# Promoter analysis of members of a plant defense-related LRR-RLK gene cluster in *Arabidopsis thaliana11*

by

Anina Mumm

(200602463)



submitted in fulfilment of the requirements for the degree

***Magister Scientiae***

in

**Biochemistry**

in the

Faculty of Science

at the

University of Johannesburg

Supervised by Prof. I.A. Dubery & Dr. B. Kemmerling

October 2013

# Summary

A 14-member, closely-spaced cluster of genes coding for leucine-rich repeat receptor-like kinases (LRR-RLKs) is located on chromosome 1 of *Arabidopsis thaliana.* Following on from previous microarray studies that found some of the members of this cluster to be upregulated in response to biotic stressors, including the bacterial elicitor flg22, the present study sought to confirm, using a luciferase-based protoplast assay, that flg22 does in fact induce the expression of the genes, and then to investigate the promoters of the genes. The promoters of At1g51790, At1g51850 and At1g51890 responded positively in this particular assay, and bioinformatic analyses determined that W-boxes are over-represented in the cloned regions. Mutational inactivation of individual W-boxes in the promoter of At1g51790 drastically reduced the flg22 response, except for the W-box closest to the start site, which seemed to increase both basal and flg22-inducible expression. In the promoter of At1g51850, mutational inactivation of either or both of its W-box dyads resulted in virtually no flg22 inducibility. The deletion of 6 W-boxes in the promoter of At1g51890, done via truncation, drastically reduced both its basal expression and its inducible response to flg22. These results provide evidence that W-box cis-elements are responsible for the upregulation of these LRR-RLKs in response to flg22. WRKYs -7, -11, -22,and -26 were found bioinformatically to have similar expression patterns to some of the genes in the cluster, and are thus good candidates to investigate as transcriptional regulators of the cluster in future studies.

# Acknowledgements

This study and my stay in Germany was made possible through funding from the National Research Foundation, the University of Johannesburg, the Bundesministerium für Bildung und Forschung (BMBF) and the German government.

Without the assistance and patience of Prof. Ian Dubery, I would never have had the opportunity to study in Germany, and I likely would not have completed this dissertation. His support and guidance through the last 5 years, since my Honours, has been invaluable.

Thanks also to Dr. Birgit Kemmerling, who welcomed me into her lab, taught me independence, and guided the progression of my research and dissertation.

Everyone I met at the Zentrum für Molekularbiologie der Pflanzen (ZMBP), University of Tübingen, were welcoming, friendly, supportive and always willing to help. I don't think I could have made it through the tough times in the lab and in a foreign country without Christine, Steve, Thierry, Sara, Luise, Dierk, Anna, Christina, Markus, Eva, Isa, Marko, Inga and many others. Thanks also to those in the UJ Biochemistry lab, especially Sherrie-Anne, Carol and Edwin.

Thanks to Prof. Dubery, Dr. Cronje, Dr. Störk and the Faculty of Science at the University of Johannesburg, my husband Stephan Mumm was able to go with me to Germany, to complete his MSc there as well. We've won together.

Last but not least, my parents, Izak and Ina Minnaar, my mother-in-law, Angela Mumm, and my dear friends, Colette and Tristan Mathiesen, kept me going and reminded me that South Africa is home.

There are three other lessons I have carried with me since our time in Germany. Firstly, always ask because the worst that can happen is that someone says no, but most of time people will surprise you and say yes. Secondly, job satisfaction to me means a challenge without being frustrating, making a tangible difference in the world and feeling constantly rewarded. I am now lucky enough to have all of that, so I am grateful for every experience that has led me here. Thirdly, life is but a game.

With this I hang up my purple lab coat, but I still have my pen, notebook and curiosity.

# Table of Contents

# Figures

# Tables

# Abbreviations

| | |
|---|---|
| ABA | Abscisic acid |
| APS | Ammonium persulfate |
| Avr | Avirulence |
| Ax21 | X*anthomonas* M/PAMP |
| BAK1 | BRI1-asscoiated kinase 1 |
| BIK1 | Botrytis-induced kinase 1 |
| BP reaction | The first reaction of Gateway cloning |
| bp | Base pair |
| BRI1 | Brassinosteroid-insensitive 1 |
| BSA | Bovine serum albumin |
| bZIP | basic-domain leucine-zipper |
| $Ca^{2+}$ | Calcium ion |
| $CaCl_2$ | Calcium chloride |
| cDNA | complementary DNA |
| CDPK | Calcium-dependent protein kinase |
| CK | Cytokinins |
| Col-0 | Columbia ecotype of *A. thaliana* |
| DELm | Deletion mutant |
| $dH_20$ | Distilled water |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide |
| DPI-ELISA | DNA-protein interaction enzyme-linked immunosorbent assay |
| DS | Double stranded |
| EDTA | Ethylenediaminetetraacetic acid |
| eFP | Electronic Fluorescent Pictograph |

| | |
|---|---|
| ELISA | Enzyme-linked immunosorbent assay |
| ER | Endoplasmic reticulum |
| ERF | Ethylene-responsive-element-binding factors |
| ES | Enzyme solution |
| ET | Ethylene |
| ETI | Effector-triggered Immunity |
| ETS | Effector-triggered susceptibility |
| flg22 | 22 amino acid peptide of bacterial flagellin |
| FLS2 | Flagellin-sensitive 2 |
| g | Gravitational acceleration |
| GA | Gibberellins |
| GCOS | Gene Chip Operating System |
| gDNA | Genomic DNA |
| GFP | Green fluorescent protein |
| GW | Gateway |
| $H_2O_2$ | Hydrogen peroxide |
| HCl | Hydrochloric acid |
| His | Histidine |
| HR | Gypersensitive response |
| HrpZ | Bacterial M/PAMP |
| IPTG | Isopropyl-β-D-thio-galactoside |
| JA | Jasmonic acid |
| KCl | Potassium chloride |
| kDa | Kilo Dalton |
| LPS | Lipopolysaccharide |
| LR reaction | The second reaction of Gateway cloning |
| LRR | Leucine-rich repeat |
| LRR-RLK | Leucine-rich receptor-like kinase |
| Luc | Luciferase |

| | |
|---|---|
| M | Molar/moles per litre |
| MAMP | Microbe-associated molecular patterns |
| MAPK | Mitogen-activated protein kinase |
| $MgCl_2$ | Magnesium chloride |
| MIMP | Microbe induced molecular patterns |
| Min | Minute |
| miRNA | Micro RNA |
| mRNA | Messenger RNA |
| MW | Molecular weight |
| NaCl | Sodium chloride |
| NB-LRR | Nucleotide-binding LRR |
| NCBI | National Center for Biotechnology Information |
| NO | Nitrogen oxide |
| OD | Optical density |
| OPD | o-Phenylenediamine dihydrochloride |
| PAGE | Polyacrylamide gel electrophoresis |
| PAMP | Pathogen-associated molecular pattern |
| PBS-T | Phospate-buffered saline with Tween |
| PCR | Polymerase chain reaction |
| PGN | Peptidoglycan |
| prom | Promoter |
| PRR | Pattern recognition receptors |
| PTI | PAMP-triggered immunity |
| *R*-gene | Resistance gene |
| RLK | Receptor-like kinase |
| RLU | Relative light units |
| RNA | Ribonucleic acid |
| RNAi | RNA interference |
| ROS | Reactive oxygen species |
| Rpm | Revolutions per minute |

| | |
|---|---|
| R-proteins | Resistance proteins |
| RSAT | Regulatory Sequence Analysis Tools |
| SA | Salicylic acid |
| SAR | Systemic acquired resistance |
| SDS | Sodium dodecyl sulfate |
| siRNA | Small interfering RNA |
| T3SS | Type III secretion systems |
| TAE | Tris-acetate-EDTA |
| TAIR | The Arabidopsis Information Resource |
| TEMED | Tetramethylethylenediamine |
| TF | Transcription factor |
| UV | Ultra-violet |
| W-box | *Cis*-element recognised by WRKYs |
| WBX | W-box |
| WRKY | Transcription factor containing WRKYGQK consensus sequence |
| Xa21 | Rice receptor of Ax21 |
| ZMBP | Zentrum für Molekularbiologie der Pflanzen |

# Chapter 1: Scope and Aims

## 1.1. Background and significance

Identifying the mechanisms used by the plant to defend itself against pathogens is crucial to developing new bio-technologies to curb plant disease and therefore increase crop yields. While it is well-known that leucine-rich receptor-like kinases (LRR-RLKs) are major components of plant immunity in terms of perceiving pathogenic threats at the cell membrane, there are still many questions surrounding their exact functions, regulatory mechanisms and signal transduction pathways (Afzal *et al*., 2008; Muthamilarasan & Prasad, 2013; Pel & Pieterse 2013). The present study will make a contribution here in that it examines an LRR-RLK-coding gene cluster on *A. thaliana* chromosome number 1. Some of the genes in this cluster, namely *At1g51790, At1g51800, At1g51820, At1g51850, At1g51860* and *At1g51890*, were found to be up-regulated upon exposure to pathogen attack, and peptidoglycan and flg22 elicitors specifically, and were suppressed after effector delivery from virulent bacterial infection (Chae *et al*., 2009; Gust *et al*., 2007; Kemmerling *et al*., 2007; Kemmerling *et al*., 2011; Postel *et al*., 2010). Furthermore, preliminary bioinformatic promoter analyses (BFF Germany/NRF South Africa Progress Report 2009) suggested potential binding sites for defense-related TFs.



These findings merit an investigation into whether this gene cluster could play an important part in plant disease resistance, and more importantly, how the cluster and/or individual genes are regulated. If scientists could artificially mimic this type of regulation by increasing the transcription of plant disease resistance genes, plants could be engineered to be resistant to infection even before a pathogenic threat is present. This project therefore aimed to define the promoter regions necessary for PAMP inducibility, the transcription factor binding sites within the promoter regions of these genes, as well as the transcription factors themselves.

## 1.2. Aims

The broad aims of this study is to investigate the *cis*-elements in the promoters of an LRR-RLK-coding gene cluster on A. thaliana chromosome number 1, and to determine whether known defense-related

transcription factors (WRKYs) may interact with these *cis*-elements to regulate the expression of the genes upon pathogenic attack.

## 1.3. Specific questions

The present study addresses the following specific questions, which represent a timeline of trial-and-error, each building on the answer to the previous, using available bioinformatic tools and databases, as well as molecular biological and biochemical experimentation:

1. Are there any structural and functional similarities between members of the LRR-RLK gene cluster of interest on chromosome 1 of *A. thaliana*?
2. Does biotic stress induce expression of the LRR-RLK gene cluster?
3. Can flg22-induced expression of the genes be assayed *in vitro* using *A. thaliana* protoplasts and a luciferase reporter system?
4. Are known *cis*-element sequences over-represented in the promoter regions of the LRR-RLK genes?
5. Do promoter truncations and/or deletions targeting these *cis*-element sequences alter the flg22-inducibility of the LRR-RLK gene cluster promoters *in vitro*?
6. Are there any WRKY transcription factors that exhibit expression patterns similar to the LRR-RLKs of interest?
7. Does over-expression of these WRKY transcription factors affect the expression kinetics of the LRR-RLK gene cluster *in vitro?*
8. Can direct protein-DNA interaction be shown between these WRKY transcription factors and the promoter regions of the LRR-RLK gene cluster *in vitro*?

# Chapter 2: Literature Review

Since the present study focuses on LRR-RLKs, major components of the signal transduction pathways involved in plant immunity, here follows a brief summary of what is known about plant defense responses and the role of LRR-RLKs. Because flagellin was used in this study, the literature review will address what is known about flagellin perception and responses as well. In addition, since the present study specifically zooms in on the promoters of the LRR-RLK gene cluster of interest, as well as possible transcription factors involved in their regulation, defense-related transcription factors (WRKYs specifically), are also covered in this literature review.

## 2.1. Plant Immunity

In addition to the cell wall, wax layers, cuticular lipids, antimicrobial enzymes, secondary metabolites and other preformed defenses, plants have evolved two layers of inducible innate immunity to defend against pathogen attack (Muthamilarasan & Prasad, 2013). The first layer is known as PAMP-triggered immunity (PTI), resulting from the interaction between the pattern recognition receptors (PRRs) of the plant and the pathogen-associated molecular patterns (PAMPs) of the pathogen (also called microbe-associated molecular patterns, or MAMPs). Pathogens may produce effector molecules to overcome this first layer of immunity, thereby making the plant susceptible to disease. This is known as effector-triggered susceptibility (ETS). In an evolutionary response to ETS, a second layer of plant defense, called effector-triggered immunity (ETI), may result from the interaction between the *R* (resistance)-gene products of the plant and the effectors of the pathogen. These pathogen effectors are thus at first virulent factors but become avirulence factors (Avr) through evolution, upon direct/indirect interaction with the plant's R-proteins. This two-layer model of plant immunity has been reviewed extensively (Consonni *et al*., 2009; Cui *et al*., 2009; Cunha *et al*., 2006; Dangl & Jones, 2001; He *et al*., 2007; Jones & Dangl, 2006; Knogge *et al*., 2009; Mackey & McFall, 2006; McDowell & Simon, 2008; Postel & Kemmerling, 2009) and its basic elements are summarised below and in Figure 2-2.

## 2.1.1. Pathogen-triggered Immunity (PTI)

### *2.1.1.1. PAMPs/MAMPs*

PAMPs/MAMPs are highly conserved, widespread and abundant molecular motifs, like sugars, peptides and lipooligosacharides, which are crucial to pathogen survival and are recognised as non-self by the plant host. In fact, all viruses, bacteria, fungi, oomycetes and insects are perceived by plants as a result of PAMPs/MAMPs (Pel & Pieterse, 2013; Sanabria *et al*., 2010). Examples of pathogen-associated molecular patterns include, but are not limited to, bacterial flagellin (or the synthetic 22-amino acid peptide, flg22, which represents a conserved domain of flagellin), bacterial peptidoglycan (PGN) and bacterial lipopolysaccharide (LPS) (Pieterse *et al*., 2012; Postel & Kemmerling, 2009). M/PAMPs of fungal/oomycete origin include ergosterol, β-glucan and chitin (Knogge *et al*., 2009) but these are not as well studied as bacterial M/PAMPs.

### *2.1.1.2. PRRs*

Membrane-bound PRRs mediate inducible defense responses by discerning M/PAMPs based on their characteristic molecular pattern (Consonni *et al*., 2009; Postel & Kemmerling, 2009). The best studied PRRs are LRR-RLKs – see 2.2 (Goff & Ramonell, 2007), like FLS2, the flagellin receptor. Nucleotide-binding LRRs (NB-LRRs), represent another type of PRR in that they recognise pathogen encoded virulence factors. And Xa21 is yet another type, recognising a specific conserved molecule from the *Xanthomonas* species known as Ax21 (Ausubel, 2005; Roux & Zipfel, 2012; Tena *et al*., 2011).

## 2.1.2. Pathogen suppression of PTI: Effector-triggered susceptibility (ETS)

### *2.1.2.1. Effectors*

Pathogen effector proteins enhance the virulence of pathogens. They have evolved to suppress the M/PAMP-induced defense response in plant hosts by modification of the host's proteins. As such, they can target PTI at any level, from the PRRs and mitogen-activated protein kinase (MAPK) cascade to transcription factors. (Consonni *et al*., 2009; He *et al*., 2007; Rasmussen *et al*., 2012).

Effectors are secreted into plant cells by bacterial type III secretion systems (T3SS). The bacterial Hrp pilus represents such a system: the HrpZ component of the pilus, which is itself secreted by the pilus, secretes effectors into the host cell. HrpZ works by associating with the lipid membrane and forming

ion pores (Chinchilla *et al*., 2007; Guttman *et al*., 2006; Lee *et al*., 2001). The effectors secreted by the T3SS suppress PTI in successful pathogens and thus render the plant susceptible to disease, hence the term 'effector-triggered susceptibility' (ETS).

These effectors may act as transcription factors (TFs) themselves or they may stimulate host TFs to activate nutrient production that favours the survival of the pathogen. They may also affect histone packaging and chromatin configuration. Examples include AvrPto, AvrPtoB, HopF2 and HopAI1 from *Pseudomonas syringae*, all of which target FLS2 signalling pathways to supress flg22-induced responses (Abramovitch *et al*., 2006; Cui *et al*., 2009; Feng & Zhou, 2012; Guttman *et al*., 2006; Muthamilarasan & Prasad, 2013).

### 2.1.3. Effector-triggered Immunity (ETI)

In resistant plants, some effectors trigger a secondary immune response through direct/indirect interaction with host *R*-gene products. This is known as effector-triggered immunity (ETI). It is difficult to define bacterial molecules like HrpZ as either a M/PAMP or an effector because, although secreted by T3SS and followed by ETI, it is also a highly conserved external bacterial ligand that is recognised by plant membrane receptors (Guttman *et al*., 2006; He *et al*., 2007; Knogge *et al*., 2009).



#### *2.1.3.1. R-proteins*

Intracellular resistance proteins (R-proteins) mediate defense responses by interacting with pathogen effectors. As with M/PAMPs versus effectors, it is sometimes difficult to distinguish M/PAMP-receptors from R-proteins, which led Mackey and McFall (2006) to suggest that pathogen effectors should rather be referred to as microbe (or pathogen) induced molecular patterns (MIMPS), and that both MAMP-receptors and MIMP-receptors should be classified as R-proteins.

The direct *R-Avr* model and the indirect "guard hypothesis" have been proposed as strategies of ETI (Abramovitch *et al*., 2006; Consonni *et al*., 2009; Dangl & Jones, 2001; Monaghan *et al*., 2009). The *R-Avr* model entails direct gene-for-gene interaction between the pathogen effector and the plant R-protein, whereas the "guard hypothesis" proposes that the association of the R-protein with the effector target allows the R-protein to indirectly trigger defense responses when it senses the binding of an effector to the target. See Figure 2-1 below.

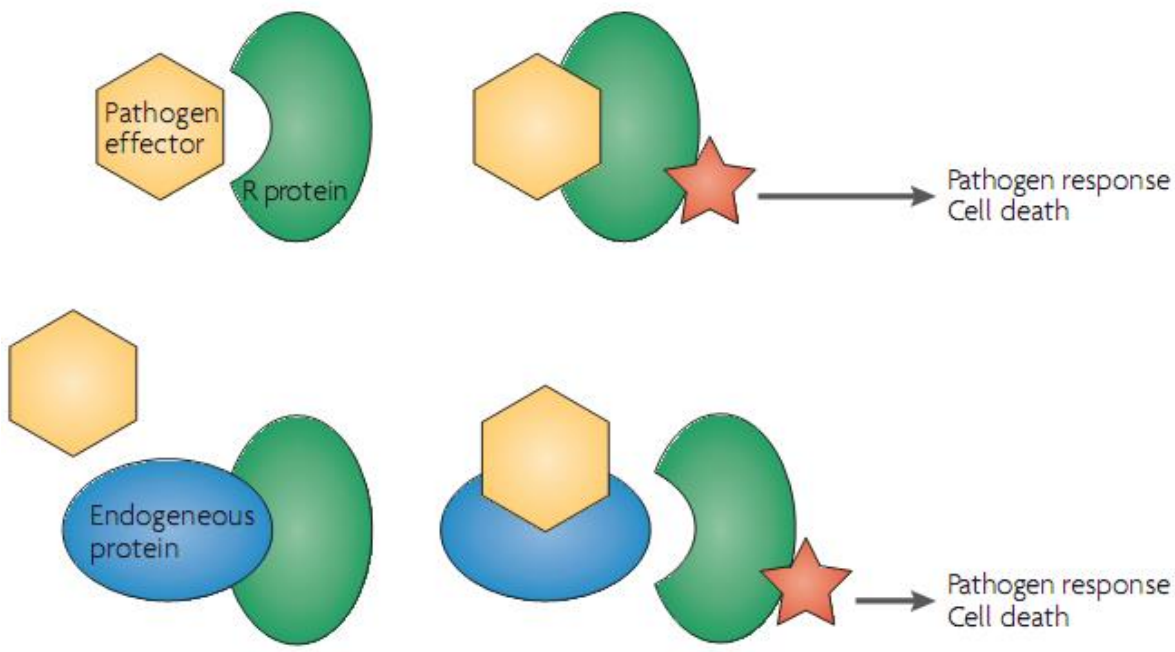



**Figure 2-1: Hypothesised strategies of ETI**

**Above:** The *R-Avr* model: direct gene-for-gene interaction between the pathogen effector and the plant R-protein. **Below**: The 'guard hypothesis': association of the R-protein with the effector target allows the R-protein to indirectly trigger defense responses when it senses the binding of an effector to the target (Bomblies & Weigel, 2007)**.**



Recent studies suggest that R-proteins may actually monitor changes in self-molecules caused by non-self-molecules (Sanabria *et al*., 2010), rather than detecting those non-self-molecules directly, which could explain how a small number of R-proteins are able to respond to a much larger number of pathogen effectors. The most abundant R-proteins are nucleotide-binding leucine-rich repeat proteins (Abramovitch *et al*., 2006; Consonni *et al*., 2009; Dangl & Jones, 2001; Knogge *et al*., 2009; Spoel & Dong, 2012).

Mutual selective pressure drives pathogens and plants to constantly evolve new effectors and new R-proteins to overcome plant defenses and pathogen attack, respectively. This has been dubbed the 'arms race' between phytopathogens and their hosts (Cui *et al*., 2009; Monaghan *et al*., 2009).

### 2.1.4. ETI- and PTI-associated defense responses

There are characteristic cellular events associated with inducible defense responses: the structural modifications, signal transduction pathways and chemical changes that lead to the hypersensitive response (HR) are local defense responses deployed at the site of infection. Furthermore, the plant is

capable of priming itself against future pathogen attack both locally and systemically (Chinchilla *et al*., 2007; Consonni *et al*., 2009; Cui *et al*., 2009; Knogge *et al*., 2009; Mackey & McFall, 2006).

#### *2.1.4.1. The hypersensitive response*

The HR is the hallmark of a resistant plant. Localised cell death occurs in the region of infection to contain the pathogen and prevent infection from spreading to other parts of the plant. An HR is R-protein-mediated and thus most often associated with ETI (Cunha *et al*., 2006).

#### *2.1.4.2. Other physical changes*

Upon interaction with M/PAMPs, the elevation of cytosolic $Ca^{2+}$ from extracellular and intracellular sources leads to the closure of stomata, through $H_2O_2$ and nitrogen oxide (NO) accumulation in guard cells (Muthamilarasan & Prasad, 2013). The plant also initiates callose deposition in the cell wall to form papillae, which reinforces it to prevent further pathogen penetration. Cytoplasm and organelles, including the nucleus, aggregate at the site of infection as well (Consonni *et al*., 2009; Cui *et al*., 2009; He *et al*., 2007; Henry *et al*., 2013).

M/PAMP-activated defenses have also been shown to lead to a reduction in non-photochemical quenching, which reduces the amount of carbon available to pathogens and thus restricting pathogen growth (Muthamilarasan & Prasad, 2013).



#### *2.1.4.3. Signal transduction and other chemical changes*

M/PAMP perception at the cell membrane immediately induces a cascade of MAPK phosphorylation. This culminates in the transcriptional activation of defense proteins, like WRKY TFs and pathogenesis-related (PR) proteins, which confer resistance against many pathogens through mostly unknown mechanisms (He *et al*., 2007; Ovecka & Lichtscheidl, 2009; Schwessinger & Ronald, 2012).

Other associated changes include variations in ion (e.g. $Ca^{2+}$) flux, calcium-dependent protein kinase (CDPK) activation and the production of reactive oxygen species (ROS), anti-microbial phytoalexins and NO. Phytohormone levels, including that of salicylic acid (SA), jasmonic acid (JA) *et*hylene (ET), gibberellins (GAs), auxins, cytokinins (CKs), brassinosteroids, and abscisic acid (ABA) also fluctuate. SA, for example, is involved in systemic acquired resistance (SAR), is required for complete PTI activation and is especially necessary for the transcription of PR proteins. ET facilitates callose deposition in the cell wall, and, along with JA and ROS, can be elicited by flg22 to induce MAPKs, stress responses and other hormonal signalling further downstream. Cross-talk between all of these hormones also

facilitates precisely tuned, efficient responses (Bari & Jones, 2009). It appears that TFs and co-factors play an important role in mediating cross-talk between SA and JA pathways, meaning that this cross-talk likely takes place at the gene transcription level. Some of these adaptive chemical changes can however come at the cost of growth and development (Consonni *et al*., 2009; Cui *et al*., 2009; He *et al*., 2007; Henry *et al*., 2013; Knogge *et al*., 2009; Pieterse *et al*., 2012; Tena *et al*., 2011).

#### *2.1.4.4. Systemic acquired resistance (SAR) and priming*

HR and the production of ROS, and even just the perception of a M/PAMP, can be associated with elevated resistance to a variety of pathogens throughout the plant. This is known as systemic acquired resistance or SAR and, as mentioned above, it is mediated by defense-related molecules like SA. A related phenomenon is known as priming, whereby the plant becomes resistant to future attack through exposure to M/PAMPs and other biotic/abiotic factors. Priming of the plant was experimentally shown when pre-treatment of Arabidopsis with flg22 caused the up-regulation of RLKs (PRRs) and R-proteins, which in turn resulted in more resistant plants upon later exposure to pathogens. Recent studies also indicate that priming can be associated with epigenetic changes and trans-generational inheritance of resistance (Abramovitch *et al*., 2006; Chinchilla *et al*., 2007; Consonni *et al*., 2009; Cui *et al*., 2009; Mackey & McFall, 2006; Slaughter *et al*., 2012).



