsWRKY-IIID4a 0	R	· · · OCYDD ·				AVHARA <mark>P</mark>		AAAA <mark>P</mark> VA	ASSSSS	DAAGEELDVMDYDMTD	
ZmWRKY-IIID4a GG	Q	QQQSP -	AGSGTT		TAVDPAVN	VGQEQEPPPPRAR	r	- PAAVD A	AGEMPR-II	VDVAGLDVMDYDVTD	IIID4-
SbWRKY-IIID4a <mark>G</mark> A	Q	QQQS <mark>P</mark> -	A <mark>G</mark> S <mark>G</mark> TT		AAVN VG	HEQDQE <mark>PPPP</mark> RAR	FRDDA <mark>P</mark> A <mark>P</mark> AP	A <mark>P</mark> AGVDDDA	P G E M P R S I I I	VDVA <mark>G</mark> LDVMDYYVTD	m
BdWRKY-IIID5a I G		- RAAA <mark>P</mark> S	SWS S P A S -		- SML P	PPPVEENM <mark>G</mark>		EMMEIYSYF	• • • • • • • • • • •	L C <mark>G</mark>	
ZmWRKY-IIID52 R <mark>G</mark>		•••• • • • • • • • • • • • • • • • • •	- WS <mark>P</mark> VAADAS	DA <mark>G</mark> AASY <mark>GP</mark> A <mark>P</mark> MQM	IQMMQAS <mark>G</mark> SGHS	ASVAQD <mark>GG</mark> ATMMT	r	MIGTDDTDF	C C WD <mark>P</mark> S I	. V <mark>g</mark> eadhqmmdedhrd	
								_	_	_	
OsWRKY-IIID6a PP	PP	- PADLATIS	NN F D E T <mark>P G</mark> L H		VCQEV <mark>PP</mark> SS	SNSSVISHYAD		E F DI	HHQM LE1	TVMEEAL	
OsWRKY-IIID6b PS	PP	- P - DLTTIS	NNFDET PGLH		VSQEV <mark>PP</mark> CS	SNSSAISHYAD		EFDI	HHQMGQQLET	TVMEEAL	
SHWRKY-IIIDGA PA		- DLQQPSAC	FPNEQIP-		- QCQKE <mark>P</mark> LFP	- TSMEQQFVCGAL	R	DHDSPVD	G - DIPSATG	SCN SGET SWWDGY SGD	IIID6-
ZmWRKY-IIID6a PP	PP	- LAEPPSAC	FPYDQMPQ		- QCHQ - LLFP	- VSMEQQFPG		N	GEEIPSAAG	CISCETS-WDGYY	
ZMWRKT-MDGD <mark>P</mark> L	FF	· L VE <mark>PPP</mark> AS			- - QCQQE <mark>P</mark> LF <mark>P</mark>	TI SMEQQLVC <mark>g</mark> al	.	0H <mark>0</mark> 5 m AD	OUCIPSATG	SCISCIEL CWWDCYSCD	
OCIMBICY IIID 7-		PEABAGEE		VEDEDIMENTODAD				CMDELVYEU			
RdWRKY.U072		SOTSWROW		VELG.	ASPSPVMEFIE	SEDWES		NS. IGEGOL	PIVAMIO	A D D D A F V	
OcWRKY-IIID78		- SGI SWROW	13333333333 12. <mark>D</mark> efeeive	I DVVGEDI AGADSS	AS PWWEFIN	G S F DWE F	VI	NSI	C C C AMLQ		
ShWRKY.IIID7a		ASMSOGGW	S. PSASSEVE	EDE FALHEWHDT		GCFGWES		ODSSDE			
ZmWRKY-IIID7a		- SSPSOGGW	SSPSASSEAG	VGFGTOLAHEWHDT	AAPDPVTEFLD		VL	QDRFDFGG -	- L S D V A T F Q		
ZmWRKY-IIID7b		- DLCSSPPE	EELRAGACDV	AEFVAEQSTTVTAE		GCLDWEL		GDSSLDVDG	FDRYYFDYS	- DLL	
OsWRKY-IIID7c		- ESGVKMSK	EEP-VDSCPG	P SAVSSPADVVSCS	SP AMEPDLL	GCLNWDD	D F	GDSSF - VDA	DEFMNFDEI	DLFQIYS	
								_			
RdWRKY-UID2a				FLAAVVVASAGPTS		GGTAAG			PSSI GERVA	SI <mark>6</mark> T <mark>6</mark>	
SEWRKY-IIID8a		DDVF355	AGS FMOVD - 0	LIGAVVGGSAGVVT	SAAAPDRGVVI	GGVASGG		GTASEPTS	PNSLGFVVG		
BdWRKY-IIID8b		DQEEVRSCL	TPGSSAVHS -	TAAAAGAG	PDQGDVS	SAPQFYE		GAADMGEFF	GLEDIFDLD	DH	
BdWRKY-IIID8c		KVEKAMVS I	WSFGAVLSS-	ST KNSVDNE	DVEPIEEQELH	DHLILTD	EKISS	FPPMVKVLE	YPLKNYELE:	SCVHD <mark>P</mark> FTIYDLDCLY	IIID8-
ShWRKY-IIIDBb		EQEEVLSSL	T PAGSSAAAE	AMRNGNAAAAAATT	TGPEPDQGDVT	s <mark>g</mark> lqlqq	F Y <mark>G</mark> AGD	DLAYMARFS	- YDDT FDLEI	DIVVF <mark>G</mark> A <mark>P</mark> DSIT-DIY	
ZmWRKY-IIID8a EQ	EL	ELEEVLSSL	T PAGSSAAT -	AEAVRT	AT <mark>PGP</mark> DQ <mark>G</mark> DVT	S <mark>G</mark> LQLQQQQQ	HWF <mark>GGG</mark>	G L S <mark>G</mark> V A H L G	- YDDTFDLEI)	·
OsWRKY-IIID8a -		DQEEVLSSL	T <mark>PG</mark> SSAAR <mark>G</mark> -	<mark>GGGGGG</mark> VA	GP F GP D Q G D V T	SSLHWSY	0	AVAGME F F KI	NDEVVFDLD	0 M <mark>G</mark> L S F	