Figure 2-2 below eloquently summarises the signalling events involved in plant immunity as has been discussed so far.

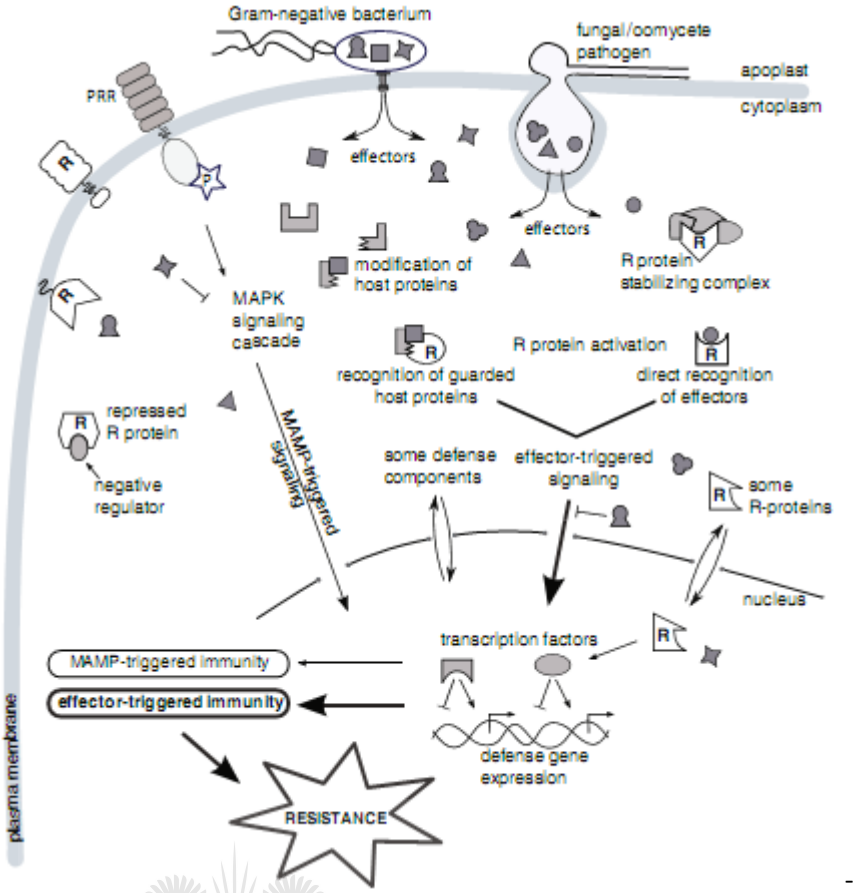





**Figure 2-2: Signalling events involved in plant innate immunity**

Plants have evolved the ability to perceive highly conserved microbe-associated molecular patterns (MAMPs) via transmembrane pattern recognition receptors (PRRs). PRR activation triggers mitogen-activated protein kinase (MAPK) signalling cascades that induce defense gene expression and hinder the growth of some microbial populations. During infection, pathogenic microbes deliver effector proteins into host cells, where they function to suppress or interfere with MAMP-triggered immunity and other defense responses. In resistant plants, cytoplasmic and membrane-associated resistance (R) proteins recognize effectors either directly or indirectly through the surveillance of guarded plant proteins and trigger effector-triggered immunity. Activated R proteins result in genetic reprogramming and pronounced physiological changes in the infected plant cell that ultimately result in resistance (Monaghan *et al*., 2009)**.**

### 2.1.5. Micro- and small interfering RNAs

Plants use RNA interference (RNAi) to achieve transcriptional and post-transcriptional gene silencing of viral RNA. In the former, the viral genome is targeted for methylation; in the latter, viral RNA or the transcripts of viral DNA are cleaved and eliminated (Muthamilarasan & Prasad, 2013). But both bacteria and viruses have evolved mechanisms to overcome RNA silencing. Researchers have in fact identified *P. syringae* effectors that inhibit the production, activation, stability or activity of miRNAs involved in basal defense responses (Navarro *et al*., 2008). In addition, *Solanaceae* studies have recently revealed that secondary siRNAs and microRNAs regulate NB-LRR/LRR immune receptor

genes, and that this may represent a conserved two-part silencing system responsible for modulating large R-gene families (Li *et al*., 2012). Other research has demonstrated that miRNAs fine-tune hormone pathways, including ABA and JA signalling, by targeting genes involved in those pathways and in hormone biosynthesis (Zhang *et al*., 2011).

Evidence is also available indicating that immunity-enhancing epigenetic modifications are heritable in plants, and that small interfering RNAs (siRNAs) may be involved in carrying this 'immune memory' to the next generation (Henry *et al*., 2013; Muthamilarasan & Prasad, 2013).

## 2.2. Leucine-rich repeat receptor-like kinases (LRR-RLKs)

More than 600 RLKs are found in *Arabidopsis*, though not all of them are PRRs. They span the plasma membrane and have an extra-cellular ligand-binding domain, as well as an intracellular kinase domain. Individually and in complexes with other family members, RLKs have been implicated in a host of signal transduction pathways relating to many biologically relevant processes, including development and defense (Greeff *et al*., 2012; Kemmerling *et al*., 2011; Muthamilarasan & Prasad, 2013; Shiu & Bleecker, 2001). However, specific functions for most RLKs remain elusive (Afzal *et al*., 2008; Pel & Pieterse, 2013).



Most of our knowledge about RLKs comes from detailed studies into LRR-RLKs, which play important roles in defense signalling and other cellular functions. Their unique structure facilitates protein-protein interactions. Specifically, in terms of defense-related interactions, LRR-RLKs can interact with the pathogen derived ligands and R-proteins. The known details of the function, location, architecture and the regulation of LRR-RLKs summarised below come from only a few characterised examples, even though there are about 235 LRR-RLKs in the *Arabidopsis* genome (Chae *et al*., 2009; Dunning *et al*., 2007; Goff & Ramonell, 2007; Gou *et al*., 2010; He *et al*., 2007; Kemmerling *et al*., 2011; Nishimura & Dangl, 2010; Shiu & Bleecker, 2001; Tax & Kemmerling, 2012; Zhang, 1998; Zipfel, 2008).

### 2.2.1. Function and Location

LRR-RLKs are majorly involved in plant defense (as PRRs and R-proteins) and in plant growth and development. Transmembrane LRR-RLKs that act as PRRs have an extracellular LRR domain to interact with M/PAMPs/effectors and an intracellular serine/threonine or tyrosine kinase domain to initiate signal transduction pathways through phosphorylation.

Folding and modifications (e.g. glycosylation) occur in the endoplasmic reticulum (ER) and ensure stable, correctly localised receptors. The localisation of FLS2, for example, changes depending on the presence or absence of a ligand. Unbound, it is constitutively recycled between the plasma membrane and the early endosome/trans-Golgi network compartment; bound by flg22, it internalises to late endosomal compartments/multivesicular bodies and is degraded. Since flg22 binds FLS2 irreversibly, it is assumed that the ligand is internalised along with its receptor (Robatzek & Wirthmueller, 2012).

The indispensable role of LRR-RLKs in defense signalling makes them perfect targets for pathogen effectors designed to suppress PTI and ETI. Interestingly, the plant in turn uses another class of LRR-domain proteins, namely nucleotide-binding LRRs (NB-LRRs), to target these effectors (Afzal *et al*., 2008; He *et al*., 2007; Johnson & Ingram, 2005; McGreal, 2009; Nishimura & Dangl, 2010; Rodakowska *et al*., 2009; Shiu & Bleecker, 2001; Zhang, 1998; Zipfel, 2008).

### 2.2.2. Architecture

The interaction between an LRR-RLK and its ligand is as a result of the characteristic horseshoe tertiary configuration of the β-α structure of the LRR-domain, which is normally 20-30 residues long. The LRR domain is also characterised by the consensus sequence LxxLxLxxNxL, in which 'L' is leucine, isoleucine, valine or phenylalanine; 'N' is aspartamine, threonine, serine, or cysteine; and 'x' is any amino acid (Afzal *et al*., 2008; Matsushima *et al*., 2010; Robatzek & Wirthmueller, 2012). The general coiled structure created by a parallel β-sheet on the concave side, with mostly helical elements on the convex side, results from the ordered arrangement of 20-29 residues rich in leucine and other hydrophobic residues. Evidence indicates that ligand binding occurs at the solvent-exposed residues along the LRR β-strand/β-turn. The repetitive nature of this domain enables rapid adaptations to new ligands – exactly what is needed in PRR and R-protein evolution (Chae *et al*., 2009; Dunning *et al*., 2007; Goff & Ramonell, 2007; Kobe & Kajava, 2001; McGreal, 2009; Rodakowska *et al*., 2009; Shiu & Bleecker, 2001; Zipfel, 2008).

### 2.2.3. Regulation

LRR-RLK activity can be enhanced or suppressed depending on the needs of the cell. Not much is known regarding the positive regulation of LRR-RLKs, but it is probably mediated by the direct binding of regulators to the kinase domain, as well as auto- and transphosphorylation. Negative regulation may be achieved by phosphorylation inhibitors, phosphatases, endocytosis and ubiquitin-mediated

protein degradation (Goff & Ramonell, 2007; Johnson & Ingram, 2005). Transcriptional activation/suppression of LRR-RLKs in response to stimuli such as pathogen components and other biotic/abiotic factors may also be very important in their regulation, but more studies are needed to elucidate these mechanisms (Chae *et al*., 2009).

### 2.2.4. FLS2 and BAK1

As its name implies, BRI1-asscoiated kinase 1 (BAK1) was first discovered for its role in brassinosteroid signalling in association with brassinosteroid insensitive 1 (BRI1). It has been established that BAK1 plays a crucial role in resistance against biotrophic and hemibiotrophic pathogens and it is now also thought that BAK1 may be loosely complexed with FLS2, the *Arabidopsis* flagellin receptor, even without flg22 binding. But once flg22 does bind, the kinase domain of FLS2 is phosphorylated and dimerisation with BAK1 occurs almost immediately. Botrytis-induced kinase 1 (BIK1), a cytoplasmic kinase, then interacts with the FLS2/BAK1 complex, undergoes a conformational change and is released in a phosphorylated state to activate two MAPK cascades (MKK4/MKK5-MPK3/ MPK6 and MEKK1/MKK1/MKK2-MPK4), which ultimately lead to the activation of WRKY TFs. BIK1 was also found to phosphorylate the kinase domains of both FLS2 and BAK1, and BIK1-FLS2 interaction subsides after flg22 signalling.



Furthermore, activated FLS2 was recently shown to be rapidly ubiquitinated by the E3 ligases PUB12 and PUB13, which are both associated with BAK1 at the plasma membrane and are phosphorylated by BAK1 upon FLS2 activation.

AtFLS2 was the first plant receptor confirmed as able to recognise and bind M/PAMPs. Though it recognises the 22-amino acid peptide known as flg22, other species have been shown to respond to different conserved sequences. In tomato, for example, a shorter peptide is sufficient while rice is more responsive to the full-length protein (Asai *et al*., 2002; Chinchilla *et al*., 2007; Gómez-Gómez *et al*., 2001; Henry *et al*., 2013; Kemmerling *et al*., 2011; Muthamilarasan & Prasad, 2013; Pandey & Somssich, 2009; Schwessinger & Ronald, 2012; Tena *et al*., 2011).

FLS2 homodimerises independent of flg22-binding, but the contribution of this phenomenon to its function as a receptor is still unknown (Chinchilla *et al*., 2007; Greeff *et al*., 2012). Recent studies also indicate that FLS2 may perceive other M/PAMPs, like the *Xanthomonas oryzae pv. oryzae*-secreted protein Ax21, whose cognate receptor in rice is Xa21 (Danna *et al*., 2011). Greef *et al.* (2012) described

this finding as “astounding” since Ax21 and flg22 are so dissimilar in sequence. Independent studies have so far failed to reproduce these results (Robatzek & Wirthmueller, 2012).

FLS2 can also perceive the CLAVATA3 peptide (CLV3p) endogenously in the shoot apical meristem. CLV3 functions as a key regulator of stem cell homeostasis in this tissue, so this discovery challenges the idea that there is a well-defined discrimination between self and non-self in innate immunity, and it exemplifies the co-evolution of plant peptides and receptor kinase signalling involved in both development and immunity (Lee *et al*., 2011).

While FLS2-flg22 is one of the best understood receptor-PAMP pairs, much remains to be uncovered about downstream signalling events, interactions and transcriptional regulation (Henry *et al*., 2013).

Figure 2.3 below shows the crystal structure of the LRR ectodomain of BAK1 (Sun et al., 2013).

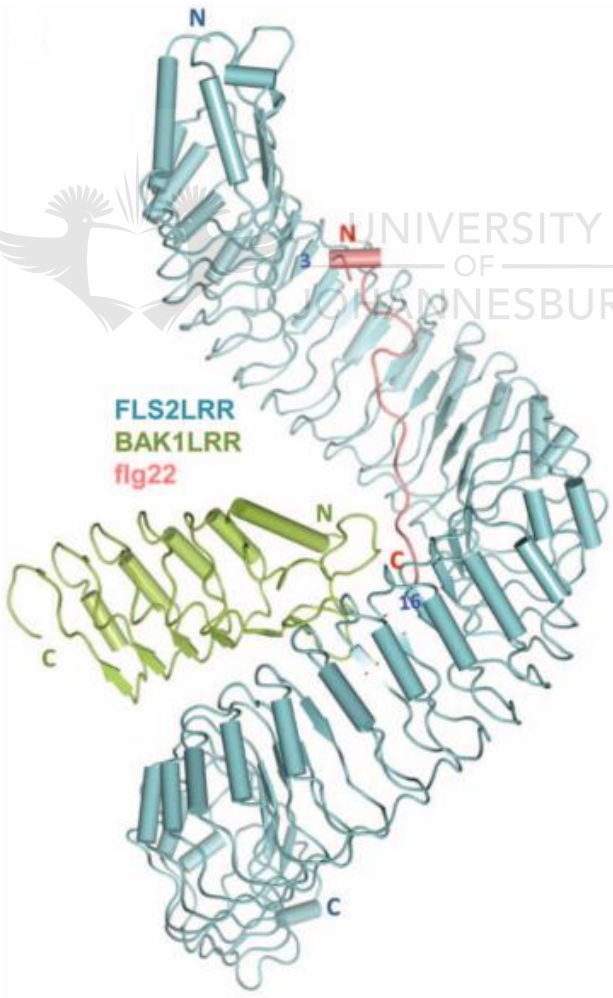





**Figure 2.2-3: Ectodomains mediate the flg22-induced heterodimerisation of FLS2 and BAK1.**

Overall structure of FLS2 LRR-flg22-BAK1 LRR. The positions of LRR3 and LRR16 are indicated by blue numbers. “N” and “C” represent the N and C terminus, respectively. Colour codes are indicated. (Sun *et al*., 2013)

## 2.3. WRKY transcription factors

PTI and ETI is associated with the upregulation of a large number of genes directly involved in defense pathways, including TFs like ethylene-responsive-element-binding factors (ERF), basic-domain leucine-zipper (bZIP) and WRKY proteins. The latter group represents a 74-member family of proteins that tightly regulate defense responses, both positively and negatively, through an intricate interaction network. Known positive regulators include AtWRKY3, -4, 33, -52, -70; known negative regulators include AtWRKY7, -11, -17, -18, -23, -25, -27, -38, -40, -41, -48, -53, -58, -60, and -62. WRKY TFs can modulate direct downstream target genes, other transcription factor genes and even other WRKY genes. They have been shown to activate early defense-related genes and are thus important weapons in the defense arsenal of plants (Muthamilarasan & Prasad, 2013; Pandey & Somssich, 2009; Rushton *et al*., 2010; Singh *et al*., 2002; Zheng *et al*., 2007).



### 2.3.1. Structure

The structure of the WRKY protein is the key to its modulation of defense genes, as well as its own regulation by other cellular components. WRKY proteins are named for the highly conserved WRKYGQK sequence within a 60 amino acid span of β-sheets comprising the N-terminal domain. Two WRKY domains (one at the N- and one at the C-terminal) and two $C_2H_2$ zinc fingers characterise Group I WRKY proteins. Members of Group II and III both contain only one WRKY domain and are distinguished as having either a $C_2H_2$ (cysteine; histidine) or a $C_2HC$ zinc finger, respectively (Journot-Catalino *et al*., 2006; Rushton *et al*., 2010; Singh *et al*., 2002). Figure 2-4 shows the arrangement of the WRKY sequence and the zinc finger within the C-terminal WRKY domain of AtWRKY4, a group I WRKY (Eulgem *et al*., 2000; Rushton *et al*., 2010). Other prominent WRKY motifs include, among others, the N-terminal leucine-zipper of some Group II WRKY TFs, a conserved N-terminal region enriched in Ser/Thr/Pro residues known as the ‘D-motif’ in many Group I WRKY TFs, and a conserved ‘C-motif’ in certain Group II WRKY TFs (Eulgem & Somssich, 2007; Eulgem *et al*., 2000; Zheng *et al*., 2007). Furthermore, Group IId members have the uniquely conserved GHARFRR sequence, the function of which is unknown, as well as a plant-specific zinc-cluster directly preceding their single

WRKY domains. Interestingly, Group IIa WRKY TFs can form heterodimers with potentially different functions. The balance of this dimer formation is disturbed by environmental conditions, mutations and overexpression, which explains why these TFs have been implicated in both positive and negative regulation of defense responses against *Pseudomonas syringae* (Eulgem & Somssich, 2007).

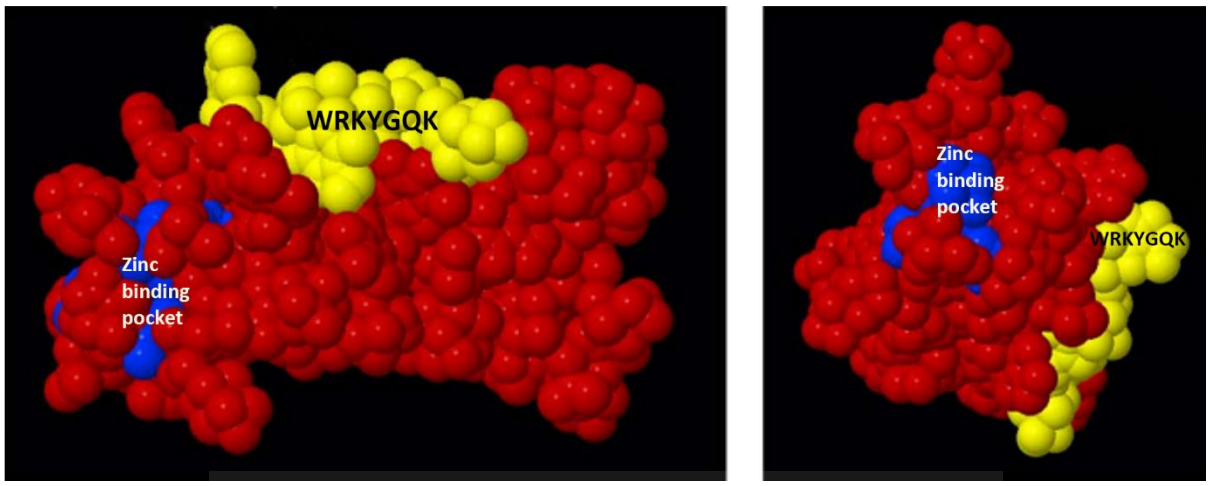



**Figure 2-4: Structural model of the C-terminal WRKY domain of AtWRKY4**

Two views of a spacefill structural model of the C-terminal WRKY domain from AtWRKY4. The model was produced using VENN (Vyas *et* al., 2009) and was based on the solution structure by Yamasaki *et al*. (2005). The WRKYGQK motif is shown in yellow and the cysteines and histidines that form the zinc-binding pocket are shown in blue (Rushton *et* al., 2010**)**.

### 2.3.1. DNA-binding action

The near invariance of the WRKY sequence of this transcription factor family is mirrored by its cognate DNA binding site, the W-box: (C/T)TGAC(C/T). Binding to this W-box is facilitated by the conserved WRKYGQK amino acid sequence of the WRKY protein as well as the tetrahedral co-ordination of a zinc atom by the cysteine and histidine residues of the WRKY domain. Many genes induced by biotic stress, including WRKY TFs themselves, have over-represented and/or clustered W-boxes or W-Box-like sequences within their promoter regions, suggesting a complex transcriptional network regulating stress responses (Eulgem & Somssich, 2007; Rushton *et al*., 2010; Singh *et al*., 2002; Ülker & Somssich, 2004).

## 2.3.2. Regulation

Although the function of WRKY proteins is to regulate genes involved in stress responses (including WRKY TFs themselves), they are also auto- and cross-regulated within the protein family, as well as by a few external mechanisms.

### *2.3.2.1. Auto- and cross-regulation*

Like the promoters of other genes involved in plant immunity, those of WRKY genes are statistically enriched in W-box elements. The implication is that many WRKY TFs are directly controlled, both positively and negatively, by other WRKY TFs via feed forward and feedback regulatory mechanisms. In parsley for example, researchers confirmed M/PAMP-dependant *in vivo* binding of *Pc*WRKY1 to *Pc*PR10, a defense-related gene, as well as its own promoter. Other co-transfection experiments also showed multiple WRKY TFs to interact not only with other WRKY promoters, but with their own promoters. For example, AtWRKY22 and AtWRKY29 can up-regulate their own expression via a positive feedback loop. Another recent study showed that WRKY18, WRKY40 and WRKY60 are not just negative regulators of ABA signalling, but also auto- and cross-regulate themselves negatively (Yan *et al*., 2013). Insertion mutant studies in *Arabidopsis* further confirmed that some WRKY TFs could induce or supress the expression of other family members and it has been shown that those WRKY TFs that enhance defense responses can be counteracted by M/PAMP-inducible WRKY TFs that supress defense responses. This auto- and cross-regulation results in complex feedback loops that allow for tight control of both basal and inducible immunity Eulgem & Somssich, 2007; (Pandey & Somssich, 2009).



### *2.3.2.2. Phosphorylation of WRKY TFs by MAP kinases*

As explained above, PTI is associated with MAPK signalling cascades. It has been shown that a MAPK known to supress salicylic acid signalling, MPK4, can indirectly phosphorylate the ‘D motif’ of both AtWRKY25 and AtWRKY33. It seems then that such WRKY TFs would be the first to be activated in response to M/PAMP-induced MAPK cascades, and may be actually responsible for up-regulating the expression of group II WRKY TFs like AtWRKY22 and AtWRKY29, the promoters of which contain a number of W-boxes. WRKY TFs have also been shown to associate with MAP kinases in the nucleus (Eulgem & Somssich, 2007; Pandey & Somssich, 2009; Rasmussen *et al*., 2012; Zheng *et al*., 2007).

#### *2.3.2.3. Calmodulin*

It has been confirmed that the conserved 'C-motif' domain of Group IId WRKY TFs can bind calmodulin, a signalling protein that binds $Ca^{2+}$ ions. The implication is that these WRKY TFs, like other defense-associated molecules, are able to respond to changes in intracellular $Ca^{2+}$ levels as a result of pathogenic attacks (Eulgem & Somssich, 2007).

#### *2.3.2.4. MicroRNAs*

Available evidence indicates that pathogen attack induces the expression of WRKY TFs that enhance or supress the expression of small RNA molecules – the promoters of several microRNA genes are rich in W-box motifs. Differentially modulated small RNAs can also control WRKY levels by targeting their transcripts (Pandey & Somssich, 2009). One recent study found that sunflower miRNA396, normally involved in development, regulates HaWRKY6: opposite expression patterns of HaWRKY6 and miR396 were observed in response to high temperatures and salicylic acid (Giacomelli *et al.*, 2012).

#### *2.3.2.5. Epigenetic regulation*

AtWRKY38 and AtWRKY62, both suppressors of basal defense responses, have been shown to interact with histone deacetylase 19 (HDA19). This interaction may explain how these WRKY TFs act to modulate gene expression, since HDA19 itself represses transcription by removing acetyl groups from histone tails (Pandey & Somssich, 2009; Rushton *et al.*, 2010).



#### *2.3.2.6. SA-dependant defense responses and pathogenic elicitors*

In a number of plant species, including *Arabidopsis*, treatments with PAMPS, pathogens or salicylic acid were found to induce or differentially regulate the expression of WRKY genes (Zheng *et al.*, 2007). In fact, at least 49 AtWRKY TFs have been shown to be differentially regulated as part of SA-dependant defense responses. The fact that pathogen effectors are able to modulate the expression of WRKY TFs may account for the functional redundancy within the family. For example, AtWRKY7, -11 and -17 mutants were all shown to be susceptible to virulent *P. syringae* (Eulgem & Somssich, 2007; Pandey & Somssich, 2009).

#### *2.3.2.7. Dimerisation*

Group IIa WRKY TFs are able to form homo- and hetero-dimers. Researchers have recorded contradictory results about whether IIa WRKY TFs enhance or suppress basal immunity in response to

*P. syringae*, suggesting that dimerisation may change the function of these TFs and that the process may be sensitive to conditions that cause concentration disturbances, such as mutations, overexpression and environmental factors. WRKY dimerisation is mediated by the N-terminal leucine zipper motif (Eulgem & Somssich, 2007).

## 2.4. Concluding remarks

The literature shows that LRR-RLK surface proteins and WRKY transcription factors play an important role in the different layers of the plant immune response. More research is needed to elucidate the exact functions of specific LRR-RLKs and WRKYs, and the manner in which they are regulated. An examination of the regulatory elements within the promoters of these proteins, as well as their interactions with one another, may be key in determining function and control. As such, the present investigation will ultimately help build a holistic picture of plant defense responses.

# Chapter 3: Materials and Methods

## 3.1. General

All chemicals and enzymes were obtained from manufacturers based in Germany, namely Merck, Carl Roth, Sigma, Fermentas, Qiagen, Invitrogen, Macherey-Nagel, Thermo Fischer Scientific and New England BioLabs, unless indicated otherwise. In cases where commercially available kits were not used, standard protocols for PCR, gel electrophoresis, restriction enzyme digestion, bacterial transformation and DNA purification were used unless indicated otherwise (Sambrook & Russel, 2001).

## 3.2. Instrumentation

**Table 3-1: Instrumentation used during this study**

| Description/Use | Name | Manufacturer |
|---|---|---|
| Bacterial shaker | INNOVA 44 | New Brunswick Science |
| Centrifuge | Biofuge Pico | Heraeus |
| Centrifuge | 5804R | Eppendorf |
| Centrifuge | RC5B Plus | Sorvall |
| DNA gel visualisation | UV light | Bachofer |
| Electrophoresis power supply | EP301 | Amersham Biosciences (now GE Healthcare) |
| Fluorescence microscope | Eclipse 80i | Nikon |
| Gel documentation | Gel Documentation | PEQLAB Biotechnologie |
| Incubator | Memmert | Bachofer |
| Incubator | Thermomixer Comfort | Eppendorf |
| Laminar flow | HERAsafe | Kendro |
| Luminometer | LB940 | Mithras |
| PCR thermocycler | DNA Engine | BioRad |
| Sonicator | Sonoplus | Bandelin |
| Spectrophotometer | Ultrospec 2100pro | Amersham Biosciences (now GE Healthcare) |
| Spectrophotometer | NanoDrop 2000 | PEQLAB Biotechnologie |
| Vacuum infiltration | Diaphragm vacuum pump | Vacubrand |

## 3.3. Bioinformatic Tools and Databases

Bioinformatics were used to compare the structure and function of the genes (4.1) to assess available data (4.2), and to find WRKY candidates for investigation (4.6). Microarray AtGenExpress Pathogen Series data were retrieved from the publically available Arabidopsis electronic Fluorescent Pictograph (eFP) Browser (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi). All data obtained online or experimentally were graphed and analysed using OriginPro 8 (OriginLab Corporation) and Microsoft Excel. Regulatory Sequence Analysis Tools (RSAT; URL: http://rsat.ulb.ac.be/) was used to find over-represented *cis*-elements in the promoters of the genes of interest and GeneSpring GX version 7.3 from Agilent was used to identify WRKY TFs that have similar expression patterns to the genes of interest. The *Arabidopsis* Information Resource (TAIR; URL: www.arabidopsis.org/) was also used to obtain known data about the gene cluster of interest.

## 3.4. Cloning

### 3.4.1. Primers and vector constructs



Below is a list of the primers used in this study, along with their sequences, purposes and melting temperatures, as well as a summary of the Gateway® vectors and constructs used in this study. These primers and vectors are of relevance to the results discussed in 4.3, 4.5, 4.7, and 4.8.

**Table 3-2: The primers used in this study, along with their sequences, purposes and melting temperatures**

Uppercase DNA sequences indicate complementary Gateway® vector sequences, whereas lowercase sequences are target-specific.

| PRIMER NAME | SEQUENCE (5' → 3') | TEMPLATE | PRODUCT DESCRIPTION | $T_M$ |
|---|---|---|---|---|
| At1g51820prom_F | AAAAAGCAGGCTTCacagctgttgctggaactcctgg | gDNA | Full length promoter: -3 bp to -1678 bp | 55°C |
| At1g51820prom_R | AGAAAGCTGGGTCtcttcttactgtccaaaagaaagagcc | | | |
| At1g51850prom_F | AAAAAGCAGGCTTCatgtgattttatgggaaagcaatcttgtt | gDNA | Full length promoter: 0 bp to -833 bp | 55°C |
| At1g51850prom_R | AGAAAGCTGGGTCtgttctccttactgtccacaggagagc | | | |
| At1g51860prom_F | AAAAAGCAGGCTTCactttagtcacacgtctgcttcc | gDNA | Full length promoter: | 55°C |

| Primer Name | Sequence (5' → 3') | Template | Product Description | $T_M$ |
|---|---|---|---|---|
| At1g51860prom_R | AGAAAGCTGGGTCaggtggtcttattcagaggagg | | -10 bp to -1452 bp | |
| At1g51790prom-WBX1m_F | gcattttgttttgggtaaaccaaagaaattaag | At1g51790prom::Luc | Mutation of W-box at -327 bp to -322 bp | 55°C |
| At1g51790prom-WBX1m_R | cttaatttctttggtttacccaaaacaaaatgc | | | |
| At1g51790prom-WBX2m_F | gtctacacatatcgttaccttgttcattaac | At1g51790prom::Luc | Mutation of W-box at -277 bp to -272 bp | 55°C |
| At1g51790prom-WBX2m_R | gttaatgaacaaggtaacgatatgtgtagac | | | |
| At1g51790prom-WBX3m_F | gatatctaccaaagagttaccattgtcttcaagttg | At1g51790prom::Luc | Mutation of W-box at -122 bp to -117 bp | 55°C |
| At1g51790prom-WBX3m_R | caacttgaagacaatggtaactctttggtagatatc | | | |
| At1g51790prom-WBX4m_F | gacattgtcttcaagttgttacctttgaaccc | At1g51790prom::Luc | Mutation of W-box at -102 bp to -97 bp | 55°C |
| At1g51790prom-WBX4m_R | gggttcaaaggtaacaacttgaagacaatgtc | | | |
| At1g51850prom-WBXtwin1m_F | gacttctcaaacatgatcgttacctttttaccgttggattcatcc | At1g51850prom::Luc | Mutation of W-box dyad at -466 bp to -454 bp | 55°C |
| At1g51850prom-WBXtwin1m_R | ggatgaatccaacggtaaaaggtaacgatcatgtttgagaagtc | | | |
| At1g51850prom-WBXtwin2m_F | cgaattgccatacaaaaggtaaaggtaacaaattcatgagaac | At1g51850prom::Luc | Mutation of W-box dyad at -385 bp to -374 bp | 55°C |
| At1g51850prom-WBXtwin2m_R | gttctcatgaatttgttacctttacctttttgtatggcaattcg | | | |
| At1g51890prom-del_F | AAAAAGCAGGCTTCcttggaaggcaagttgggta | At1g51890prom::Luc | Shortened Promoter: -334 bp to -10 bp | 55°C |
| WRKY7_Ful_F | AAAAAGCAGGCTTGatgactgttgagctg | gDNA | Full length coding sequence with stop codon | 65°C |
| WRKY7_stop_R | AGAAAGCTGGGTCctaaagagttttgtcatg | | | |
| WRKY7_C-dom_F | AAAAAGCAGGCTTGatgtcatcaggattc | gDNA | C-terminal domain of coding sequence without stop codon | 65°C |
| WRKY7_no-stop_R | AGAAAGCTGGGTCaagagttttgtcatgattc | | | |
| WRKY22_Full_F | AAAAAGCAGGCTTGatggccgacgattgg | gDNA | Full length coding sequence with stop codon | 65°C |
| WRKY22_stop_R | GAAAGCTGGGTCtcatattcctccggtgg | | | |
| WRKY22_C-dom_F | AAAAAGCAGGCTTGatggcttcttccggtagc | gDNA | C-terminal domain of | |

| Primer Name | Sequence (5' → 3') | Template | Product Description | $T_M$ |
|---|---|---|---|---|
| WRKY22_no-stop_R | GAAAGCTGGGTCatcatcgctaaccaccgt | | coding sequence without stop codon | 65°C |
| WRKY26_Full_F | AAAAAGCAGGCTTGatgggctcttttgatc | gDNA | Full length coding sequence with stop codon | 65°C |
| WRKY26_stop_R | AGAAAGCTGGGTCttatgtctctgtttttcc | | | |
| WRKY26_C-dom_F | AAAAAGCAGGCTTGatgattgagattgtc | gDNA | C-terminal domain of coding sequence without stop codon | 65°C |
| WRKY26_no-stop_R | AGAAAGCTGGGTCtgtctctgtttttccaag | | | |

**Table 3-3: A summary of the Gateway® vectors and constructs used in this study**

| Name | Type | Description | Antibiotic Resistance |
|---|---|---|---|
| pDONR201 | Entry Vector | Donor of cloned fragment to destination vector | Kanamycin |
| pDONR207 | Entry Vector | Donor of cloned fragment to destination vector | Gentamycin |
| pBGWFS7 | Destination Vector | Promoter cloning; GFP expression driven by promoter | Spectinomycin |
| pBGWL7 | Destination Vector | Promoter cloning; Luc expression driven by promoter | Spectinomycin |
| pK7FWG2 | Destination Vector | Constitutive protein expression; C-terminal GFP | Spectinomycin |
| pK7WGF2 | Destination Vector | Constitutive protein expression; N-terminal GFP | Spectinomycin |
| pDEST42 | Destination Vector | Inducible protein expression; C-terminal 6xHis | Ampicillin |
| At1g51790-pDONR201 | Donor construct | Full length promoter fragment | Kanamycin |
| At1g51800-pDONR207 | Donor construct | Full length promoter fragment | Gentamycin |
| At1g51820-pDONR207 | Donor construct | Full length promoter fragment | Gentamycin |
| At1g51850-pDONR201 | Donor construct | Full length promoter fragment | Kanamycin |
| At1g51860-pDONR207 | Donor construct | Full length promoter fragment | Gentamycin |
| At1g51890-pDONR207 | Donor construct | Full length promoter fragment | Gentamycin |

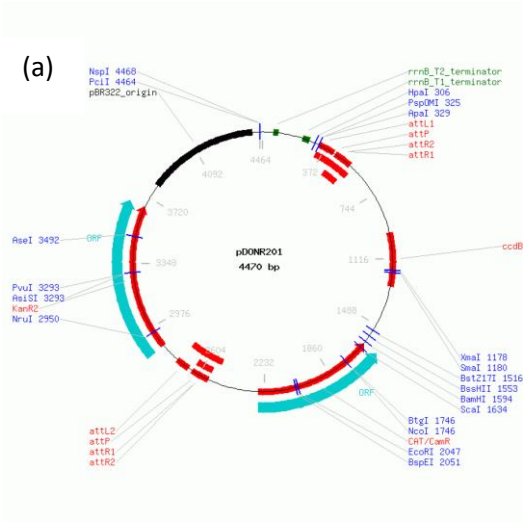
| NAME | TYPE | DESCRIPTION | ANTIBIOTIC RESISTANCE |
|---|---|---|---|
| WRKY7_Full-pDONR207 | Donor construct | Full length protein coding region with stop codon | Gentamycin |
| WRKY7_C-dom-pDONR207 | Donor construct | C-domain of protein coding region without stop codon | Gentamycin |
| WRKY11_Full+Stop-pDONR207 | Donor construct | Full length protein coding region with stop codon | Gentamycin |
| WRKY22_Full+Stop-pDONR207 | Donor construct | Full length protein coding region with stop codon | Gentamycin |
| WRKY22_C-dom-pDONR207 | Donor construct | C-domain of protein coding region without stop codon | Gentamycin |
| WRKY26_Full+Stop-pDONR207 | Donor construct | Full length protein coding region with stop codon | Gentamycin |
| WRKY26_C-dom-pDONR207 | Donor construct | C-domain of protein coding region without stop codon | Gentamycin |
| At1g51790prom::GFP | Reporter construct (pBGWFS7) | Full length promoter driving GFP expression (-1 bp to -1565 bp | Spectinomycin |
| At1g51850prom::GFP | Reporter construct (pBGWFS7) | Full length promoter driving GFP expression (0 bp to -833 bp) | Spectinomycin |
| At1g51890prom::GFP | Reporter construct (pBGWFS7) | Full length promoter driving GFP expression (-10 bp to -1349 bp) | Spectinomycin |
| At1g51790prom::Luc | Reporter construct (pBGWL7) | Full length promoter driving Luc expression (-1 bp to -1565 bp) | Spectinomycin |
| At1g51800prom::Luc | Reporter construct (pBGWL7) | Full length promoter driving Luc expression (-38 bp to -1708 bp) | Spectinomycin |
| At1g51820prom::Luc | Reporter construct (pBGWL7) | Full length promoter driving Luc expression (-3 bp to -1678 bp) | Spectinomycin |
| At1g51850prom::Luc | Reporter construct (pBGWL7) | Full length promoter driving Luc expression (0 bp to -833 bp) | Spectinomycin |
| At1g51860prom::Luc | Reporter construct (pBGWL7) | Full length promoter driving Luc expression (-10 bp to -1452 bp) | Spectinomycin |
| At1g51890prom::Luc | Reporter construct (pBGWL7) | Full length promoter driving Luc expression (-10 bp to -1349 bp) | Spectinomycin |
| At1g51790prom-DELm1::Luc | Reporter construct (pBGWL7) | Shortened promoter driving Luc expression (-3 bp to -712 bp) | Spectinomycin |
| At1g51850prom-DELm1::Luc | Reporter construct (pBGWL7) | Shortened promoter driving Luc expression (0 bp to -630 bp) | Spectinomycin |
| At1g51850prom-DELm2::Luc | Reporter construct (pBGWL7) | Shortened promoter driving Luc expression (0 bp to -368 bp) | Spectinomycin |

| NAME | TYPE | DESCRIPTION | ANTIBIOTIC RESISTANCE |
|---|---|---|---|
| At1g51850prom-DELm3::Luc | Reporter construct (pBGWL7) | Shortened promoter driving Luc expression (0 bp to -134 bp) | Spectinomycin |
| At1g51890prom-DELm1::Luc | Reporter construct (pBGWL7) | Shortened promoter driving Luc expression (-10 bp to -742 bp) | Spectinomycin |
| At1g51890prom-DELm2::Luc | Reporter construct (pBGWL7) | Shortened promoter driving Luc expression (--10 bp to 334 bp) | Spectinomycin |
| At1g51790prom-WBX1m::Luc | Reporter construct (pBGWL7) | Full length promoter with mutation of W-box at -327 bp to -322 bp | Spectinomycin |
| At1g51790prom-WBX2m::Luc | Reporter construct (pBGWL7) | Full length promoter with mutation of W-box at -277 bp to t-272 bp | Spectinomycin |
| At1g51790prom-WBX3m::Luc | Reporter construct (pBGWL7) | Full length promoter with mutation of W-box at -122 bp to -117 bp | Spectinomycin |
| At1g51790prom-WBX4m::Luc | Reporter construct (pBGWL7) | Full length promoter with mutation of W-box at -102 bp to -97 bp | Spectinomycin |
| At1g51850prom-WBXtwin1m::Luc | Reporter construct (pBGWL7) | Full length promoter with mutation of W-box at -466 bp to -454 bp | Spectinomycin |
| At1g51850prom-WBXtwin2m::Luc | Reporter construct (pBGWL7) | Full length promoter with mutation of W-box at -385 bp to -374 bp | Spectinomycin |
| At1g51850prom-WBXtwin1x2m::Luc | Reporter construct (pBGWL7) | Full length promoter with mutation of W-box at -385 bp to -374 bp and at -466 bp to -454 bp | Spectinomycin |
| 35S::WRKY7_Full+Stop | Constitutive protein expression construct (pK7FWG2) | Full length protein; untagged (stop codon) | Spectinomycin |
| 35S::WRKY11_Full+Stop | Constitutive protein expression construct (pK7FWG2) | Full length protein; untagged (stop codon) | Spectinomycin |
| 35S::GFP-WRKY11_Full+Stop | Constitutive protein expression construct (pK7WGF2) | Full length protein; N-terminal GFP tagged | Spectinomycin |
| 35S::WRKY22_Full+Stop | Constitutive protein expression construct (pK7FWG2) | Full length protein; untagged (stop codon) | Spectinomycin |
| 35S::WRKY26_Full+Stop | Constitutive protein expression construct (pK7FWG2) | Full length protein; untagged (stop codon) | Spectinomycin |
| WRKY7_C-dom-pDEST42 | Inducible protein expression construct | C-terminal domain of protein; 6xHis tagged | Ampicillin |
| WRKY11_C-dom-pDEST42 | Inducible protein expression construct | C-terminal domain of protein; 6xHis tagged | Ampicillin |
| WRKY22_C-dom-pDEST42 | Inducible protein expression construct | C-terminal domain of protein; 6xHis tagged | Ampicillin |
| WRKY26_C-dom-pDEST42 | Inducible protein expression construct | C-terminal domain of protein; 6xHis tagged | Ampicillin |

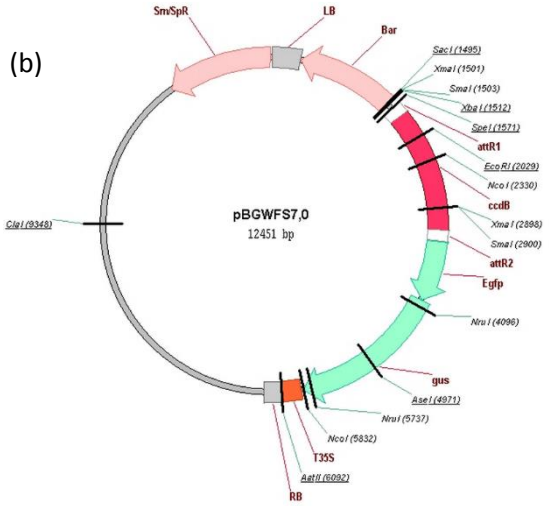
## 3.4.2. Gateway® vector maps

Vector maps and descriptions were obtained from the URLs https://www.lablife.org/ll and http://gateway.psb.ugent.be/search.

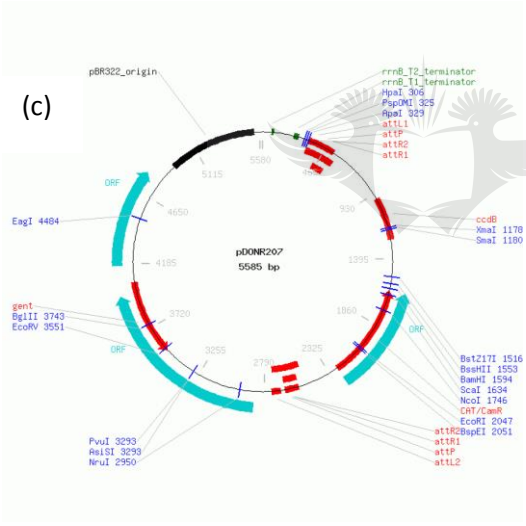
(a)



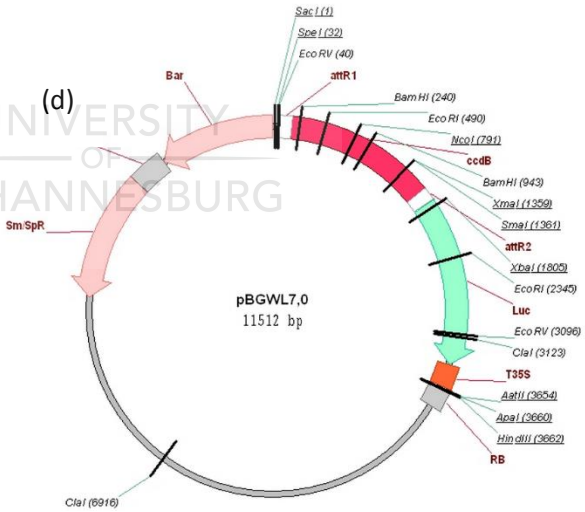

(b)

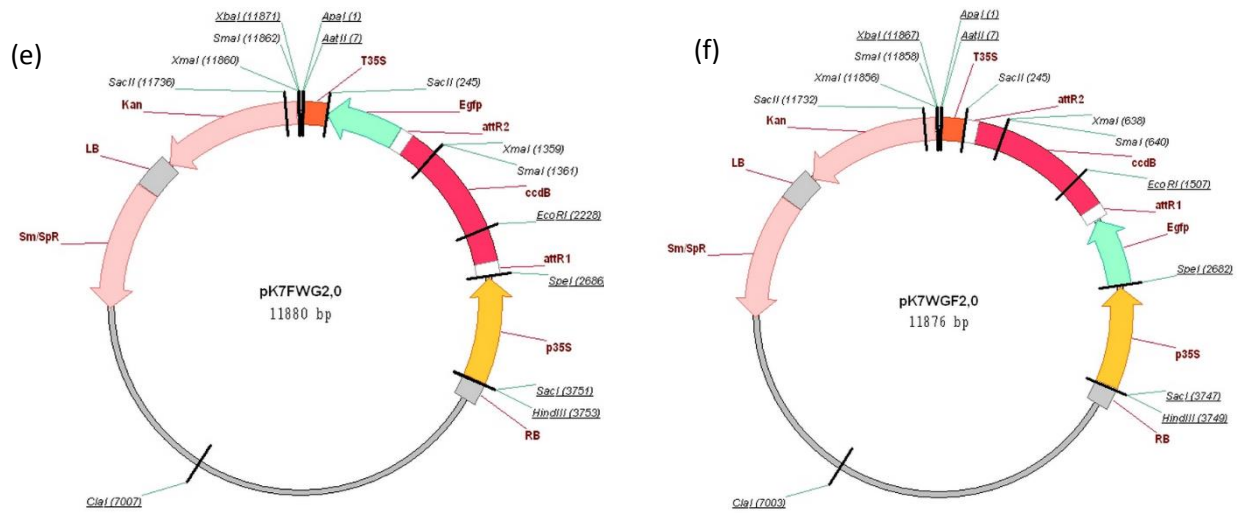



(c)




(d)





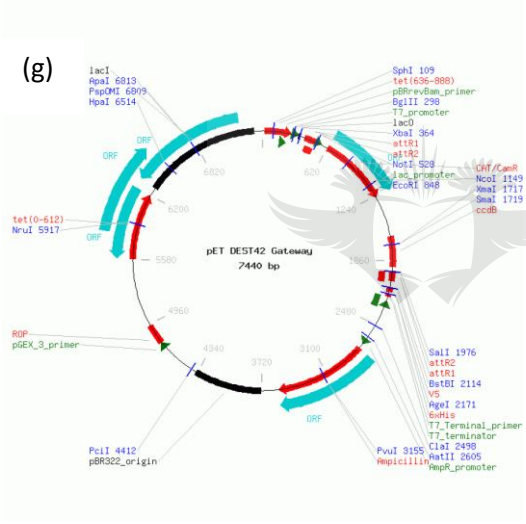






**Figure 3-1: The maps of the vectors used to create the constructs listed in Table 3-3.**

**(a) pDONR201,** (**b**) **pBGWFS7,** (**c**) **pDONR207 ,** (**d**) **pBGWL7,** (**e**) **pK7FWG2,** (**f**) **pK7WGF2,** and **(g) pDEST42. Vector maps and descriptions were obtained from the** URLs https://www.lablife.org/ll **and** http://gateway.psb.ugent.be/search.

### 3.4.3. Gateway® cloning

The Invitrogen Gateway® cloning system was used for all cloning reactions, to create the constructs mentioned previously, using the manufacturer's instructions as a guideline. Here follows the general procedure as well as alterations that were made to the recommended protocol:

Gene-specific primers were designed with part of the attenuation sites needed for cloning. Two rounds of PCR were required to produce the desired product: the first round of PCR produces a product with incomplete attenuation sites, so a second round of PCR is performed with standard attenuation site primers to produce the product shown in Figure 3-2 below:

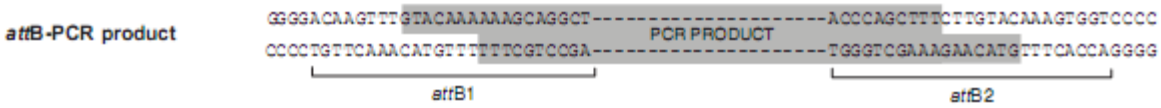



**Figure 3-2: The product obtained after the two rounds of Gateway® PCR**.

(Life Technologies Corporation, 2010)



Once the final PCR product was produced, the first cloning (BP) reaction was performed overnight at 25°C to insert the fragment into an entry vector. Four µl of the BP reaction was then transformed using heat shock treatment into a 250 µl aliquot of competent DH5α/Top10 *E. coli* cells, shaken for 1 h at 37°C and plated onto an LB-agar plate (15 g/l of Agar in LB media, which was made using 10 g/l Bacto-Tryptone, 5 g/l Bacto-Yeast and 5 g/l NaCl in Milli-Q $H_2O$) with the relevant antibiotic (see Table 3-3) to be grown overnight at 37°C. Colonies were then picked and placed into 2 ml LB media with the relevant antibiotic. The plasmids were purified the next day using a Qiagen mini-prep kit and positive clones were confirmed by restriction analysis and subsequent sequencing (GATC Biotech, Germany).

The second cloning (LR) reaction transfers the fragment from the entry vector to the destination vector. This reaction was also performed at 25°C overnight using the purified entry clone and destination vector of choice (a quarter of the recommended reaction components were used with satisfactory efficiency). This reaction was then transformed as above and clones were confirmed by restriction analysis. Figure 3-3 summarises the recombination procedure.

**Gateway® Recombination Reactions**

The Gateway® Technology uses the lambda recombination system to facilitate transfer of heterologous DNA sequences (flanked by modified *att* sites) between vectors (Hartley *et al.*, 2000). Two recombination reactions constitute the basis of the Gateway® Technology:

- **BP Reaction:** Facilitates recombination of an *att*B substrate (*att*B-PCR product or a linearized *att*B expression clone) with an *att*P substrate (donor vector) to create an *att*L-containing entry clone (see diagram below). This reaction is catalyzed by BP Clonase™ enzyme mix.


- **LR Reaction:** Facilitates recombination of an *att*L substrate (entry clone) with an *att*R substrate (destination vector) to create an *att*B-containing expression clone (see diagram below). This reaction is catalyzed by LR Clonase™ enzyme mix.


**Figure 3-3: A graphical representation of the Gateway® cloning system.**

Adapted from the Invitrogen Gateway® Technology Manual (Life Technologies Corporation, 2010).



### 3.4.4. Primer design

Gateway® compatible primers that had to be manually designed and ordered from Eurofins MWG Operon, were checked for primer pair compatibility and target specificity using software, the URLs of which are as follows:

http://blast.ncbi.nlm.nih.gov (Blastn)

http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/ (Integrated DNA Technologies' Oligo Analyzer 3.1)

http://frodo.wi.mit.edu/primer3/ (Primer 3)

### 3.4.5. Promoter cloning

The promoter lengths (around 1500 bp or less) were selected based on putative regulatory elements, previous experiments, personal communication and readily available constructs. The promoters are named according to their gene identity, with the suffix 'prom' to differentiate promoter constructs

from gene constructs (e.g. At1g51790prom; see Table 3-3). The promoters of At1g51790 and At1g51890 were cloned using primers and entry clones already available (Kemmerling Lab, ZMBP). At1g51800prom was donated in entry vector pDONR207 by Harald Keller (Institut Sophia Agrobiotech, France). At1g51800prom, At1g51820prom and At1g51860prom were newly cloned for this study. The primer sequences used to clone At1g51850prom were originally designed by Andrea Gust (ZMBP) for a related study (Gust *et* al., 2007), but were adapted in this study for Gateway® compatibility. All of the above promoter clones are henceforth referred to as the "full length" promoters. Primers were also designed to clone a shorter promoter fragment of At1g51890 as it had no convenient restriction sites in this region. The resulting construct was named 'At1g51890prom-DELm2::Luc' ('Luc' being short for luciferase) to fit in with the deletion mutants created by restriction enzymes (see 3.4.9).

## 3.4.6. WRKY TF cloning

The full length untagged and the functional C-terminal domains of each WRKY TF were required for protoplast transient expression assays (4.7), and DPI-ELISA experiments (4.8) respectively.

### *3.4.6.1. Full length WRKY TFs*



Gateway compatible primers were designed in frame to span the entire coding sequence of WRKY7, WRKY22 and WRKY26. A stop codon was included in the reverse primer to ensure an untagged product for constitutive expression in pK7FWG2 (without a stop codon a C-terminal GFP tag would have resulted). These WRKY TF constructs were named to indicate constitutive expression of the full length coding sequence with its stop codon e.g. 35S::WRKY7_Full+Stop. The primers and entry constructs are similarly named (see Table 3-2 and

Table 3-3). WRKY11 was already available as a full length construct with a stop codon in pDONR207 (Wanke lab, ZMBP) and this was used to create an untagged version (in pK7FWG2) and an N-terminal GFP-tagged version (in pK7WGF2 - 35S::GFP-WRKY11_Full+Stop).

### *3.4.6.2. Functional C-terminal domain WRKY TFs*

The functional C-terminal domains of WRKY7, WRKY22 and WRKY26 were cloned using primers designed to include the C-terminal region coding for the 'WRKYGQK' amino acid motif. A start codon was included in frame. These constructs were named to indicate that it is the functional C-terminal domain only in the inducible expression vector pDEST42 e.g. WRKY7_C-dom-pDEST42. Again the C-

terminal domain of WRKY11 was already available in pDEST42. The destination vector pDEST42 was chosen because the product is 6xHis-tagged at the C-terminal.

## 3.4.7. PCR

A variety of PCR techniques were employed during this project to confirm melting temperatures, amplify fragments from genomic DNA, cDNA and plasmids, validate sequencing primers and perform site-directed mutagenesis.

### *3.4.7.1. Gradient and optimisation PCR*

Temperature gradient PCR was performed using *Taq* polymerase to optimise the annealing temperatures of almost all primers designed for this study. Standard PCR reaction components were used: $dH_20$, dNTPs, *Taq* buffer, gDNA, forward and reverse primers and DMSO if needed. A basic PCR program was used and adapted based on the results. Initial denaturation took place at 95°C for 2 min, followed by denaturation at 95°C for 15 s and primer annealing at the appropriate temperature for 45 s, and elongation was then performed at 72°C for 2 min/kilobase of target DNA. The denaturation, primer annealing and elongation steps were repeated 40 times, with a final elongation at 72°C for 10 min. The samples were kept at 10°C until needed.



### *3.4.7.2. Amplification for cloning*

High fidelity DNA polymerases *Pfu* and Phusion was used to create the fragments for cloning to eliminate as far as possible the chance of base mutations during amplification. Standard reaction components and concentrations were used according to the DNA polymerase manufacturer's instructions. The two step Gateway® PCR program was used with temperatures and times depending on the length of the target and DNA polymerase used (according to the manufacturer's instructions) and other optimisation variations.

The first Gateway® step involved initial denaturation at 95°C for *Pfu* or 98°C for Phusion, followed by 40 cycles of denaturation, primer annealing (this temperature was empirically determined by *Taq* PCR) and elongation. The samples were kept at 10°C until the second step could be performed. The second step was similar, except that primer annealing was at 45°C for 5 cycles, followed by 55°C for 20 cycles with a final elongation and storage at 10°C.

Twenty µl reactions were used for both reactions; 4 µl of the GW-1 reaction was used as template for the GW-2 reaction. The GW-2 reaction uses the standard Gateway adapter primers to amplify the template from GW-1 (see Figure 3-2). Depending on the success of the reaction, it was sometimes necessary to perform GW-1 as a full 30 cycle PCR, run the product on an agarose gel, purify it and then use it as a template for GW-2.

#### *3.4.7.3. Site-directed mutagenesis*

To create the W-box mutations (TTGAC to TTACC and GTCAA to GGTAA) in At1g151790prom and At1g51850prom, *Pfu* polymerase was used according to an established PCR protocol (Fisher & Pei, 1997). Eighteen cycles of PCR were performed at a standard annealing temperature of 55°C and an elongation temperature of 68°C, using the destination vectors of the full length promoters as template (At1g51790prom::Luc and At1g51850prom::Luc). The complementary mutagenesis primers were designed to be around 30 bp with the target centrally located (see Table 3-2). The PCR products were digested overnight with 1 µl *DpnI* to remove methylated template DNA. The digested PCR products were then transformed as previously described. Positive clones were confirmed using a sequencing primer designed specifically for pBGWL7 during this study.



### 3.4.8. Positive clone selection

As mentioned, all clones were checked by restriction analysis before sequencing was done. The freeware program pDRAW32 (AcaClone) was used to determine the restriction sites, and hence product sizes, on vector maps constructed using the sequences of the plasmids and PCR products. Digests were then performed using 0.25 µl of enzyme in a 20 µl reaction for 3 or more hours. Other reaction components were added according to the manufacturer's recommendation for each enzyme/enzyme combination. The full list of NEB and Fermentas restriction enzymes used during this study for various purposes is as follows: *EcoRI, EcoRV, PstI, XbaI, HindIII, HindII, BamHI, PvuI, FspI, NheI, PsiI, BglII, SpeI, ApaLI, MfeI, PflMI, NdeI, SacI, NotI, SnaBI (Eco105I), BspHI (PagI), AseI (VspI), DpnI.*

### 3.4.9. Generation of deletion mutants

Some restriction sites were conveniently located in the promoters of At1g51790, At1g51850 and At1g51890, enabling the creation of deletion constructs by restriction digests of the relevant pBGWL7

constructs. The following enzymes were used in combination with *SacI* (*SacI* cuts in pBGWL7 itself, near the start of the promoter): *BglII* created At1g51790prom-DELm1::Luc; *SnaBI* created At1g51850prom-DELm1::Luc; *BspHI* created At1g51850prom-DELm2::Luc; *AseI* created At1g51850prom-DELm2::Luc and At1g51890prom-DELm1::Luc. Large amounts of purified DNA (10 to 15 µg) were subjected to an overnight digestion at 37°C by a total of 1.5 µl restriction enzyme (in ratios as per the manufacturer's recommendation). The DNA was purified from the reaction the next day and run on an agarose gel to separate the excised fragment from the remaining linear deletion mutant. A blunting reaction using 1 µl of *Pfu* polymerase, 1 µl dNTPs and 25 µl of the purified linear plasmid (total volume = 30 ul) was then performed at 74°C for 30 min. Again the DNA was purified from this reaction and incubated with T4 DNA ligase (the reaction mix was pipetted according to the manufacturer's instruction) for 1 to 5 h to generate a circular plasmid. This reaction was then transformed and clones were screened as described before.

Below is a representative restriction analysis gel showing the successful creation of the At1g51790prom deletion mutant, At1g51790prom-DELm1.

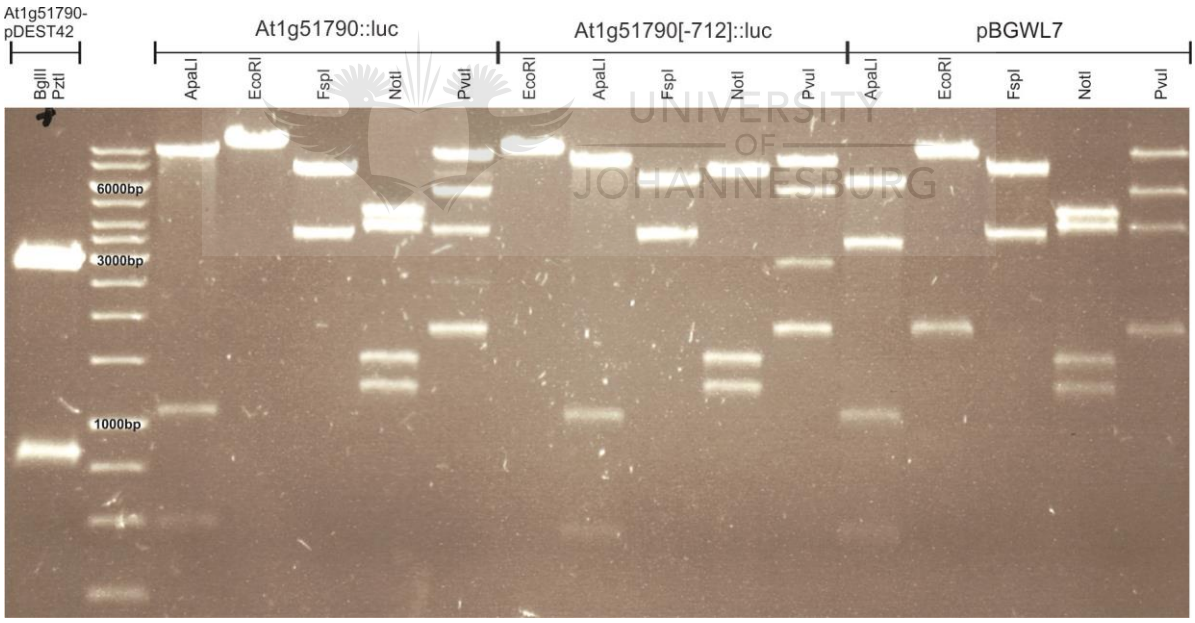



**Figure 3-4: A representative gel showing the restriction digests of the original At1g51790::Luc construct, the deletion mutant At1g51790[-712]::Luc and the vector pBGWL7.**

### 3.4.10. Agarose gel electrophoresis

Agarose gels were made using Tris-acetate-EDTA (TEA) buffer at concentrations ranging between 0.8% and 2% depending on the expected DNA product sizes. Ethidium bromide was added at a concentration of 0.005%. The gels were used to check the sizes and, if necessary, excise DNA from PCR products and restriction digest reactions. Most often, 20 µl of sample (with a few µl of Fermentas loading dye) was loaded, depending on the size of the sample and the well. Standard DNA ladders from Fermentas were used and all electrophoresis was done at a limiting 100 V (unlimited Amperes) for a time required for the desired resolution. Gels were viewed under a UV light and clean scalpels were used to excise DNA bands.

### 3.4.11. DNA purification

All DNA purification for cloning was performed and validated using commercially available kits from Qiagen, Thermoscientific, Machery Nagel and Fermentas, according to the manufacturer's instructions and depending on the source and purpose of the DNA.



## 3.5. Transient expression assays: promoter analysis

The standard protocol devised by the Sheen Lab at Harvard (Yoo *et* al., 2007) was adapted for the present study by the Brunner Group at the ZMBP (URL: http://www.zmbp.uni-tuebingen.de/plant-biochemistry/research-groups/brunner.html). All results obtained in 4.5 to 4.7 were based on this method.

### 3.5.1. Plant growth

Col-0 *Arabidopsis thaliana* plants were grown in an environmentally-controlled phytochamber (22 °C, 16 h of light). The plants were treated regularly for nematode infections, were watered twice per week and were kept without lids in order to prevent high humidity. Seedlings were kept in the fridge for two days before growing them to a four-leaf state and separating into individual pots.

### 3.5.2. Protoplast isolation

Well expanded leaves from 5 week old plants were cut into strips as narrow as possible (0.5 mm or less) using a razor blade. Two leaves were cut at once, discarding the top and stalk of the leaf. Using clean flat-tip forceps and making sure that all leaf surfaces were exposed to the enzyme solution (ES), as many leaf strips as possible were transferred into the freshly prepared ES (made using 20 mM MES at pH 5.7, 0.4 M mannitol, 20 mM KCl, 10 mM $CaCl_2$, 0.1% w/v protease-free BSA, and from Yakult pharmaceuticals, 0.4% w/v macroenzyme R10 and 1.5% w/v cellulase R10). A Petri dish of desired size was used for the ES depending on the volume of protoplasts needed: 10 ml ES yielded 5 to 10 ml of protoplasts at a concentration of $2x10^5$ protoplasts/ml. After a 30 min vacuum infiltration in the dark, the solution was left in the dark without shaking for 3 h at room temperature (20 to 22°C) in an air-conditioned room. Next the solution was gently shaken until it turned green due to the release of protoplasts from the leaf strips. Ten millilitres of W5 (2 mM MES at pH 5.7, 154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl), which had been stored at -20°C, was then added for every 10 ml of ES and this mixture was filtered through a nylon mesh into 12 ml cell culture tubes. The tubes were then centrifuged at 200 x g (4°C) for 1 min. The supernatant was subsequently removed and discarded. Next W5 was added to the tubes such that the total volume of protoplasts would be 2 ml regardless of the starting volume of ES. The tubes were left on ice for about 2 h after a sample of the protoplasts was taken to be counted, using a haemocytometer and microscope, in order to calculate the solution volume required for a working concentration of $2x10^5$ protoplasts per ml of MMG (4 mM MES at pH 5.7, 0.4 M mannitol, 15 mM $MgCl_2$). After this 2 h rest, the supernatant was removed from the settled protoplasts, which were then resuspended in the calculated volume of freshly prepared MMG.



### 3.5.3. Protoplast transformation

The protoplasts (suspended in MMG at a concentration of $2x10^5$ protoplasts/ml) were pipetted into pre-prepared tubes containing 10 µg of plasmid for every 100 µl of protoplasts. After gently shaking the tubes to mix the DNA and the protoplasts, a freshly prepared PEG solution (40% w/v PEG4000, 0.2 M mannitol, 100 mM NaCl, 100 mM $CaCl_2$) was added in a v/v ratio of 1.1:1 to the protoplast volume. Following gentle mixing by slowly inverting the tubes a few times, they were left to transform for 5 to 10 min. The W5 solution was then added in a v/v ratio of 4.4:1 of the protoplast volume in order to stop the transformation reaction. After gentle mixing, the tubes were centrifuged at 200 x g for 1 min (4°C). The supernatant was then removed and a volume of freshly prepared W1 (4 mM MES at pH 5.7, 0.5 M mannitol, 20 mM KCl) equal to the original volume (such that the concentration is again $2x10^5$

protoplasts/ml) was pipetted into the tube to resuspend the protoplasts. For the luciferase-based assays, 1 µl of 20 mM luciferin was added per 100 µl protoplasts before resting the tubes horizontally overnight. To make the luciferin solution, 5 mg of the free acid form of D-Luciferin (PJK) was converted to the soluble potassium salt by incrementally adding 17.8-18.1 µl of diluted KOH equivalent, with shaking, until the solution turned bright yellow (Brunner Lab, ZMBP). It was stored in 100 µl aliquots at -20°C.

### 3.5.4. Fluorescence microscopy

Where protoplasts were transformed with GFP expressing constructs, they were visualised under a fluorescent microscope to confirm transformation.

### 3.5.5. Luminometry

As mentioned in 3.5.3, luciferin was added to the transformed protoplasts before resting overnight. The following day the tubes were inverted carefully to re-suspend the protoplasts, which were then aliquoted into an opaque white 96-well micro-titre plate (100 µl per well). After measuring the background luminescence in a luminometer, 1 µl of distilled water and 1 µl of 10 µM Flg22 (dissolved in $dH_2O$) were added to the negative control (untreated) and test (treated) samples respectively. Three technical replicates each of treated and untreated aliquots were used. Luminometric readings were taken every hour for 6 h. The kinetic time points and Flg22 concentrations were empirically optimised and the data were converted into spreadsheets and graphed using OriginPro 8.

## 3.6. DNA-Protein-Interaction (DPI)-ELISA

This method applies to the results discussed in 4.8.

### 3.6.1. WRKY protein expression

All WRKY TF expression and purification were done according to an established method (Brand *et* al., 2010). After cloning and propagating the WRKY#_C-dom-pDEST42 constructs as described in 3.4.6, they were all transformed into BL21-AI inducible expression *E. coli* from Invitrogen, except for

WRKY11_C-dom-pDEST42, which was already available in BL21(DE3) cells (Wanke Laboratory, ZMBP). After standard positive clone selection based on restriction enzyme digests, said clones were inoculated into 5 ml of LB-media (containing strain- and plasmid-specific antibiotics – see Table 3-3) in duplicate. The culture was shaken overnight at 37°C (180 rpm). The next day the 5 ml cultures were added to 200 ml of LB media without selection antibiotics and the optical density at 595 nm ($OD_{595}$) was measured to be between 0.05 and 0.1. After shaking at 37°C (180 rpm) for about 2 h, when the $OD_{595}$ reached between 0.4 and 0.6, the inducing agent IPTG (Isopropyl-β-D-thio-galactoside) was added at a final concentration of 1 mM. For the BL21-AI cells, arabinose was also added, according to the manufacturer's recommendation, to a final concentration of 0.2%. An un-induced negative control was included for confirmation of expression by Western blotting. The cultures were then shaken at 30°C (180 rpm) until the $OD_{595}$ reached between 0.8 and 1.3 (around 4 to 6 h). The cultures were then centrifuged at 2200 x g for 20 min (4°C) after removing 1 ml for Western blotting, and the supernatant was removed before freezing the pellet at -20°C overnight.

### 3.6.2. Western blotting

The 1 ml induced and un-induced samples were centrifuged at maximum speed in a bench top centrifuge before re-suspending the pellet in 150 µl sample buffer (375 mM Tris/HCl, 38.7% v/v glycerol, 12.3% w/v SDS, 6% β-mercaptoethanol, 0.06% bromophenol blue). The samples were then incubated at 95°C for 10 min in order to extract the proteins under denaturing (non-native) conditions. Ten µl of each sample were loaded, along with Fermentas' PageRuler™ Plus Prestained Protein Ladder marker, onto a 15% SDS-PAGE gel. The resolving gel (5 ml) was made with 1.1 ml $dH_2O$, 2.5 ml 30% acrylamide, 1.3 ml 1.5 M Tris (pH 8.8), 50 µl 10% Sodium dodecyl sulphate (SDS), 50 µl 10% Ammonium persulfate (APS) and 2 µl N, N, N', N'-tetramethylethylenediamine (TEMED). The stacking gel (2 ml) comprised of 1.4 ml $dH_2O$, 330 µl 30% acrylamide, 250 µl 1.0 M Tris (pH 6.8), 20 µl 10% SDS, 20 µl 10% APS and 2 µl TEMED.

The gel, submerged in SDS running buffer (250 mM Tris, 1.9 M glycine, 0,1% SDS), was run at a limit of 170 V (unlimited Amperage) for about 1 h. The proteins were then blotted onto a nitrocellulose membrane activated by methanol at 200 V for 2 h (4°C). The membrane was then blocked in PBS-T containing 5% milk powder overnight at 4°C. The following day the membrane was washed three times by shaking for 10 min in PBS-T at room temperature.

### 3.6.3. Crude protein extraction

The frozen bacterial pellets (from 0) were thawed on ice, re-suspended in 50 ml of ice cold Tris/NaCl and transferred into 50 ml Falcon tubes. The tubes were then centrifuged at 2200 x g for 20 min (4°C) before discarding the supernatant. DPI-Ex-buffer (for 11 extractions: 1 ml 0.2 M HEPES in KOH at pH 7.4, 5 ml 1 M KCl, 4 ml v/v glycerol, 32.75 ml $ddH_20$ and 1 tablet of Roche Complete Proteinase Inhibitor w/o EDTA) was used to resuspend the pellet for sonication on ice at 75% power (6 repeats of 15 s each). The tubes were again centrifuged at 2200 x g for 20 min (4°C), after which the supernatant was transferred into a new tube.

In order to measure the protein concentration of the extract, a standard Bradford assay was performed. A standard curve was compiled using 5 µg to 30 µg of BSA dissolved in Bradford (BioRad) reagent such that the readings of 3 dilutions of extract could be averaged. To obtain a final volume of 4 ml crude protein extract, 400 µl of 2% biotin-free BSA and 20 µl of 1 M DTT were added to 3.58 ml of the extract. This was aliquoted into smaller volumes as required by the DPI-ELISA experiments, and stored at -20°C for up to 4 weeks.



### 3.6.4. DPI-ELISA

Double stranded (DS) oligo probes were prepared in advance using single stranded 30 bp primers. Only the promoter of At1g51850 was analysed using this technique. Three oligos had to be designed for each W-box dyad and its mutant under investigation: one biotin tagged sense oligo (denoted with the suffix '-S-B'), one untagged sense oligo (denoted with the suffix '-S') and one 100% complementary untagged anti-sense oligo (denoted with the suffix '-A'). These primers were ordered from biomers.net GmbH with HPLC purification. Refer to Table 3-2 for the primer sequences. The experiment required a biotinylated DS probe for basic interaction studies and another un-biotinylated DS probe for the competition assays. The DS probes were prepared by heating the following reaction components at 95°C (in a thermocycler) for 3 min and then gradually decreasing the temperature to 25°C: 20 µl of 10 µM biotinylated/unbiotinylated sense oligo, 20 µl of 10 µM anti-sense oligo, 10 µl of 10X Annealing Buffer (8 ml 1 M Tris/HCl at pH 7.5 – 8, 2 ml 2 M $MgCl_2$, 5ml 2 M NaCl and 5 ml $ddH_2O$) and 50 µl $ddH_2O$.

The DS oligos were stored at -20°C until needed, at which time they were diluted 1:60 in TBS-T (1X solution with 0.1% Tween of a 10X solution of 12 g Tris base, 52.6 g NaCl added up to 500 ml $ddH_2O$

and adjusted to pH 7.5 using HCl). Sixty µl of the diluted DS oligos were pipetted per well of a 96-well streptavidin-coated plate (Thermo Fisher Scientific), in as many wells as required by the experiment. The plate was incubated for 1 h at 37°C to immobilise the DS-oligos to the plate. After incubation the DS oligo solution was discarded and the plate was vigorously tapped to remove residual liquid. The plate was then washed 3 times using 150 µl TBS-T per well. Qiagen blocking solution (0.1 g His-specific blocking reagent, 2 ml blocking buffer and 18 ml dd$H_2O$ stirred at 70°C for 10 min, after which 20 µl Tween was added) was then added (100 µl per well) and the plate was incubated for 30 min at room temperature. The blocking solution was discarded as described before with another 3 washes of TBS-T. Next 60 µl of the crude protein (diluted in DPI-buffer – DPI-Ex-buffer without glycerol and proteinase inhibitor – such that 3 samples of 0.3, 3 and 30 µg respectively per well could be used for each WRKY and the negative control). The plate was again incubated for 1 h at room temperature after which the protein extract was discarded and the plate was washed 3 times with 150 µl TBS-T. The antibody solution was then added (60 µl/well; Qiagen His-epitope antibody conjugated with horseradish peroxidase; 1:1000 in TBS-T) and the plate was left at room temperature for 1 h. Next the OPD solution was made by dissolving 1 OPD tablet from Sigma in 6 ml CP Buffer (12.5 ml 0.2 M $Na_2HPO_4$ and 125 ml 0.2 M citric acid, made up to 250 ml with dd$H_2O$ and adjusted to pH 5 with NaOH) with 3 µl 30% $H_2O_2$.

After discarding the antibody solution and washing 4 times with TBS-T (avoiding any bubbles), 60 µl per well of the OPD solution was added The plate was left at room temperature for up to an hour or until background signal was detectable. The colour reaction was stopped by adding 60 µl per well of 2 N HCl. Absorption readings were taken in an ELISA plate reader at 492 nm with a reference at 650 nm.

# Chapter 4: Results and Discussion

The present study focussed on a 14-member gene cluster on *A. thaliana* chromosome 1. Based on the results of a number of *in silico* and *in vitro* analyses, the original group of 14 was narrowed down to 13, then to 10, then 6, then 3, and then to a final interesting candidate, *At1g51850*,. The genes all code for LRR-RLKs and previous studies have shown that some of them could be responsive to biotic stressors (Chae *et al*., 2009; Gust *et al*., 2007; Kemmerling *et al*., 2007; Kemmerling *et al*., 2011; Postel *et al*., 2010). Here, the regulation of these genes upon PAMP treatment was evaluated using bioinformatic analyses and transient expression assays in protoplasts. The *cis*-elements and transcription factors potentially involved in the response of the genes to biotic stress were identified and characterised.

## 4.1. Structural and functional similarities within the gene cluster

A search for *A. thaliana* LRR-RLKs on PLAZA (url: http://bioinformatics.psb.ugent.be/plaza) yielded 14 closely-spaced genes on chromosome 1: *At1g51620, At1g51790, At1g51800, At1g51805, At1g51810, At1g51820, At1g51830, At1g51840, At1g51850, At1g51860, At1g51870, At1g51880, At1g51890, At1g51910.* The figure below shows the relative positions of the genes, except for *At1g51620*, which lies much further to the left of *At1g51790*.

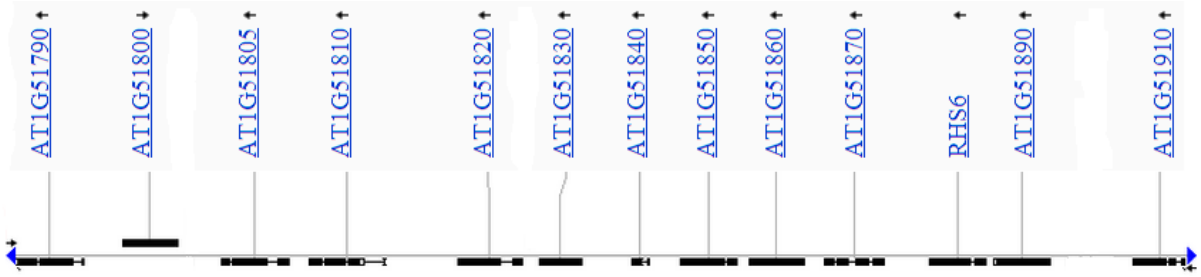



**Figure 4-1: The LRR-RLK gene cluster on chromosome 1 of *A. thaliana.***

This figure was created using the NCBI *Arabidopsis thaliana* Map viewer based on TAIR10 (url: http://www.ncbi.nlm.nih.gov/projects/mapview/). The search range was from position 19 206 858 to 19 288 385, the start and end positions of At1g51790 and At1g51910 respectively. At1g51620 is not shown on the diagram as it lies much further from the 13 shown here, at position 19 140 218 to 19 141 638. Root Hair Specific 6 (RSH6) is At1g51880.

The following figure shows a phylogenetic tree of the 14 genes, and again *At1g51620* seems to be an outlier, so it will no longer be considered part of the cluster for the present analysis. Gene duplication events may explain the clustering patterns observed (Thomas, 2006; Yi *et al*., 2007).




**Figure 4-2: Phylogenetic tree of the 14 genes within the cluster**

The tree was constructed using the basic input ("one click") features of *Phylogeny.fr: robust phylogenetic analysis for the non-specialist* (Dereeper *et* al., 2008), with the coding sequences (CDS) retrieved in Fasta format using the *Sequence* bulk data retrieval tool (url: http://www.arabidopsis.org/tools/bulk/sequences/index.jsp)

As seen in Table 4-1, to date only 2 of the genes in the cluster have been assigned specific functions related to defense responses according to TAIR annotations: *At1g51800* and *At1g51890*. The former plays a role in the plant's defense against downey mildew disease, while the latter is reportedly involved in the defense response against fungus and bacteria, as well as in the HR (Hok *et al,*. 2011; TAIR).

The question remains whether evolution of this cluster has resulted in functional redundancy or functional divergence.

**Table 4-1: Gene annotations according to TAIR**

| | ROLE IN IMMUNITY? | CELLULAR LOCAL-ISATION | PLANT STRUCTURE WHERE EXPRESSED | OTHER NAMES |
|---|---|---|---|---|
| ***At1g51790*** | respiratory burst involved in defense response | endo-membrane system/plasma membrane | cauline leaf, collective leaf structure, cotyledon, leaf apex, leaf lamina base, petiole, root, stem, vascular leaf | none |
| ***At1g51800*** | defense response to fungus, respiratory burst involved in defense response | plasma membrane | collective leaf structure, cotyledon, flower, hypocotyl, leaf apex, petal, root, sepal, stamen, stem, vascular leaf | IMPAIRED OOMYCETE SUSCEPTIBILITY 1, IOS1 |
| ***At1g51805*** | unknown | endomembrane system | carpel, cauline leaf, collective leaf structure, cotyledon, flower, guard cell, hypocotyl, inflorescence meristem, leaf apex, leaf lamina base, pedicel, petal, petiole, plant embryo, seed, sepal, shoot apex, shoot system, stamen, stem, vascular leaf | none |
| ***At1g51810*** | unknown | plasma membrane | unknown | none |
| ***At1g51820*** | respiratory burst involved in defense response | endo-membrane system | collective leaf structure, flower, inflorescence meristem, leaf apex, petiole,plant embryo, sepal, stamen, stem, vascular leaf | none |
| ***At1g51830*** | unknown | nucleus | root | none |
| ***At1g51840*** | unknown | extracellular region/endomembrane system | unknown | none |
| ***At1g51850*** | unknown | plasma membrane, plasmodesmata | guard cell, root | none |
| ***At1g51860*** | unknown | endomembrane system | guard cell, root, stamen, stem | none |
| ***At1g51870*** | Unknown | endomembrane system/extracellular region | Unknown | none |
| ***At1g51880*** | unknown | endomembrane system/extracellular region | only in root hair cell | ROOT HAIR SPECIFIC 6, RH6 |
| ***At1g51890*** | defense response to bacterium, defense response to fungus, regulation of plant-type hypersensitive response, salicylic acid mediated signalling pathway | endomembrane system/extracellular region | cauline leaf, collective leaf structure, cotyledon, guard cell, hypocotyl, root, sepal, stamen, stem, vascular leaf | none |
| ***At1g51910*** | unknown | endomembrane system/extracellular region | carpel, cauline leaf, collective leaf structure, cotyledon, flower, inflorescence meristem, leaf apex, leaf lamina base, petal, petiole, plant embryo, root, seed, shoot system, stamen, stem, vascular leaf | none |

It is also interesting that *At1g51880* is only expressed in root hairs, though its exact function is unknown. Future studies on the cluster could focus in detail on the expression patterns of the genes in specific tissues in response to other biotic and abiotic stimuli. Such investigations would shed light on the evolutionary history of the cluster, in terms of environmental pressures. *At1g51805* is annotated as a pseudogene according to NCBI, and it would be interesting to find out whether any of the other genes in the cluster started out as redundant copies, only to develop novel functions later on through mutations (Conant & Wolfe, 2008).

## 4.2. The known effects of biotic stress on the gene cluster of interest

The experimental part of this study was conceived after examining microarray results provided by the Nürnberger lab at the ZMBP. This AtGenExpress Pathogen Series data is publically available and can be accessed on the Arabidopsis electronic Fluorescent Pictograph (eFP) Browser (url: http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) as part of the Stress Series (Winter *et al*., 2007).

Figure 4-3 shows a summary of the absolute expression values in response to 1 mM flg22 as per the eFP browser. At the time of this particular analysis, only 10 genes were considered part of this cluster. The graph seems to indicate that of the 10 genes in the cluster, 6 are upregulated to different degrees upon the addition of the biotic stress elicitor flg22. These 6 genes are *At1g51790, At1g51800, At1g51820, At1g51850, At1g51860* (only slightly upregulated) and *At1g51890,* and it was decided to limit the rest of the investigation to only these genes.

Other biotic stressors tested during the microarray study included elicitors HrpZ, LPS and oomycyte-derived GST-NPP1, and *Botrytis cinerea*, *Pseudomonas syringae, Phytophthora infestans* and *Erysiphe orontii.* For the present study however, only flg22 was used as it was the best defined and studied (Kemmerling lab, ZMBP).

It is worth noting how the microarray data was collected and represented since it formed the basis of this dissertation. According to eFP Browser, 5 week old *Arabidopsis thaliana* Columbia-0 leaves were grown under 8/16 hour light/dark conditions and were infiltrated in triplicate with 1 µM flg22. The data shown on the eFP Browser represents the average of the triplicate values, and results can be filtered to account for high standard deviations. Isolated RNA was hybridised to the ATH1 GeneChip (Affymetrix) and results were normalised by global scaling via the Gene Chip Operating System (GCOS). The normalisation procedure involves averaging data from independent replicates and then scaling

the “Treatment” sample data using the “Control” sample as the reference. Regression analysis is used to determine the scaling factors (Hunter *et al*., 2002).

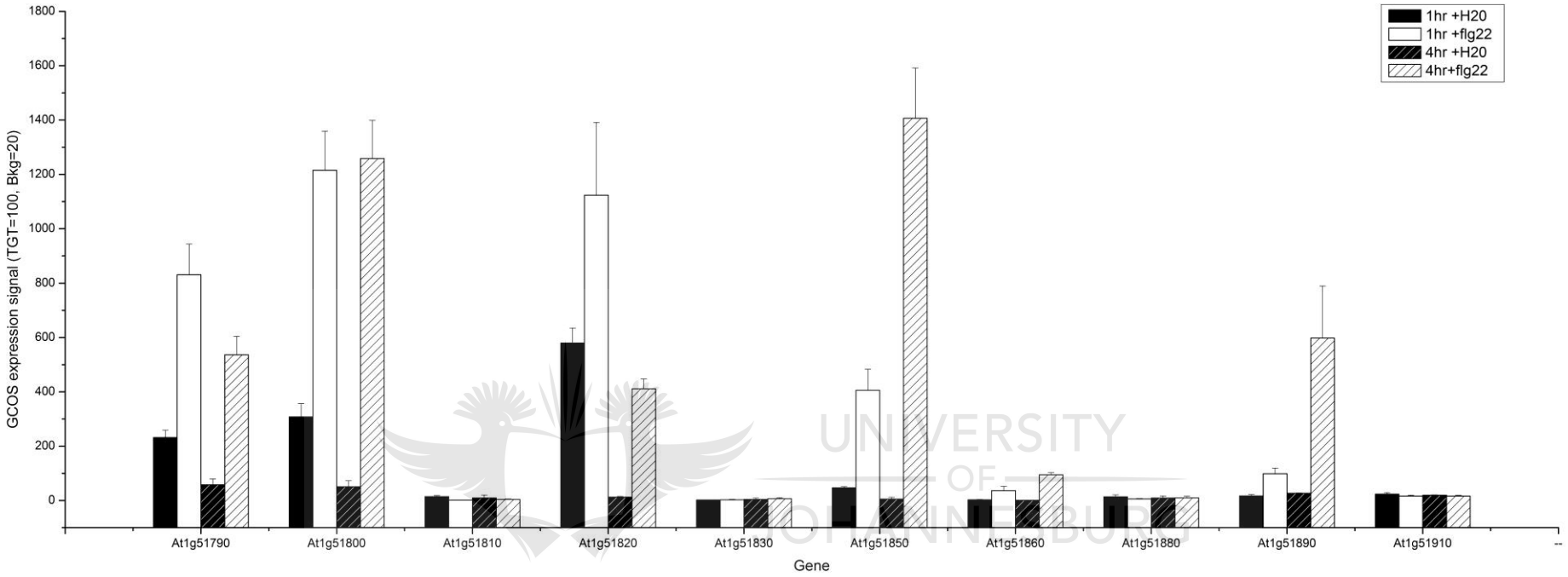



**Figure 4-3: Response of the gene cluster to flg22 as per microarray gene expression data sourced from eFP browser**

This graph summarises the absolute values of expression of the gene cluster according to available data on the BAR Arabidopsis eFP browser. Most genes have a basal expression signal between 0 and 100 (represented on the y-axis). Genes At1g51790, At1g51800, At1g51820, At1g51850, At1g51860 and At1g51890 show differences in expression levels between time points and treatment with flg22. The other four genes, At1g51810, At1g51830, At1g51880 and At1g51910 appear to be unaffected by flg22 treatment. The error bars indicate standard deviation.

It is possible to represent the data either as absolute values, or as relative values. Absolute values refer to the "raw" normalised expression values, whereas the relative values are so in relation to the mock treatment of 10mM $MgCl_2$. This study was originally based on the relative values which give an indication of "fold induction"; that is, how much larger the expression of the treated samples are in comparison to the mock treated samples. What this representation omits is the fact that induced expression levels are exaggerated when starting with a low basal expression. For example, a basal expression of 0.1 units that is induced to a level of 0.5 seems to have a 5 fold higher expression; this is however insignificant when considering that most genes have a basal expression of between 0 and 100. In other words, graphing relative expression values may exaggerate induction by stressors.

Nevertheless, the microarray data hints at interesting gene expression in response to stressors, thus informing the work presented here.

## 4.3. Measuring flg22 induction of At1g51790, Atg151800, At1g51820, At1g51850, Atg151860 and At1g51890 *in vitro*



The promoters of the 6 genes that were shown to be induced by flg22 in the microarray study (see 4.2 and Figure 4-3) were cloned into pBGWL7 vectors to drive the transient expression of luciferase in *A. thaliana* mesophyll protoplasts (Karimi *et al*., 2007). Using luminometry to measure the expression in relative light units (RLUs), the full length promoters (see Table 3-3) were assessed as to their *in vitro* response to flg22. (Attempts were made to develop a similar assay using green fluorescent protein (GFP), but this approach failed as the sensitivity of the assay was not high enough to detect basal gene expression with a fluorescent microscope).

Based on pooled results from all experiments, Figure 4-4 shows the fold induction of the 6 promoters using the luciferase-based assay. From these results, three promoters were selected for further study: At1g51790prom, At1g51850prom and At1g51890prom. At1g51850 and At1g51890 showed fairly consistent inducibility, while the At1g51790prom showed inconsistent induction – this is clear from error bars indicated in Figure 4-4.

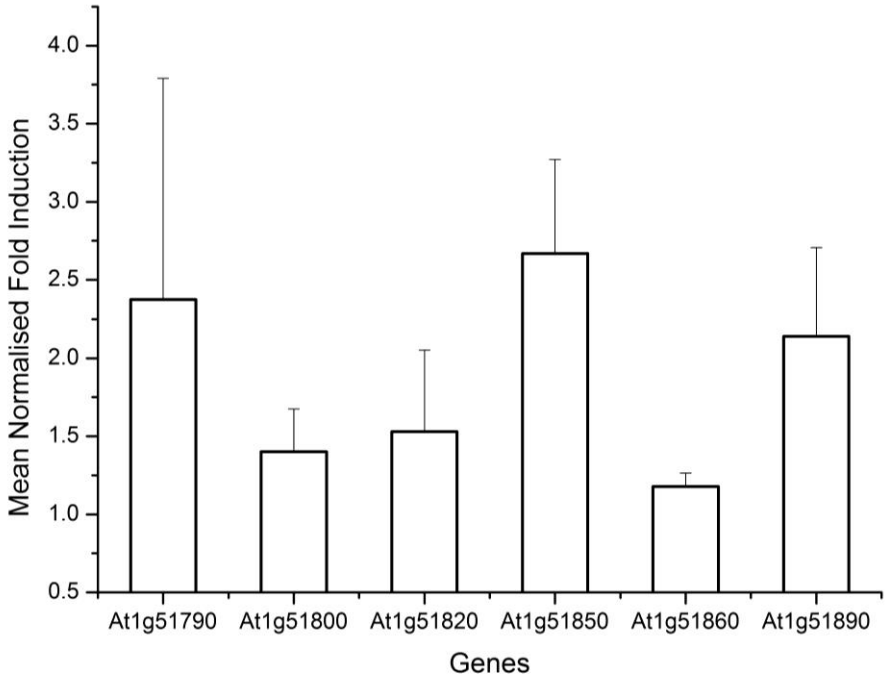



**Figure 4-4: The average normalised flg22 inducibility of the full length promoters of At1g51790, Atg151800, At1g51820, At1g51850, Atg151860 and At1g51890.**

This graph was created using the mean maximum fold induction relative to the water controls for the 4 h time point of all experiments in which the genes were tested themselves or used as controls to test mutants. The promoters were cloned into pBGWL7, which was transformed into *A. thaliana* mesophyll protoplasts. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water. The error bars indicate standard deviation.

At this point it serves well to note that the transient expression assays performed during this study gave highly variable results. Basal expression levels seemed to have been affected by transformation efficiency, environmental stress and protoplast viability. The latter was aggravated by the growth conditions of the plant, including temperature, humidity and occasional *Thysanoptera* (thrips) infections, which were detected too late. Indeed, it is well-known that plant defense responses are sensitive to both biotic and abiotic stressors, including circadian disruptions and developmental factors (Schwessinger & Ronald, 2012). It is for this reason that pooled results had to be normalised as a fold induction rather than using the raw expression values in RLUs.

The same criticism as mentioned before stands with the representation of fold induction: results are skewed depending on the values of basal expression/untreated controls. However, a subjective, visual comparison between the raw graphs and this single normalised graph did not seem to indicate a compromise in individual induction trends.

Still, it is an unfortunate consequence of assay variability then that the microarray data cannot be directly compared to the *in vitro* data. Other factors, such as a different flg22 concentration, may account for discrepancies in the expression levels seen in microarray data versus the *in vitro* data. Furthermore, while the microarray data used only two time points (1 h and 4 h), the protoplast assays were done as a kinetic study over 6 h. The maximal expression did however mostly occur at 4 h so this was the time point used for the normalised graph. When drawing indirect comparisons between Figure 4-3 and Figure 4-4, it must be remembered that the microarray detects mRNA levels whereas the transient expression assays detect protein levels – there is thus an inherent delay in the conversion of mRNA to protein. In addition, the microarray studies are not limited by the somewhat arbitrarily truncated "full length" promoter in the way that the *in vitro* protoplast assays might be.

Earlier it was pointed out that At1g51850 is located in the plasma membrane, like FLS2, but unlike FLS2 it is not ubiquitously expressed in every tissue. In fact, TAIR asserts that it is majorly expressed in the root of the plant (see Table 4-1: Gene annotations according to TAIRTable 4-1). In the case of all genes tested here, the assay relies on leaf tissue and thus clearly ignores tissue-specific basal expression variability. Furthermore, available evidence indicates that PAMP perception and receptor kinase signalling could be tissue-specific as well: flg22, PGN and chitin responses were shown to differ between distinct root zones (Tena *et al*., 2011).



All these factors may have contributed to the fact that At1g51800prom, At1g51820prom and At1g51860prom hardly showed a response in terms of basal or induced expression. Since no difference in expression could be detected for those three promoter constructs, it was decided not to continue with further studies on them, but rather to focus on At1g51790prom, At1g51850prom and At1g51890prom.

## 4.4. Over-represented *cis*-elements in the promoters of At1g51790, At1g51850 and At1g51890

The cloned promoter sequences of *At1g51790, At1g51850* and *At1g51890* were analysed using Regulatory Sequence Analysis Tools (RSAT, url: rsat.ulb.ac.be) in order to find over-represented hexamers that might act as *cis*-elements. Figure 4-5 shows the approximate locations of putative W-boxes – the only over-represented hexamer that RSAT was able to find in the sequences. It also indicates the restriction sites used to create deletion mutants, which were used to eliminate putative non-functional W-boxes before examining them individually.

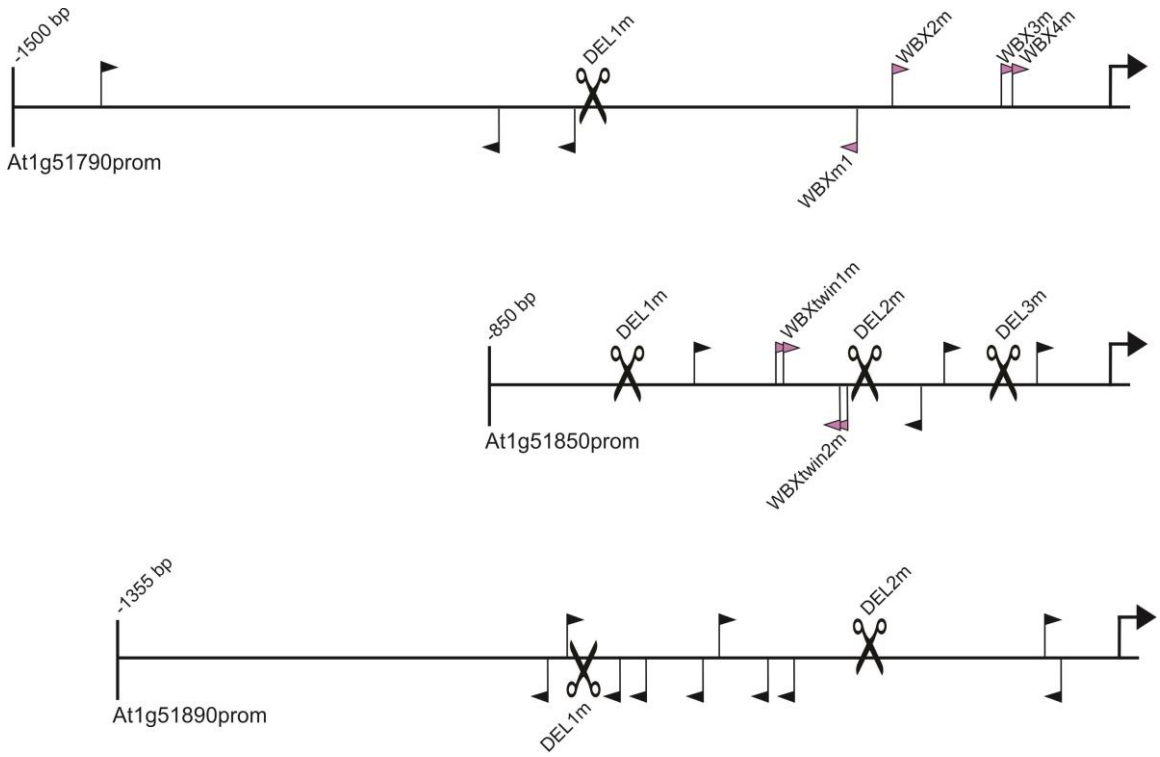



**Figure 4-5: Graphical representation of the promoters indicating W-Boxes and deletion target sites.**

W-boxes are indicated as directional flags, and the restriction sites used to create the deletion mutants are indicated as scissors. The cloned promoter of *At1g51790* was 1500bp in length and those of *At1g51850* and *At1g51890* were 850 and 1335 bp, respectively. The exact promoter sequences and locations of the W-boxes and relevant restriction sites are available in Appendix I and in Chapter 3, Table 3-3.

## 4.5. The effect of truncations and/or mutations on the promoters of At1g51790, At1g51850 and At1g51890

The promoters were systematically truncated using restriction enzymes on the pBGWL7 vectors in which the promoters were cloned (see 3.4.8). The resulting constructs were then religated and compared to the full length version in protoplast assays as described before.

### 4.5.1. Induction patterns of the modified At1g51790 promoter

Only one deletion mutant (DELm1::Luc) was tested for At1g51790prom, truncated at -712 bp upstream of the start codon. Figure 4-6 shows a kinetic that compares the response to flg22 of the mutant promoter to that of the full length promoter, graphed from a representative experiment.

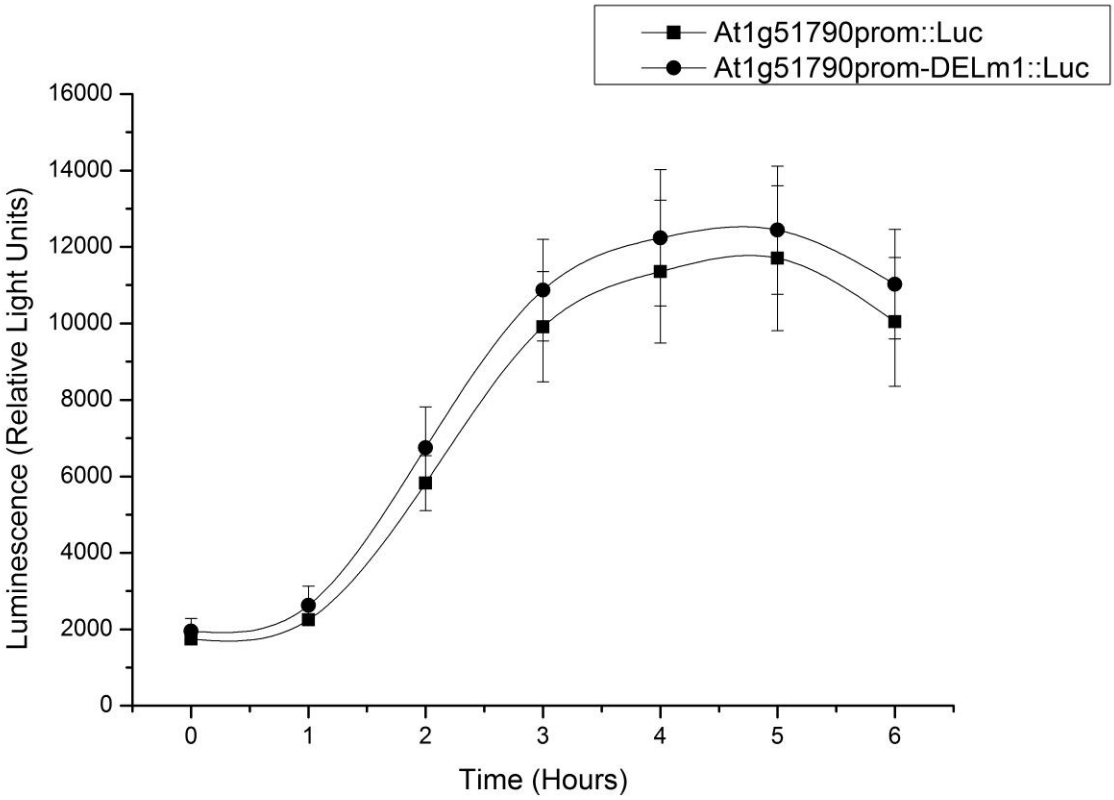



**Figure 4-6: A representative example of the relative Flg22 inducibility of At1g51790prom: The full length promoter (~-1500 bp) compared to the truncated (-712 bp) promoter.**

The promoters were cloned into pBGWL7, which was transformed into *A. thaliana* mesophyll protoplasts. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.



It is clear that there is no significant difference in induction between the deletion mutant and the full length promoter. The mutant is missing three W-boxes according to Figure 4-5, and the above result thus suggests that these 3 W-boxes are not responsible for the inducible response of At1g51790prom to flg22.

It was hence decided to inactivate each of the remaining four W-boxes located ahead of the deletion site by individual mutation. This approach is reported in the literature. For example, one group investigated the W-boxes within the *PR-1* (pathogenesis-related 1) promoter. They found, using luciferase-based transient *A. thaliana* mesophyll protoplast assays, that the deletion of W-box sequences leads to a considerable reduction in reporter gene activity (Pape *et al*., 2010).

The mutations were done on the full-length promoter. Figure 4-7 shows how these mutants responded to flg22 as compared to the unmutated promoter.

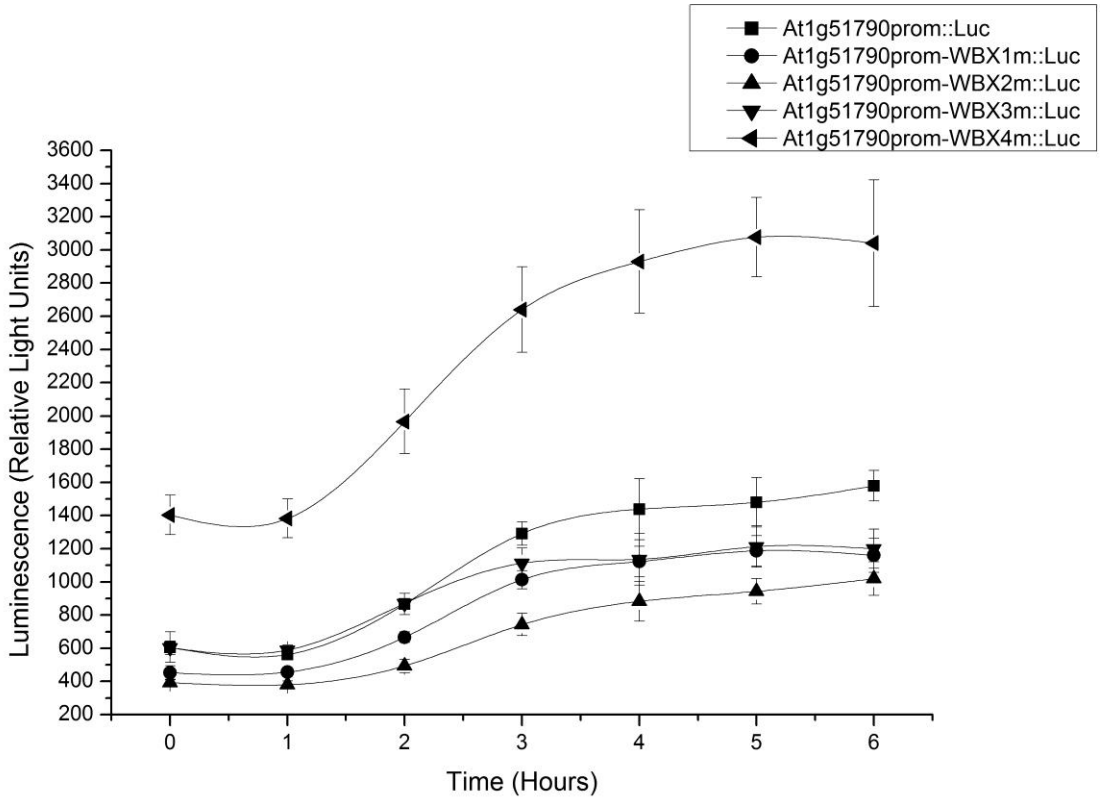



**Figure 4-7: A representative example of the relative Flg22 inducibility of At1g51790prom: The full length promoter (~-1500 bp) versus four individual W-box mutants (refer to Fig.4-5).**

The promoters were cloned into pBGWL7, which was transformed into *A. thaliana* mesophyll protoplasts. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.

W-box mutants 1 to 3 showed slightly lower flg22 inducibility than the unmutated promoter. This suggests that they may play a role in the stress response of this gene.

Of more interest is the fact that the 4$^{th}$ W-box mutant showed slightly greater flg22 inducibility than the unmutated promoter, even when one considers that its basal response is higher. It thus appears that inactivation of this W-box, the one closest to the transcription start site, increases the promoter's stress response in the case of flg22, but the reasons for a higher basal response remain unclear.

One implication may be that this W-box is not responsible for flg22 inducibility, but that it may be involved in transcriptional repression, which plays a crucial role in balancing stress responses with normal cellular functions (see 4.6).

Future investigation could focus on double, triple and quadruple mutants to begin to shed light on whether the W-boxes act synergistically or individually as repressors/enhancers of transcription, as has been done for 3 closely-spaced W-boxes in the promoter of the *Arabidopsis* pathogenesis-related-1 (PR-1) gene. In that case, the simultaneous deletion of all 3 W-boxes had the same effect as deleting the W-box furthest away from the start codon between positions −550 and −480 (Pape *et* al., 2010). For the promoter of At1g51790, it might be that the inactivation of a single W-box, only slightly reduces the flg22 response, which might indicate that it doesn't matter which are active, as long as one or two out of the three are. It would be interesting to find out how exactly the response to flg22 is modulated if W-box 4 is in fact a binding site for a repressor and W-boxes 1 – 3 are binding sites for enhancers. Research has indeed shown that interacting WRKY TFs or WRKY complexes may bind to closely-spaced W-box clusters, like those present in the promoters discussed here, to regulate gene expression cooperatively or antagonistically (Chi *et al*., 2013).

### 4.5.2. Induction patterns of the modified At1g51850 promoter

Three deletion mutants were tested for At1g51850prom, namely DELm1, DELm2 and DELm3, truncated at -630 bp, -368 bp and -134 bp upstream of the start codon, respectively. DELm1 did not show a difference in expression compared to the full length promoter (data not shown). Figure 4-8 shows results from a kinetic study that compares the flg22 response of the other two mutant promoters, DELm2 and DELm3, and the full length promoter, graphed from a representative experiment.

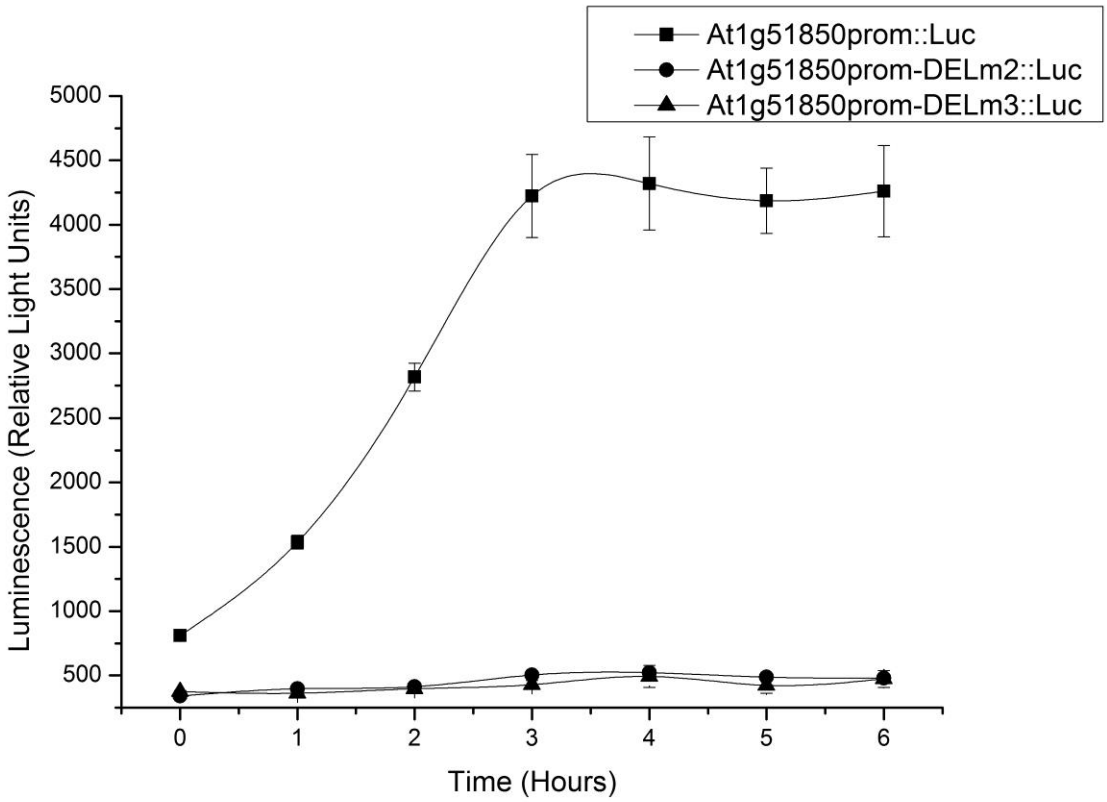



**Figure 4-8: A representative example of the relative Flg22 inducibility of At1g51850prom: The full length promoter (-850 bp) versus the second (-368 bp) and third (-134 bp) truncated promoters (Refer to Fig. 4-5).**

The promoters were cloned into pBGWL7, which was transformed into *A. thaliana* mesophyll protoplasts. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.

Clearly, flg22 inducibility is no longer a feature of the promoter truncated at -368 bp. The further truncation at around -134 bp is therefore a redundant reflection of the same.

This data suggests that the transcription factor binding sites responsible for the upregulation of the genes in response to flg22 are located between deletion sites 1 and 2. As shown in Figure 4-5, there are 5 putative W-boxes in this region. The four closest to the second deletion site represent two W-box dyads, a functional configuration that has been found in a number of promoters: In parsley for example, target dyads of *Pc*WRKY1 were found to have a synergistic effect on transcription (Eulgem *et al*., 1999), and in barley, efficient binding of *Hv*WRKY38 necessitates two proximal W-box motifs (Mare *et al*., 2004). In the *Arabidopsis* genome, W-box dimers spaced 0 to 30 nucleotides apart are statistically abundant in promoter regions (Ciolkowski *et al*., 2008).

On speculation that the arrangement of the W-box dyads in the promoter of At1g51850 may have functional significance as well, it was decided to inactivate each of the dyads individually and as a unit by mutation. This yielded the constructs named by the suffixes -WBXtwin1m and -WBXtwin2m, and -WBXtwin1x2m respectively. Figure 4-9 shows the resulting flg22 inducibility.

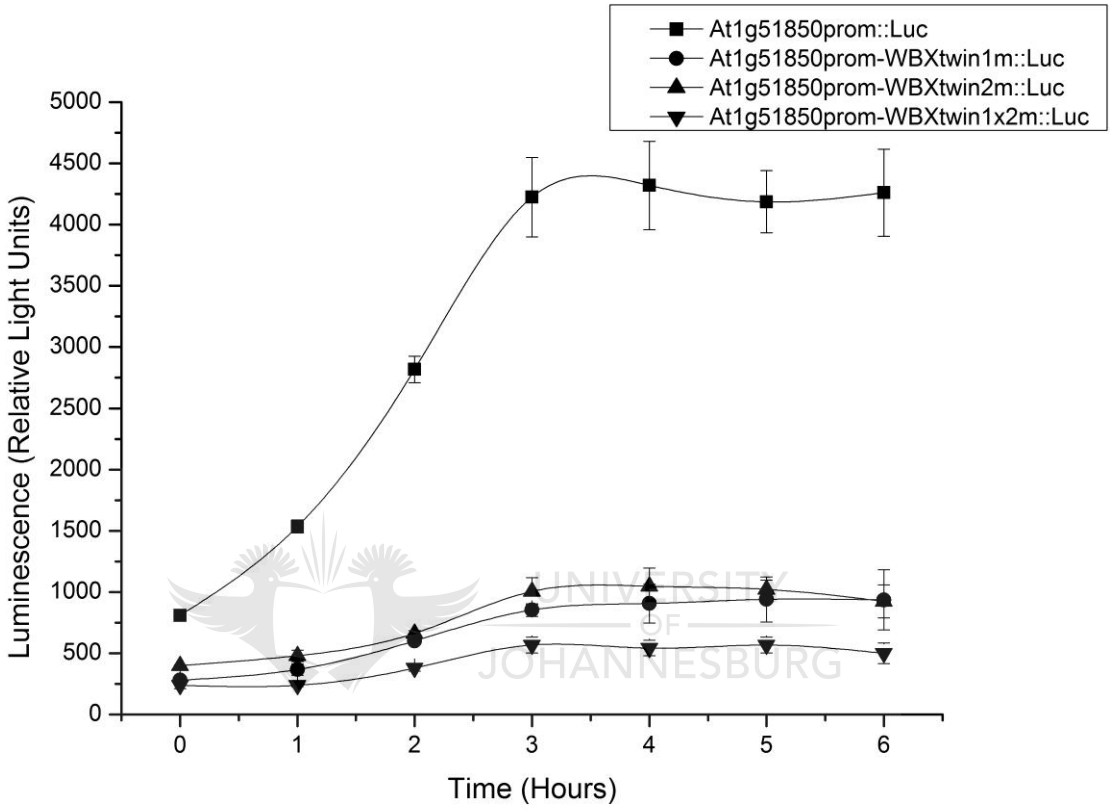



**Figure 4-9: A representative example of the relative Flg22 inducibility of At1g51850prom: The full length promoter (-850 bp) versus the W-box dyad mutants (refer to Fig. 4-5).**

The promoters were cloned into pBGWL7, which was transformed into *A. thaliana* mesophyll protoplasts. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.

Individual mutation of the dyads results in a markedly lower response to flg22, and it is lower still upon mutation of both dyads. This provides strong *in vitro* evidence that these W-Box dyads are in fact important for the induction of At1g51850 by flg22. It does however not exclude the possibility that other *cis*-acting elements may be involved.

## 4.5.3. Induction patterns of the modified At1g51890 promoter

The deletion mutant DELm1, truncated at around -700 bp, did not show a clear difference in expression compared to the full length promoter (data not shown). Figure 4-10 shows results from a kinetic study that compares the flg22 response of the promoter truncated at -334 bp (-DELm2) to its full length counterpart, At1g51890prom (-1335 bp), graphed from a representative experiment.

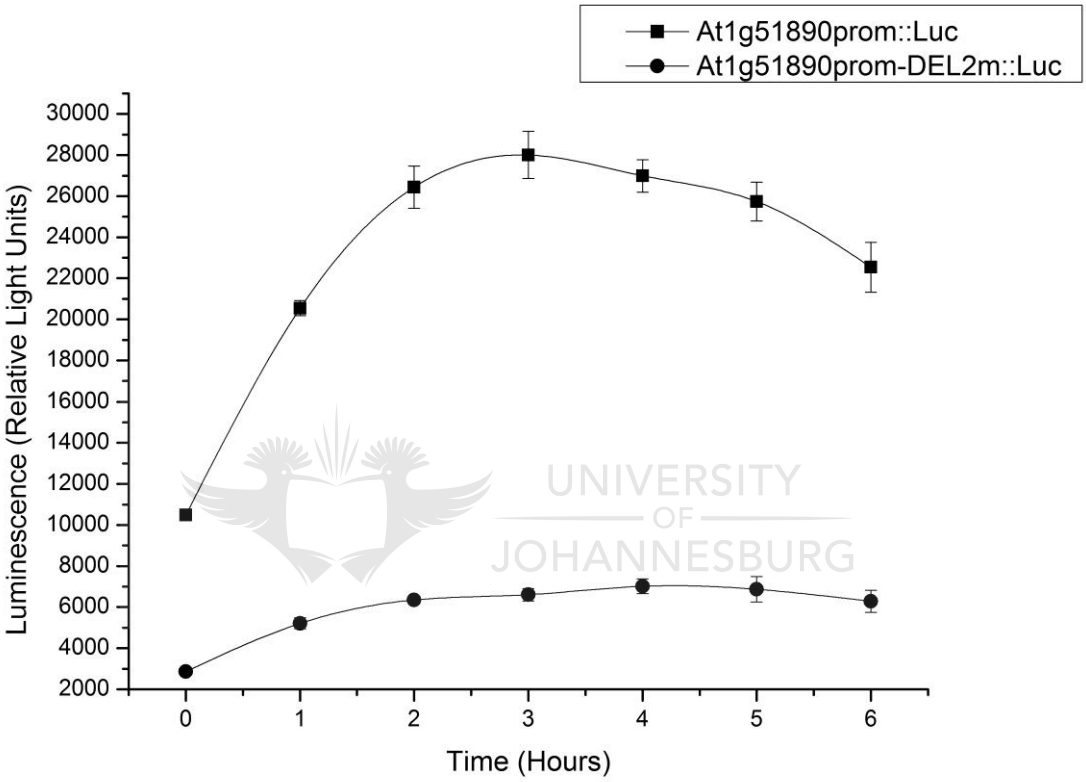



**Figure 4-10: A representative example of the relative Flg22 inducibility of At1g51890prom: The full length promoter (-1335 bp) versus the second truncated promoter (-334 bp). (Refer to Fig. 4-5)**

The promoters were cloned into pBGWL7, which was transformed into *A. thaliana* mesophyll protoplasts. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.

It is clear that flg22 inducibility is prominently reduced in the DELm2 promoter, truncated at around 300 bp. There are still 2 W-boxes between the truncation site and the start codon so this might account for the slight induction observed in the truncated promoter. This data suggests that the 6 W-boxes located between 850 and 1335 bp (between the deletion sites; see Figure 4-5) may play a role in the

transcriptional upregulation of At1g51890 in response to flg22. However, it is once again important to differentiate between basal expression and fold induction. For At1g5890 in particular, the removal of the W-boxes also resulted in a lower basal expression. Further investigations would reveal whether WRKY TFs are in fact responsible here, or whether it was merely a side effect or stress response related to the preparation of these mutants.

In summary then, truncations that deleted W-boxes resulted in markedly lower responses to flg22 in At1g51790prom, At1g51850prom and At1g51890prom. Mutational inactivation of individual W-boxes in At1g51790prom also drastically reduced the flg22 response and mutational inactivation of W-box dyads in At1g51850 resulted in virtually no flg22 inducibility. The deletion of 6 W-boxes in the promoter of At1g51890, done via truncation, drastically reduced both its basal expression and its inducible response to flg22. These results provide strong *in vitro* evidence that W-boxes may play a key role in the transcriptional regulation of these genes in response to flg22.

## 4.6. WRKY transcription factor genes co-expressed with the induced RLK genes



GeneSpring GX version 7.3 from Agilent was used to identify WRKY7, -11, -22 and -26 as having expression patterns similar to those of At1g51790, At1g51850 and At1g51890 (Wanke Lab, ZMBP; see 3.3).

WRKY7 and -11 are closely related and are known to suppress defense responses, but it is not known whether they do so directly or indirectly (Ciolkowski *et al*., 2008; Eulgem & Somssich, 2007). WRKY11 loss-of-function mutants (though, interestingly, not WRKY7 loss-of-function mutants) have been shown to enhance basal resistance and WRKY11 has also been implicated in SA-JA crosstalk (Pieterse *et al*., 2012; Journot-Catalino *et al*., 2006).

WRKY26 is involved in the regulation of salt and heat stress tolerance (Van Aken *et al*., 2013). It is also induced by bacterial pathogens and/or SA, and transgenic plants constitutively overexpressing WRKY26 were found to be more susceptible to bacterial pathogens (Zheng, 2005). Interestingly, unlike most other WRKY TFs, WRKY26 can simultaneously bind to both W-box motifs of closely spaced dyads (Ciolkowski *et al*., 2008).

Although it might seem counterintuitive that negative regulators of resistance would be induced during pathogen attack, they are necessary to maintain a balance between the allocation of resources

to defense responses and to other general functions of the plant such as growth and reproduction. As such, these repressive TFs may be important for attenuating the defense response, and they may also ensure an equilibrated response against a range of pathogens rather than just one class of attacker (Journot-Catalino *et al*., 2006).

WRKY22, or At4g01250, has been implicated in specific defense responses as well. It was shown to be upregulated by chitooctaose, a chitin octamer (Libault *et al*., 2007) and functions downstream of FLS2 upon flg22 perception through a MAPK cascade (Asai *et al*., 2002).

## 4.7. The effect of WRKY over-expression on the flg22 inducibility of At1g51850

Those WRKY TFs determined to have a similar expression pattern to the gene cluster in question were cloned into expression vectors (see 3.4.5) to determine if their over-expression had any effect on the flg22-inducibility of the genes. Following on from the previous experiments that revealed At1g51850 to be the most interesting gene in terms of its response to flg22 and its promoter's W-box dyads, the luciferase-driving promoter of At1g51850 was tested in the presence and absence of the over-expressed WRKY TFs using the protoplast transient expression assay.



### 4.7.1. The cellular localisation of over-expressed WRKY TFs

The figure below shows GFP-conjugated WRKY11 is localised in the nucleus of the plant protoplast. It is assumed that the same localisation takes place in the luciferase assays (results to follow) where untagged WRKY TFs were used to ensure functionality. Confirmation that the cloned protein is localised in the nucleus indicates that it is in the right area to act as a TF on the cell's DNA.

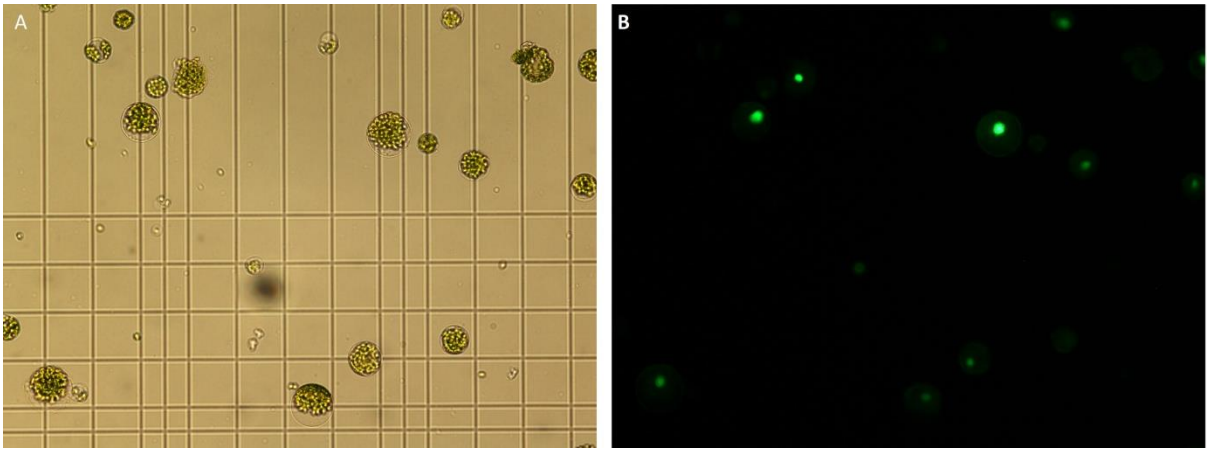



**Figure 4-11: Localisation of over-expressed GFP-tagged WRKY11 in protoplasts.**

A light microscope image of the protoplast transformed with WRKY11 cloned into pK7WGF2 is shown on the left (A). The fluorescence microscope image of exactly the same frame is shown on the right (B). The GFP fluorescence clearly originates from the nuclei of the protoplasts.

### 4.7.2. The effect of over-expression of WRKY7, -11, -22 and -26 on the flg22 inducibility of At1g51850



As mentioned, WRKY7, -11, -22 and -26 were overexpressed to determine if they might interact with the W-box dyads on the promoter of At1g51850. Figure 4-12 shows how the basal expression and the flg22 inducibility of full length At1g51850prom are affected by the presence of WRKY7, -11, -22 and -26.

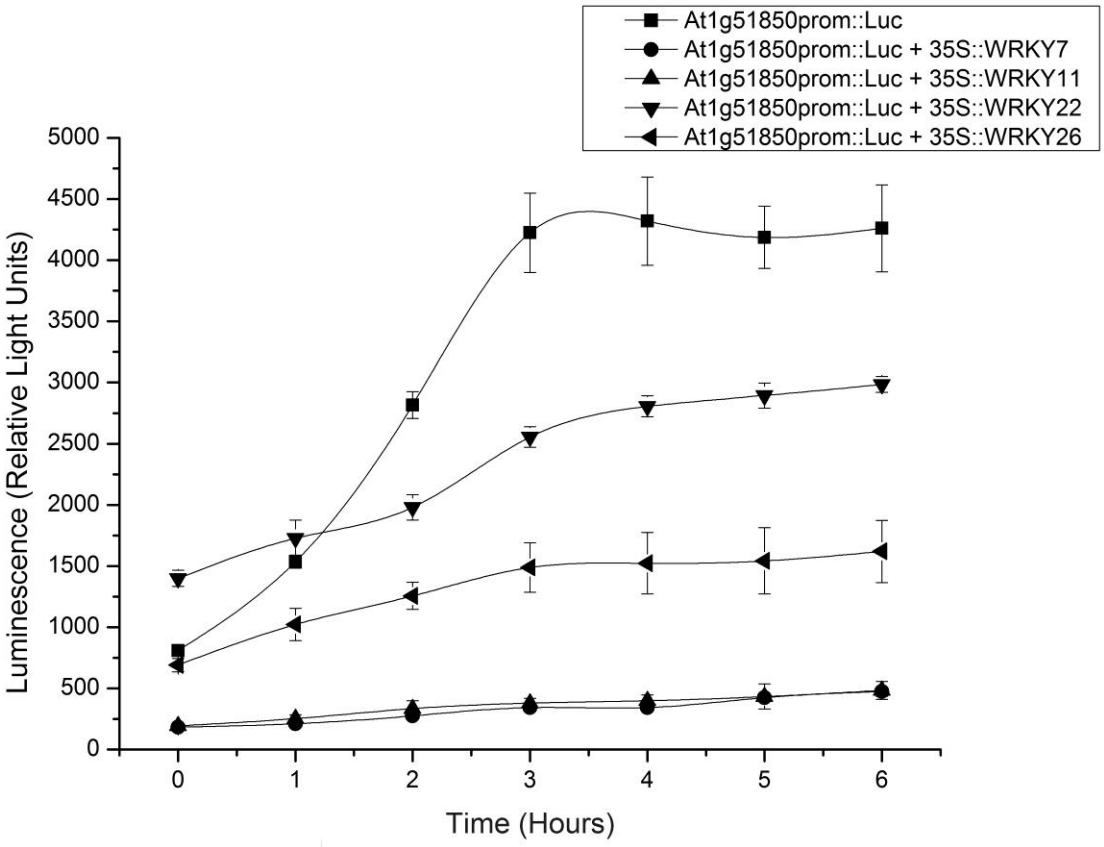



**Figure 4-12: A representative example of the relative Flg22 inducibility of the At1g51850prom: the full length promoter only and in the presence of over-expressed WRKY7, -11, -22 or -26.**

The promoters were cloned into pBGWL7 and the WRKY genes were cloned into pK7FWG2. The vectors were transformed into *A. thaliana* mesophyll protoplasts as required by the experiment. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.

As expected, known transcriptional repressors WRKY7 and -11 minimise basal expression as well as flg22 inducibility to negligible levels. WRKY26 appears not to affect basal levels but does suppress inducibility to some extent. Interestingly, WRK22, the known transcriptional enhancer (see 4.6), increases basal transcription but seems to suppress flg22 induction; because WRKY22 has been shown to mimic flg22 signalling (Asai *et al*., 2002), an artificial excess of the protein in these experiments was expected to result in much higher inducibility.

Of the selected WRKY TFs, only WRKY22 is known to enhance defense responses (see 4.6), and only WRKY22 showed an interesting effect on the expression kinetics of the At1g51850prom W-box dyad mutants upon over-expression. The next two figures show that WRKY22 enhances the basal expression of the second W-Box mutant to the same level as the unmutated promoter.

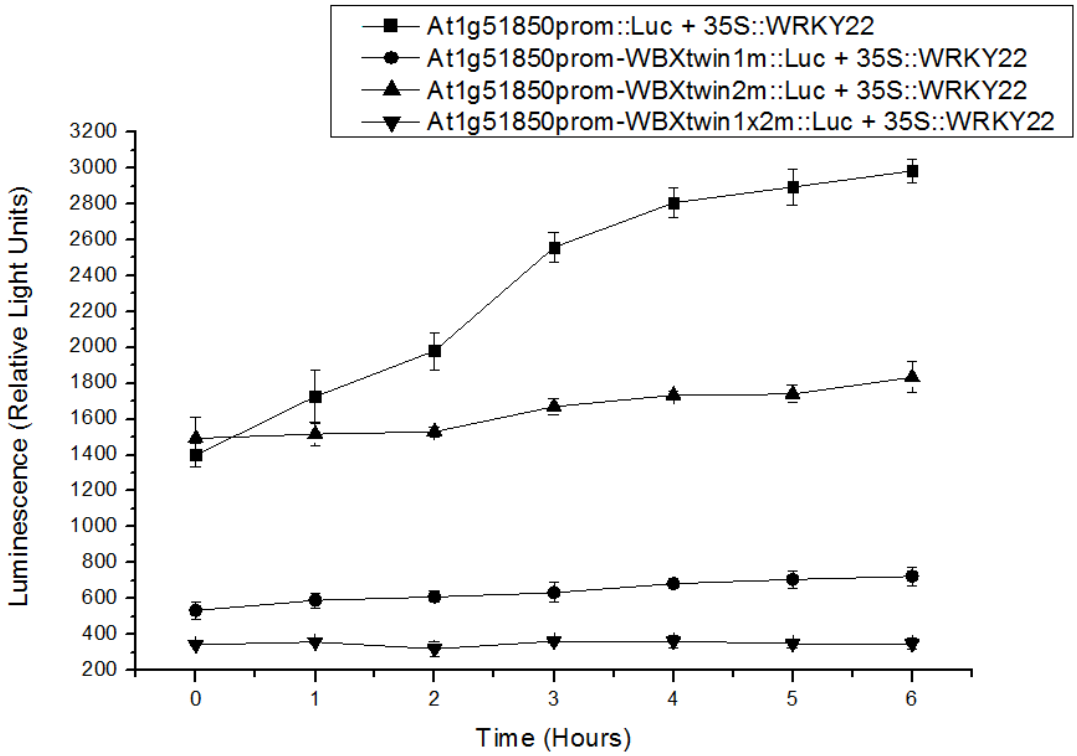



**Figure 4-13: A representative example of the relative Flg22 inducibility of At1g51850prom in the presence of over-expressed WRKY22 – the full length promoter versus the W-box dyad mutants.**

The promoters were cloned into pBGWL7 and the WRKY genes were cloned into pK7FWG2. The vectors were transformed into *A. thaliana* mesophyll protoplasts as required by the experiment. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.

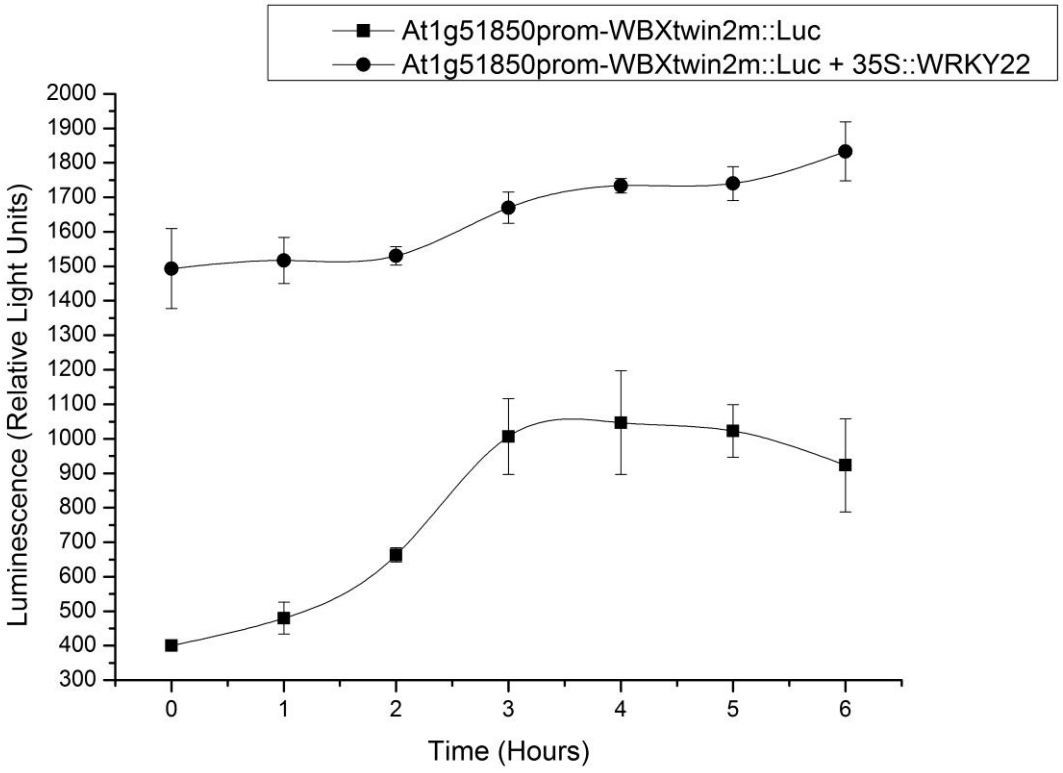



**Figure 4-14: A representative example of the relative Flg22 inducibility of At1g51850prom-WBXtwin2m – the second W-box dyad mutant in the presence and absence of over-expressed WRKY22.**

The promoters were cloned into pBGWL7 and the WRKY genes were cloned into pK7FWG2. The vectors were transformed into *A. thaliana* mesophyll protoplasts as required by the experiment. The protoplasts, 100 µl samples in triplicate, were subjected to induction with either 1 µl of 10 µM flg22 or 1 µl of distilled water for 6 hours, with luminometric readings taken every hour.

Taken together, the experiments seem to show the following with regard to WRKY22 and At1g51850prom: First, WRKY22 enhances the basal expression to the same degree in both At1g51850prom and At1g51850prom-WBXtwin2m; second, WRKY22 suppresses flg22 inducibility in At1g51850prom and has no effect on that of At1g51850prom-WBXtwin2m; and finally, WRKY22 has no real effect on the expression or inducibility of At1g51850prom-WBXtwin1m or At1g51850prom-WBXtwin1x2m .

What this suggests is that the first W-Box dyad interacts with WRKY22 to enhance the basal expression level, since this enhancement is not observed in the mutant. It is also clear that the second dyad cannot be responsible for this because the enhancement is still observable when it is mutated.

One thing to note however, is that it is not possible to use the mutant experiments to attribute the suppression of flg22 inducibility directly to the W-box dyads because their inactivation already results in this suppression.

Though these results are a promising starting point for investigating the role of these or other WRKYs in regulating the expression of Atg51850 via its promoter's W-box dyads, certain issues cannot be addressed in this artificial experimental set up. For example, the natural expression levels of individual WRKYs may be different depending on the developmental stage of the plant, and some WRKYs, including WRKY14, -25, -36, -69 and-72, are expressed almost exclusively in roots; here only leaf tissue was used from 4 week old plants (Dong *et al*., 2003). In other words, if interaction could be shown here, it may not be a reflection of natural circumstances. A more comprehensive approach could be to generate knock-out or knock-down WRKY lines, and to then observe the effects on the expression of all the genes in the cluster, with and without pathogen elicitation (Nakayama *et al*., 2013), in different areas of the plants and at different stages of development.

That being said, others have used transient expression assays and W-box mutations to confirm transcriptional activation by specific WRKYs. One group transformed plasmids containing WRKY45 cDNA driven by the cauliflower mosaic virus 35S promoter into rice coleoptiles to examine the effect on the luciferase-driving promoter of *OsNAC4*, a transcription factor that acts as a positive regulator of hypersensitive cell death (Kaneda *et al*., 2009). In this case, the WRKY44 effector gene clearly activated the reporter gene, and subsequent mutations of the promoter's W-boxes markedly reduced transactivation (Nakayama *et al*., 2013).



Another aspect to investigate would be the fact that multiple WRKYs may be able to interact with a W-box on a promoter, though each WRKY may have its own preferential binding domain. This was shown for WRKY18, WRKY40 and WRKY60, which interact with the W-box in the promoters of *ABI4* and *ABI5*, ABA-responsive genes. In that case researchers proposed that the WRKYs may cooperate to regulate the repression of the genes (Liu *et al*., 2012).

## 4.8. Demonstrating an *in vitro* interaction between WRKY7, -11, -22 and -26 and the W-box dyads of the promoter of At1g51850

In order to determine whether WRKY7, -11, -22 and/or -26 are able to physically interact with the W-box dyads in the promoter of At1g51850, the proteins were expressed with His-tags for use in a DNA-protein interaction enzyme-linked immunosorbent assay (DPI-ELISA; see 3.6).

### 4.8.1. WRKY protein expression

All four tagged WRKY TFs were successfully cloned, expressed, and purified by means of a crude protein extraction from *E. coli* cultures, as shown in the figure below:

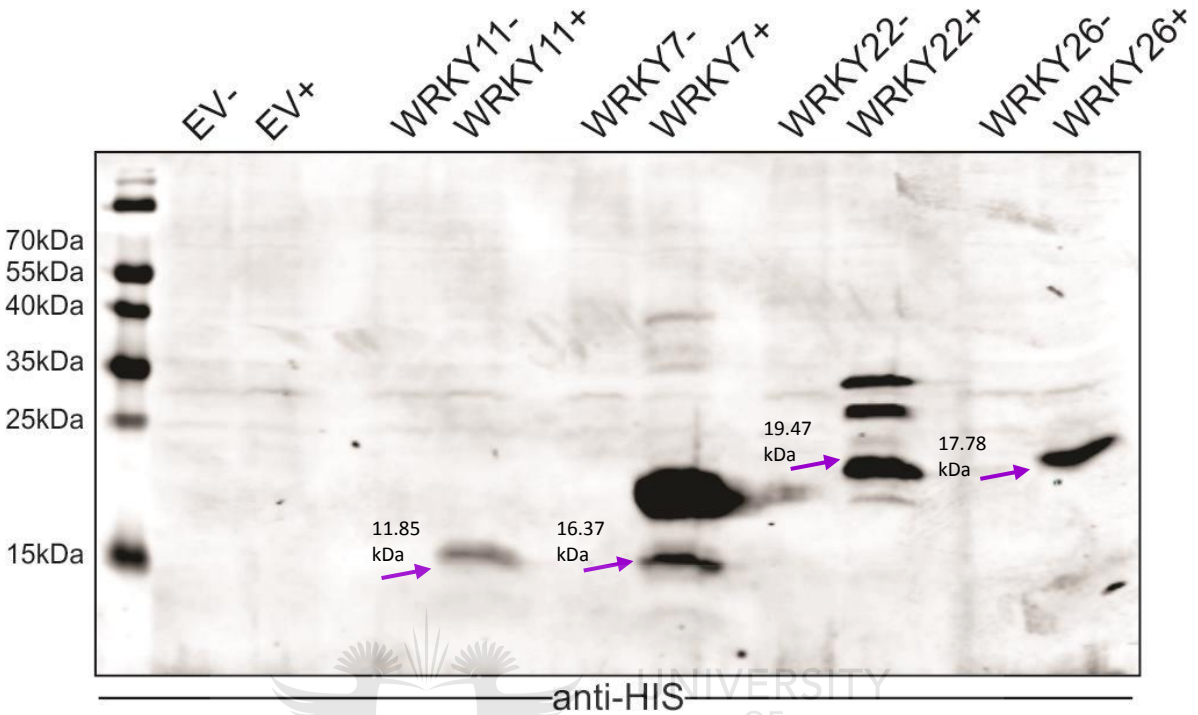



**Figure 4-15: A Western blot showing the successful expression of WRKY TFs -7, -11, -22 and -26 in BL21-AI *E. coli* after 4 to 6 hours of induction with IPTG.**

'EV' refers to the empty vector (pDEST42) negative control. The positive sign indicates the addition of IPTG. The expected size of the correct protein bands are indicated on the figure.

### 4.8.2. DPI-ELISA

The next step involved incubating the His-tagged WRKY proteins with biotinylated, ds oligos of the promoter regions containing the dyads, which were immobilised on a streptavidin-coated ELISA plate. The method detects protein-DNA interaction using an anti-His antibody conjugated with horseradish peroxidase (see 3.6.4).

Unfortunately, due to time constraints and a lack of resources, this part of the investigation could not be completed. Although multiple attempts were made to show an interaction between the WRKY TFs and the W-box dyads of At1g51850, the assay itself did not respond appropriately. It is an established method (Brand *et al*., 2010), but in this case, even the positive controls failed. Troubleshooting (Wanke

lab, ZMBP) determined that the most likely reason for this may have been faulty streptavidin-coated plates.

The DPI-ELISA technique remains a simple and attractive method to confirm the protein-DNA interactions suggested by the transient protoplast expression assays (see 4.7.2), so it is unfortunate that it failed here. Future studies could focus on tailoring this assay or other similar assays in order to prove that the W-box dyads on At1g51850 are directly responsible for its upregulation in response to flg22 and other P/MAMPS, and to determine the potentially responsible WRKY TFs.

# Chapter 5: Conclusion

At1g51790, At1g51800, At1g51820, At1g51850, At1g51860 and At1g51890, were investigated here using flg22 as a biotic stressor. These genes are part of a 14-member LRR-RLK cluster on *Arabidopsis thaliana* chromosome 1, and while they are themselves similar in sequence, their promoters are not. This means that their transcriptional regulation in response to elicitors of pathogenic origin likely varies from one to another. Microarray data that was already available gave the first hint that biotic stress induces the expression of these 6 genes. During this study, the luciferase-driving promoters of *At1g51790*, *At1g51850* and *At1g51890* were successfully cloned, and gene expression induced by flg22 was qualitatively measurable using a transient protoplast expression system. For the other three promoters the experimental set-up failed to detect that flg22 induced expression.

Bioinformatic analyses were used to determine that W-boxes are over-represented in At1g51790prom, At1g51850prom and At1g51890prom. This is significant since the WRKY transcription factors that bind to W-boxes are known to be involved in the transcriptional regulation of plant defense responses. Indeed, truncations that deleted W-boxes from the three promoters resulted in markedly lower responses to flg22 in the protoplast assay. And, mutational inactivation of individual W-boxes in the At1g51790prom also drastically reduced the flg22 response, except for the W-box closest to the start site, which actually seemed to increase both basal and flg22-inducible expression. In At1g51850prom, mutational inactivation of either or both of its W-box dyads resulted in virtually no flg22 inducibility. The deletion of 6 W-boxes in the promoter of *At1g51890,* done via truncation, drastically reduced both its basal expression and its inducible response to flg22. These results provide strong *in vitro* evidence that W-boxes, and W-box dyads in the case of At1g51850prom, may play a key role in the transcriptional regulation of these genes in response to flg22.

Further bioinformatic analyses indicated that four WRKY TFs show expression patterns similar to At1g51790, At1g51850 and At1g51890. These are WRKY7, -11, -22 and -26. WRKY22 is a known enhancer of pathogen resistance that acts downstream of the FLS2 receptor, while WRKY7, -11 and -26 enhance pathogen susceptibility. Here, WRKY7 and -11 were found to suppress basal as well as flg22-inducible expression of luciferase driven by the At1g51850 promoter. WRKY26 had no effect on basal expression but did suppress flg22 inducible expression. WRKY22 on the other hand, enhanced basal expression but suppressed flg22 inducibility. It seems that WRKY22 binds to the first W-box dyad to enhance basal expression, but, despite being a known transcriptional enhancer downstream of the FLS2 receptor, it is probably not responsible for the upregulation of At1g51850 in response to flg22.

Up to now we failed to show direct protein-DNA interaction in vitro between the W-box dyads of At1g51850prom and any of the WRKY TFs in question.

In conclusion then, this study provided substantial *in vitro* evidence that the W-boxes located upstream of the start sites of *At1g51790*, *At1g51850* and *At1g51890*, and especially the W-box dyads of *At1g51850*'s promoter, are involved in the regulation of these genes in response to flg22. Further work is needed to determine exactly which WRKY TFs bind to these W-boxes, but based on the findings presented here, WRKY TFs 7, 11, 22 and 26 are good candidates to investigate.

## 5.1. Future directions

The most prominent limitation of the present study was assay variability and/or failure, so other techniques, including quantitative real-time PCR, constitutive plant expression and the electrophoretic mobility shift assay could be applied in future to validate the expression kinetics and protein-DNA interaction results obtained here. Further validation of the involvement of these LRR-RLK genes in *Arabidopsis* immunity could also be gained by investigating the effects and signalling pathways of other M/PAMPs like LPS, HrpZ and chitin on the expression of the genes, before undertaking a comprehensive study of the structure and functional interactions of the LRR-RLK protein products.

# Chapter 6: References

Abramovitch, R. B., Anderson, J. C., Martin, G. B., 2006. Bacterial elicitation and evasion of plant innate immunity. *Nature Reviews Molecular Cell Biology,* 7, 601-6011.

Afzal, A. J., Wood, A. J., Lightfoot, D. A., 2008. Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Molecular Plant-Microbe Interactions,* 21, 507-517.

Asai, T., Tena G., Plotnikova, J., Willmann, M. R., Chiu, W. L., Gomez-Gomez, L., Boller, T., Ausubel, F. M., Sheen. J., 2002. MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature,* 415, 977-983.

Ausubel, F. M., 2005. Are innate signalling pathways in plants and animals conserved? *Nature Immunology,* 6, 973-979.

Bari, R., Jones, J. D. G., 2009. Role of plant hormones in plant defence responses. *Plant Molecular Biology,* 69, 473-488.

Bomblies, K., Weigel, D., 2007. Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nature Reviews Genetics,* 8, 382-393.

Brand, L. H., Kirchler, T., Hummel, S., Chaban, C., Wanke, D., 2010. DPI-ELISA: a fast and versatile method to specify the binding of plant transcription factors to DNA *in vitro*. *Plant Methods,* 6, doi:10.1186/1746-4811-6-25.

Chae, L., Sudat, S., Dudoit, S., Zhi, T., Luan, S., 2009. Diverse transcriptional programs associated with environmental stress and hormones in the *Arabidopsis* receptor-like kinase gene family. *Molecular plant,* 2, 84-107.

Chi, Y., Yan, Y., Zhou, Y., Zhou, J., Fan, B., Yu, J., Chen, Z., 2013. Protein–Protein Interactions in the Regulation of WRKY Transcription Factors. *Molecular Plant,* 6, 287-300.

Chinchilla, D., Boller, T., Robatzek, S., 2007. Flagellin signalling in plant immunity. *Advances in experimental medicine and biology: Current topics in innate immunity,* 598, 358–371.

Ciolkowski, I., Wanke, D., Birkenbihl, R., Somssich, I., 2008. Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Molecular Biology,* 68, 81-92.

Conant, G. C., Wolfe, K. H., 2008. Turning a hobby into a job: how duplicated genes find new functions. *Nature Reviews Genetics,* 9, 938-950.

Consonni, C., Humphry, M., Panstruga, R., 2009. Defence responses in plants. *Plant Relationships (2nd Edition),* Edited by Deising, H. B., Berlin Heidelberg, Springer, Chapter 18, 363-385.

Cui, H., Xiang, T., Zhou, J., 2009. Plant immunity: a lesson from pathogenic bacterial effector proteins. *Cellular Microbiology,* 11, 1453-1461.

Cunha, L., McFall, A. J., Mackey, D., 2006. Innate immunity in plants: a continuum of layered defences. *Microbes and Infection,* 8, 1372-1381.

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

Danna, C. H., Millet, Y. A., Koller, T., Han, S., Bent, A. F., Ronald, P. C., Ausubel, F. M., 2011. The *Arabidopsis* flagellin receptor FLS2 mediates the perception of Xanthomonas Ax21 secreted peptides. *Proceedings of the National Academy of Sciences,* 108, 9286-9291.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet F., Dufayard, J.F., Guindon, S., Lefort, V., Lecot, M., Claverie, J.M., Gascuel, O., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research,* 36, doi: 10.1093/nar/gkn180.

Dong, J., Chen, C., Chen, Z., 2003. Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response. *Plant Molecular Biology,* 51, 21-37.

Dunning, F. M., Sun, W., Janser, K. L., Helft, L., Bent, A. F., 2007. Identification and mutational analysis of *Arabidopsis* FLS2 leucine-rich repeat domain residues that contribute to flagellin perception. *The Plant Cell,* 19, 3297-3313.

Eulgem, T., Rushton, P. J., Robatzek, S.,Somssich, I. E., 2000. The WRKY superfamily of plant transcription factors. *Trend in Plant Science,* 5, 199-206.

Eulgem, T., Rushton, P. J., Schmelzer, E., Hahlbrock, K., Somssich, I. E., 1999. Early nuclear events in plant defense: rapid gene activation by WRKY transcription factors. *The EMBO Journal,* 18, 4689-4699.

Eulgem, T., Somssich, I., 2007. Networks of WRKY transcription factors in defense signalling. *Current Opinion in Plant Biology,* 10, 366–371.

Feng, F., Zhou, J. M., 2012. Plant-bacterial interactions mediated by type III effectors. *Current Opinion in Plant Biology,* 15, 469-476.

Fisher, C. L., Pei, G. K., 1997. Modification of a PCR-based site-directed mutagenesis method. *Biotechniques,* 23, 570-574.

Giacomelli, J. I., Weigel, D., Chan, R. L., Manavella, P. A., 2012. Role of recently evolved miRNA regulation of sunflower HaWRKY6 in response to temperature damage. *New Phytologist,* 195, 766-773.

Goff, K. E., Ramonell, K. M., 2007. The role and regulation of receptor-like kinases in plant defense. *Gene Regulation and Systems Biology,* 1, 167-175.

Gómez-Gómez, L., Baue, Z., Boller, T., 2001. Both the extracellular leucine-rich repeat domain and the kinase activity of FLS2 are required for flagellin binding and signalling in *Arabidopsis*. *Plant Cell,* 13, 155-163.

Gou, X. *et al*., 2010. Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in *Arabidopsis thaliana*. *BMC Genomics,* 11, 1-15.

Greeff, C., Roux, M., Mundy, J., Petersen, M., 2012. Receptor-like kinase complexes in plant innate immunity. *Frontiers in Plant Science,* 3, doi: 10.3389/fpls.2012.00209.

Gust, A. A., Biswas, R., Lenz, H. D., Rauhut, T., Ranf, S., Kemmerling, B., Götz, F., Glawischnig, E., Lee, J., Felix, G., Nürnberger, T., 2007. Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in *Arabidopsis*. *Journal of Biological Chemistry,* 282, 32338-32348.

Guttman, D. S., Gropp, S. J., Morgan, R. L., Wang, P. W., 2006. Diversifying selection drives the evolution of the type III secretion system pilus of *Pseudomonas syringae*. *Molecular Biological Evolution,* 23, 2342–2354.

He, P., Shan, L., Sheen, J., 2007. Elicitation and suppression of microbe-associated molecular pattern-triggered immunity in plant–microbe interactions. *Cellular Microbiology,* 9, 1385-1396.

Henry, E., Yadeta, K. A., Coaker, G., 2013. Recognition of bacterial plant pathogens: local, systemic and transgenerational immunity. *New Phytologist,* 199, 908-915.

Hok S., Danchin, E. G., Allasia, V., Panabières, F., Attard, A., Keller, H., 2011. An *Arabidopsis* (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease, *Plant Cell and Environment*, 34, 1944-1957.

Hothorn, M., Belkhadir, Y., Dreux, M., Dabi, T., Noel, J. P., Wilson, I. A., Chory, J., 2011. Structural basis of steroid hormone perception by the receptor kinase BRI1. *Nature,* 474, 467-471.

Hunter, B. G., Beatty, M. K., Singletary, G. W., Hamaker, B. R., Dilkes, B. P., Larkins, B. A., Jung, R., 2002. Maize opaque endosperm mutations create extensive changes in patterns of gene expression. *Plant Cell,* 14, 2591-2612.

Johnson, K. L., Ingram, G. C., 2005. Sending the right signals: regulating receptor kinase activity. *Current Opinion in Plant Biology,* 8, 648-656.

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

Journot-Catalino, N., Somssich, I. E., Roby, D., Kroj, T., 2006. The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis thaliana*. *The Plant Cell,* 18, 3289-3302.

Kaneda, T., Taga, Y., Takai, R., Iwano , M., Matsui, H., Takayama, S., Isogai, A., Che, F.S., 2009. The transcription factor OsNAC4 is a key positive regulator of plant hypersensitive cell death. *The EMBO journal,* 28, 926-936.

Karimi, M., Depicker, A., Hilson, P., 2007. Recombinational cloning with plant Gateway vectors. *Plant Physiology,* 145, 1144-1154.

Kemmerling, B., Halter, T., Mazzotta, S., Mosher, S., Nürnberger, T., 2011. A genome-wide survey for Arabidopsis leucine-rich repeat receptor kinases implicated in plant immunity. *Frontiers in Plant Science,* 2, doi: 10.3389/fpls.2011.00088.

Kemmerling, B., Schwedt, A., Rodriguez, P., Mazzotta, S., Frank, M., Qamar, S. A., Mengiste, T., Betsuyaku, S., Parker, J. E., Müssig, C., Thomma, B. P., Albrecht, C., de Vries, S. C., Hirt, H., Nürnberger, T., 2007. The BRI1-associated kinase 1, BAK1, has a brassinolide-independent role in plant cell-death control. *Current Biology,* 17, 1116-1122.

Knogge, W., Lee, J., Rosahl, S., Scheel, D., 2009. *Plant Relationships (2nd Edition): Signal Perception and Transduction in Plants,* Edited by Deisinng H. B., Berlin Heidelberg, Springer, Chapter 17, 338-361.

Kobe, B., Kajava, A., 2001. The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology,* 11, 725-732.

Lee, H., Chah, O. K., Sheen, J., 2011. Stem-cell-triggered immunity through CLV3p-FLS2 signalling. *Nature*, 473, 376-379.

Lee, J., Klüsener, B., Tsiamis, G., Stevens, C., Neyt, C., Tampakaki, A. P., Panopoulosi, N. J., Nöller, J., Weiler, E. W., Cornelis, G. R., Mansfield, J. W., Nürnberger, T., 2001. HrpZPsph from the plant pathogen *Pseudomonas syringae pv. phaseolicol*a binds to lipid bilayers and forms an ion-conducting pore *in vitro*. *Proceedings of the National Academy of Sciences,* 98, 289-294.

Li, F., Pignatta, D., Bendix, C., Brunkard, J. O., Cohn, M. M., Tung, J., Sun, H., Kumar, P., Baker, B., 2012. MicroRNA regulation of plant innate immune receptors. *Proceedings of the National Academy of Sciences,* 109, 1790-1795.

Libault, M., Wan, J., Czechowski, T., Udvardi, M., Stacey, G., 2007. Identification of 118 *Arabidopsis* transcription factor and 30 ubiquitin-ligase genes responding to chitin, a plant-defense elicitor. *Molecular Plant Microbe Interactions,* 20, 900-911.

Liu, Z., Yan, L., Wu, Z., Mei, C., Lu, K., Yu, Y., Liang, S., Zhang, X., Wang, X., Zhang, D., 2012. Cooperation of three WRKY-domain transcription factors WRKY18, WRKY40, and WRKY60 in repressing two ABA-responsive genes ABI4 and ABI5 in *Arabidopsis*. *Journal of Experimental Botany,* 63, 6371-6392.

Mackey, D., McFall, A. J., 2006. MAMPs and MIMPs: proposed classifications for inducers of innate immunity. *Molecular Microbiology,* 61, 1365-1371.

Mare, C., Mazzucotelli, E., Crosatti, C., Francia, E., Stanca, M. A., Cattivelli, L., 2004. Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. *Plant Molecular Biology,* 55, 399-416.

Matsushima, N., Miyashita, H., Mikami, T., Kuroki, Y., 2010. A nested leucine rich repeat (LRR) domain: The precursor of LRRs is a ten or eleven residue motif. *BMC Microbiology,* 10, doi: 10.1186/1471-2180-10-235.

McDowell, J. M., Simon, S. A., 2008. Molecular diversity at the plant-pathogen interface. *Developmental and Comparative Immunology,* 32, 736-744.

McGreal, E. P., 2009. Target pattern recognition in innate immunity: structural basis of pattern recognition by innate immune molecules. *Advances in Experimental Medicine and Biology (Volume 653)*, Edited by Kishore, U., Landes Bioscience and Springer Science, Business Media, Chapter 10, 139-161.

Monaghan, J., Weihmann, T., Li, X., 2009. Plant-Environment Interactions: Plant innate immunity. In: *Signaling and Communication in Plants.* Berlin Heidelberg: Springer, 119-136.

Muthamilarasan, M., Prasad, M., 2013. Plant innate immunity: an updated insight into defense mechanism. *Journal of Biosciences,* 38, 433-449.

Nakayama, A., Fukushima, S., Goto, S., Matsushita, A., Shimono, M., Sugano, S., Jiang, C., Akagi, A., Yamazaki, M., Inoue, H., Takatsuji, H., 2013. Genome-wide identification of WRKY45-regulated genes that mediate benzothiadiazole-induced defense responses in rice. *BMC Plant Biology,* 13, doi:10.1186/1471-2229-13-150.

Navarro, L., Jay, F., Nomura, K., He, SY., Voinnet, O., 2008. Suppression of the microRNA pathway by bacterial effector proteins. *Science,* 321, 964-967.

Nishimura, M. T., Dangl, J. L., 2010. *Arabidopsis* and the plant immune system. *Plant Journal,* 61, 1053-1066.

Ovecka, M., Lichtscheidl, K., 2009. Signaling in plants: MAPK signalling. *Signalling and Communication in Plants,* Edited by Baluska F., Mancuso, S., Berlin Heidelberg, Springer, 51-69.

Pandey, S., Somssich, I., 2009. The role of WRKY transcription factors in plant immunity. *Plant Physiology,* 150, 1648-1655.

Pape, S., Thurow, C., Gatz, C., 2010. The *Arabidopsis* PR-1 promoter contains multiple integration sites for the coactivator NPR1 and the repressor SNI1. *Plant Physiology,* 154, 1805-1818.

Pel, M. J. C., Pieterse, C. M. J., 2013. Microbial recognition and evasion of host immunity. *Journal of Experimental Botany,* 64, 1237-1248.

Pieterse, C. M. J., Van der Does, D., Zamioudis, C., Leon-Reyes, A., Van Wees, S. C. M., 2012. Hormonal modulation of plant immunity. *Annual Review of Cell and Developmental Biology,* 28, 489–521.

Postel, S., Kemmerling, B., 2009. Plant systems for recognition of pathogen-associated molecular patterns. *Seminars in Cell and Developmental Biology,* 20, 1025-1031.

Postel, S., Küfner, I., Beuter, C., Mazzotta, S., Schwedt, A., Borlotti, A., Halter, T., Kemmerling, B., Nürnberger, T., 2010. The multifunctional leucine-rich repeat receptor kinase BAK1 is implicated in *Arabidopsis* development and immunity. *European Journal of Cell Biology,* 89, 169-174.

Rasmussen, M. W., Roux, M., Petersen, M., Mundy, J., 2012. MAP kinase cascades in *Arabidopsis* innate immunity. *Frontiers in Plant Science,* 3, doi: 10.3389/fpls.2012.00169.

Robatzek, S., Wirthmueller, L., 2012. Mapping FLS2 function to structure: LRRs , kinase and its working bits. *Protoplasma,* 250, 671-681.

Rodakowska, E., Derba-Maceluch, M., Kasprowicz, A., Zawadzki, P., Szuba, A., Kierzkowski, D., Wojtaszek., 2009. Signalling and cell walls. *Signalling and Communication in Plants*, Edited by Baluska F., Mancuso, S., Berlin Heidelberg, Springer, 173-192.

Roux, M., Zipfel, C., 2012. Receptor kinase interactions: complexity of signalling. *Signalling and Communication in Plants,* 13, 145-172.

Rushton, P. J., Somssich, I. E., Ringler, P., Shen, Q. J., 2010. WRKY transcription factors. *Trends in Plant Science,* 15, 247-258.

Sambrook, J., Russel, D. W., 2001. *Molecular Cloning: A laboratory manual (3rd edition)*, New York, Cold Spring Harbor Laboratory Press.

Sanabria, N. M., Huang, J. C., Dubery, I. A., 2010. Self/nonself perception in plants in innate immunity and defense. *Self Nonself,* 1, 40-54.

Schwessinger, B., Ronald, P. C., 2012. Plant innate immunity: perception of conserved microbial signatures. *Annual Review of Plant Biology,* 63, 451-482.

She, J., Han, z., Kim, T., Wang, J., Cheng, W., Chang, J., Shi, S., Wang, J., Yang, M., Wang, Z., Chai, J., 2011. Structural insight into brassinosteroid perception by BRI1. *Nature,* 474, 472-476.

Shiu, S., Bleecker, A. B., 2001. Plant receptor-like kinase gene family: diversity, function, and signalling. *Science Signaling,* 113, doi: 10.1126/scisignal.1132001re22.

Singh, K. B., Foley, R. C., Oñate-Sánchez, L., 2002. Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology,* 5, 430-436.

Slaughter, A., Daniel, X., Flors, V., Luna, E., Hohn, B., Mauch-Mani, B.m 2012. Descendants of primed *Arabidopsis* plants exhibit resistance to biotic stress. *Plant Physiology,* 158, 835-843.

Spoel, S. H., Dong, X., 2012. How do plants achieve immunity? Defence without specialized immune cells. *Nature Reviews Immunology,* 12, 89-100.

Sun, Y., Li, L., Macho, A. P., Han, Z., Hu, Z., Zipfel, C., Zhou, J., Chai, J., 2013, Structural basis for flg22-induced activation of the *Arabidopsis* FLS2-BAK1 immune complex. *Science*, 342, 624-628

Tax, F., Kemmerling, B., 2012. Receptor-like kinases in plants: from development to defense. *Signalling and Communication in Plants, 13.*

Tena, G., Boudsocq, M., Sheen, J., 2011. Protein kinase signaling networks in plant innate immunity. *Current Opinion in Plant Biology,* 14, 519-529.

Thomas, J. H., 2006. Analysis of homologous gene clusters in *Caenorhabditis elegans* reveals striking regional cluster domains. *Genetics,* 172, 127-143.

Ülker, B., Somssich, I. E., 2004. WRKY transcription factors: from DNA binding towards biological function. *Current Opinion in Plant Biology,* 7, 491-498.

Van Aken, O., Zhang, B., Law, S., Narsai, R., Whelan, J., 2013. AtWRKY40 and AtWRKY63 modulate the expression of stress-responsive nuclear genes encoding mitochondrial and chloroplast proteins. *Plant Physiology,* 162, 254-271.

Vyas, J., Gryk, M. R., Schiller, M. R., 2009. VENN, a tool for titrating sequence conservation onto protein structures. *Nucleic Acids Research,* 37, doi: 10.1093/nar/gkp616.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G., Provart, N., 2007. An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data set. *PLoS ONE,* 8, doi: 10.1371/journal.pone.0000718.

Yan, L., Liu, Z., Xu, Y., Lu, K., Wang, X., Zhang, D., 2013. Auto- and cross-repression of three *Arabidopsis* WRKY transcription factors WRKY18, WRKY40, and WRKY60 negatively involved in ABA signaling. *Journal of Plant Growth Regulation,* 32, 399-416.

Yi, G., Sze, S., Thon, M. R., 2007. Identifying clusters of functionally related genes in genomes. *Bioinformatics,* 23, 1053-1060.

Yoo, S., Cho, Y., Sheen, J., 2007. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols,* 2, 1565-1572.

Zhang, W., Gao, S., Zhou, X., Chellappan, P., Chen, Z., Zhou, X., Zhang, X., Fromuth, N., Coutino, G., Coffey, M., Jin, H., 2011. Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks. *Plant Molecular Biology,* 75, 93-105.

Zhang, X., 1998. Leucine-rich repeat receptor-like kinases in plants. *Plant Molecular Biology Reporter,* 16, 301-311.

Zheng, Z., Fan, B., Chen, Z., 2005. *Roles of structurally related WRKY 20, WRKY25, WRKY26 and WRKY33 transcription factors in plant defense responses.* 16th International Conference on Arabidopsis Research, TAIR.

Zheng, Z., Mosher, S. L., Fan, B., Klessig, D.F., Chen, Z., 2007. Functional analysis of Arabidopsis WRKY25 transcription factor in plant defense against *Pseudomonas syringae*. *BMC Plant Biology,* 7, doi: 10.1186/1471-2229-7-2.



Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., Gruissem, W., 2004. Genevestigator: *Arabidopsis* microarray database and analysis toolbox. *Plant Physiology,* 136, 2621-2632.

Zipfel, C., 2008. Pattern-recognition receptors in plant innate immunity. *Current Opinion in Immunology,* 20, 10-16.

# Appendix I: W-box locations

The sequences below indicate mutated W-boxes in blue and red, as well as position where the deletion mutants were created in square brackets and in purple. Correlate with Figure 4-5: Graphical representation of the promoters indicating W-Boxes and deletion target sites.Figure 4-5.

## I.I. The promoter region of At1g51790

>AT1G51790 at1g51790; upstream from -1500 to 2; size: 1503; feature type:cds; location: Arabidopsis_thaliana:NC_003070.9:19210572-19212074:R; upstream neighbour: AT1G51800 (distance: 3628)

TATAAAATATCTTATTTAGAAATAATAGCATTTTAGTAGAAAATGTTGAAGTAAAGAAAAATCATATCATTAGCT
GTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGTT**TTGAC**CAAAAAAATCATCATACTGTATTACTGTAC
AGTGTACATGTACTTTAATTATAAATCAGAATCATGGATGATTCGATAAATAATACTTTCTTAATATGCATATAT
CAAGTAACAGGAGGAGAGGAGACTTTAATTTCAATGGCACAACGAGTCCATTGTCTACGATGTGGGACAGATTAG
CCTAAAGGTATGAAACAAATCTGAGGGTGAATTTCTGAAAACAATGAAAACTATGAGAAACCCTGATAGCCATTA
CATTCACAGGACGTCCAAGTGGAAGAATGACCTGTTTTATAACATGGAAGTCGCTCAGGAATCTATG**[1086]**GA
TCCAATACTGCTGCACCTTTCATTATCGCTGCCACGTTTTTCTTCTTCTCCTCTACTTGTCTTGTCTTCTTAACT
TTTTTGTGTGTTTTTATTGTCTTTCTCTTTAGGTTTAGTTGCAAAGCCTCTTGTGAAATTAGTAAAGACATTGTG
AGAATAATGGTTGTTTAAGGATGTACGAAGAATCCTATGAGACATAACTCCAGCACCTGCTGGTGAAGTCAAACC
ACAACCGGTGGTGAAGACCTTTCTCATCAGCCAAAAACGTTGCAAATTCGCCATTATTCCGACAGACTTCAGGAG
AAGGAAAAAAAAAGTAAGCGTGTAGCTAGGTCAACTTACGTGAA[722]GATCTTAATTAGGTTTTCTCTTTTGT
ACTAGTCGTAGGCATAACAACTTTTTTCTATCAAAAAAAGGAAATTTGCAAGAATTTACAAAAAATAAACTAATC
ACTTCAAAATAACCCTTAATTTACTACAATTTTAACTATTTTCGGTATCAGAATATCGATTTGATACAAATTCTC
AGCAACAATCTCCCTTTATATTGATTCTTCCGACACAAGTTTCTGATTTACACACCATTTACCAAGTGGTTTTGT
TCAAACAAATCATGGTGCAAGACACTTTTTACCCAATTGTCTTGTTAGGTTACAAAGTTATAGATTTTCTGACCC
ATTTTTTACAGAACAATTACAGAGTAGAGAACCCATGTGAACATAAGCATTTTGTTTTGGTCAAACCAAAGAAAT
TAAGAAAGTCTCTGGTTTTTGTCTACACATATCGTTGACTTGTTCATTAACATATACACAAGTTATCTAAAGAAA
TCAACTAAAGAAAAAACAAGTTGGCTTTCTTGTTGATTAGTGTTCTACTTGCTGTTAGATAATGATTTATACGTT
GTGACTTGTGAGCACACAAAATTGATATCTACCAAAGAGTTGACATTGTCTTCAAGTTGTTGACTTTGAACCCAA
AAAAAACTCTGGTGTTTCTCAGAAATGGATAAACCAGTTTTAAGAAATAGGAGATGATGGTTGTTTCTGCTGTTT
GAAGAACCGAAATG

## I.II. The promoter region of At1g51850

>AT1G51850 at1g51850; upstream from -850 to 2; size: 853; feature type:cds; location: Arabidopsis_thaliana:NC_003070.9:19256781-19257633:R; upstream neighbour: AT1G51860 (distance: 850)

ATTTGAGAAATTCTTTTATGTGATTTTATGGGAAAGCAATCTTGTTGTTATTTTGTTTCTTTGTGTATTTCCCAA
TGTATGTTTTGAAGAAACAATGTACATTCTATCTTCGTGGGTGCAGGTGTTTTGATTACTCTGTTTTGAAACTTT
TTGTGTGTTACTATATAAATATTTTTCTTAAAAGGATTAATTATGTAATACTCTTTATTTTAATTTTTAC[641
]GTATTAATGCATCAAGCTAGTTAACAACTAGTGACAGCTCAATTTTCGGAATCATACACATTTAGAAGTCAAGG
GCTGACTAAAATCATTGGAGGTTGCATTGGCATGCACAAACATTACAGGCCACGCCCACATTGATATAGGAGACT
TCTCAAACATGATCGTTGACTTTTGACGTTGGATTCATCCAGGACGCCAAAGACACAAACAAAGAGACAAATCAA
CCAACGAATTGCCATACAAAAGTCAAAGTCAACAAATT[379]CATGAGAACATTTGTGGATTAATAGTAATAA
TACACCGAAATTACATTATGTGGATTGATAATACACCGTTGTTTTGAGTTTACTTATAGTCAACGGTGTTTTTTG
TTCTAAAGAACATAACAATAATAACTCTTGTCATTGTTGACGAAAAACAAATACTACTTCTTATATCATTTTCAC
TCAACTTGCTCTATCGAATCATTTTGCATCGTGAAAACATATTCTCAATCAAT[145]TAATTTATTCAATTTC
AATTATGTTATAAGCAGTTGACTTGGACTTTTGAGTAGTATAGCTCATATAGTCATATTCTCATATGCATAACTT
GCTATCTTAAAAAAAAGCTCTCCTGTGGACAGTAAGGAGAACAATG

## I.III. The promoter region of At1g51890

>AT1G51890 at1g51890; upstream from -1355 to 2; size: 1358; feature type:cds; location: Arabidopsis_thaliana:NC_003070.9:19278526-19279883:R; upstream neighbour: AT1G51900 (distance: 1355)

GCTTCCAGGATGAGACACGAAGATTATCAACACCAAAAGGAAGATGAAGAAGACTATAACCAAACCACTAGGGTT
TTTTGATTCGGAGTCGGACCAACCTTAAGGTTGGAAAGCTAACTAAAGCTTCGCCTCTGTCTTTTTGAATTTTCC
ACAGTCCAATTTAGGATTCAAAGTCTCTTTTTGTGTTGAAGAATGAGAGACGAAGATTATTAACCCAAAAAGAAG
AAGAAGAAGAAAGACGAAGATAATTAACCAAACCGCTAGTTTTTTTTTTTGGTTTTGATTCGGAGTCGGACCAAC
TTTAAGGTTGGAAAGCTATTTAAAGCTTCAGCTCAGTCTTTTTTGAATGAAGTCTTAATTGATGTACTATATAAT
CGAAAATGAAAATAGTGTATATGA[957]GATCTTTAATTTTGTTACCGTATAAAATATGTTGATTCATTTTAC
CAATTTGGGTATATCTTACAAAAAAGTTTATGTTACTCGTTTTTAATTAAATAAGCAATCAAATGACTTAGTCAT
AAAATTTAGTGAAAATTATCCAGATTTGAGTCATGGATTTATAAGGGAATATTATAGGCGACGTCAAAAATCAAA
ATCACCTAAGTTGACTTAT[743]TAATTTCTAAATTTTTCTTCACTTTATTTTTAAAAAAGAATTTCAAACTC
ATCGAAATTTGTCAAAATAATTTCAGATAAATATATACCATATACGTCAACGGACTATAATATCGAGTTTTTTTC
CACACTTAATATTAGAAATTTCAGAAATGAATATTATAGGCAACGACGACGTCAAAAATCAAAATCACCTAAGTT
GACTTATTCTTTTCTAATTTTTCTTCACTTTATTTTTTAAAAAGAATTTTAAACACATCGAAATTTGTCAAAATA
ATTCCAAATAAATATATACAATATACGTCAACGAACTATAATATCGAGATTTTTTTCACACAAAATAAAAGTACG
ATTGTTTAGTAGTTTATTTACGCTCTCTTTTTGTTATTATAGTTGCTTTTATTGGAAG*(334bp)*CTTGGAAGGC
AAGTTGGGTATTTTTGTTTTGTTTAAGCAAGATTCTGCAACTTTCTTTGTTTGTATCTTTCCATTCTTTATTTAC
TTTCGTTTGATTTCTCATTTTGTTTTTATATATTTGCCACCAATTCTAGTTTGAACATTTCCACAAGACACATAA
TTTCGTAGGCAAATAAGAGTGCTGTTACAATCACCTGAATAAGTCTAATGTTGAAGAAGAAGACTTGGCCATCTT
AATTGACCAATACTTTGCTCTTATGTCAATCTTTACCCACCAAGAACCTATGAAAAATGAATAGCCTCCTCTGAT
TCAGAACAAACATCCCCAAAAAACATG