Investigation of Apoplastic Effectors from *Arabidopsis* Downy Mildew Pathogen *Hyaloperonospora arabidopsidis*

by

Gülin Boztaş Demir

Submitted in partial fulfilment of the requirements

for the degree of Doctorate of Philosophy

University of Worcester

May, 2015

ABSTRACT

Downy mildew is a plant disease, caused by a group of obligate oomycete pathogens, which results in significant economic loss by affecting plants such as, tomato, potato, grapevine, lettuce, onion and brassicas. The downy mildew pathogens secrete molecules called effectors, which modulate plant innate immunity and enable parasitic colonisation. Despite the increasing attention, the knowledge of these effectors is highly limited. Especially, the apoplastic effectors are understudied with respect to cytoplasmic effectors. Therefore, in this study, we focused on identifying novel apoplastic effectors, taking advantage of an established model pathosystem of the *Arabidopsis thaliana* and its adapted downy mildew pathogen *Hyaloperonospora arabidopsidis* (*Hpa*). With this purpose, two different approaches were followed: Firstly, five candidate genes were determined from the Expressed Sequence Tag library of *Hpa*-Emoy2 isolate using bioinformatic tools. The candidates were *Hpa* 804480, *Hpa* 806249, *Hpa* 814231, *Hpa* 814014 and *Hpa* 813915, all of which contained signal peptides and lacked known motifs of cytoplasmic effector and intracellular functions. The candidate genes were validated for their expression by challenging the susceptible Ws-*eds*1 plants with the pathogen. It was observed that the candidates were expressed at varying times and levels during the infection suggesting that some encoded proteins by the candidate genes can be put forward by the pathogen right away or as a second layer of attack to suppress the PTI later during the infection. Additional assays also showed that the candidate genes were not triggering hypersensitive response in *Nicotiana* plants. Moreover, the candidate genes were examined for their variations across other *Hpa* isolates, and the results suggested they are under less pressure to evolve with respect to cytoplasmic effectors; thus it is possible that they may be avoiding recognition. Secondly, the intercellular washing fluid (IWF) of the *Hpa* infected *Arabidopsis* plants were scrutinized for the capability of triggering defence responses in wild type and receptor-impaired plants. GUS reporter assays and cumulative quantification of Reactive Oxygen Species assays suggested that the infected IWF, indeed, was inducing immune

activation on both wild-type and mutant plants. The complex IWF was then simplified to identify the active fraction. One fraction, the flow-through, was found to be responsible for triggering defence responses with the same assays. In addition, the IWF sample was analysed via MALDI-TOF mass spectrometry, and results showed that there were cysteine-rich proteins from *Hpa* with putative apoplastic effector characteristics. Two genes from the cysteine-rich proteins, *Hpa* 806256 and *Hpa* 813024 were tried in *in vitro* expression assays. Additionally, MALDI-TOF revealed *A. thaliana* that were annotated as LRR family proteins with signal peptides; which carried high potential of recognizing apoplastic effectors. T-DNA insertional *A. thaliana* mutant lines for those LRR protein encoding genes, AT1G33610.1, AT3G20820.1 and AT1G49750.1 were use in assessment of interaction phenotypes and recognition of the molecules within the active fraction. The assays suggested that these candidate receptors may not be the corresponding receptors, or they may need co-receptors to function properly. Overall, we believe, the findings of this research will indeed contribute to future studies on apoplastic effectors and their recognition.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

LIST of ABBREVIATIONS

A	Alanine
A. candida	*Albugo candida*
A. thaliana	*Arabidopsis thaliana*
A. tumefaciens	*Agrobacterium tumefaciens*
ABRC	Arabidopsis Biological Resource Centre
Amp	Ampicillin
APS	Ammonium persulfate
ATR	*Arabidopsis thaliana* recognized
Avr	Avirulence
B. lactucae	*Bremia lactucae*
BAK1	Brassinosteroid insensitive 1-associated kinase 1
BLAST	Basic Local Alignment Search Tool
BLASTP	Protein-protein BLAST
BP	Border primer
BRI1	Brassinosteroid insensitive 1
BSA	Bovine serum albumin
C (Cys)	Cysteine
Ca^{++}	Calcium ion
CBEL	Cellulose binding elicitor lectin
cDNA	Complementary DNA
CEBiP	Chitin elicitor binding protein
CERK1	Chitin elicitor receptor kinase 1
cm	Centimetre

Col	Columbia
CRN	Crinkler
D	Aspartate
dai	Days after infection/inoculation
DAMP	Damage-associated molecular pattern
DEPC	Diethylpyrocarbonate
dH$_2$O	Distilled water
DNA	Deoxyribonucleic acid
E	Glutamate
E. amylovora	*Erwinia amylovora*
E. coli	*Escherichia coli*
Ed	Edinburgh
EDS1	*Enhanced disease susceptibility 1*
EDTA	Ethylenediaminetetraacetic acid
EFR	Elongation Factor-Tu Receptor
EGase	Endo-β-1,3-glucanases
EHM	Extrahaustorial matrix
Elf	Elongation factor-Tu
Em	East Malling
EPI	Extracellular protease inhibitor
EST	Expressed sequence tag
ET	Ethylene
ETI	Effector-triggered immunity
ETS	Effector-triggered susceptibility
flg	Flagellin
FLS2	Flagellin Sensing 2

F	Phenylalanine
FT	Flow through
G	Glycine
GC	Guanine Cytosine
Gen	Gentamicin
GFP	Green fluorescent protein
GIP	Glucanase inhibitor protein
GUS	β-glucuronidase
GW	Gateway
h	Hour
H (His)	Histidine
H$_2$O	Water
H$_2$O$_2$	Hydrogen peroxide
H. parasitica	*Hyaloperonospora parasitica*
HCl	Hydrogen chloride
Hi	Hilliers
HaNLP	*Hyaloperonospora arabidopsidis* NEP1-like protein
Hpa	*Hyaloperonospora arabidopsidis*
HR	Hypersensitive response
ID	Identity
in	inch
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITS	Internal transcribed spacer
IWF	Intercellular washing fluid
JA	Jasmonic acid
K	Lysine

K_2HPO_4	Dipotassium phosphate
$K_3Fe(CN)_6$	Potassium ferricyanide
Kan	Kanamycin
kb	Kilobase
KCl	Potassium chloride
kDa	Kilodalton
KOH	Potassium hydroxide
Ks	Keswick
L	Leucine
LB	Lysogeny broth/Left border
LecRK	Lectin receptor kinase
Ler	Landsberg erecta
LP	Left border primer
LRR	Leucine-rich repeats
Lys-M	Lysin motif
M	Molar
mA	Milliampere
MALDI-TOF	Matrix-assisted laser desorption/ionisation-Time of flight
MAMP	Microbe-associated molecular pattern
MAPK	Mitogen-activated protein kinase
MCS	Multiple cloning site
mg	Milligrams
min	minutes
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
min	Minutes

ml	Millilitre
mM	Millimolar
mm	Millimetre
MS	Murashige-Skoog
N$_2$	Nitrogen
N. benthamiana	*Nicotiana benthamiana*
N. tabacum	*Nicotiana tabacum*
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NASC	Nottingham Arabidopsis Stock Centre
NB	Nucleotide binding
NCBI	National Center for Biotechnology Information
NEP	Necrosis and ethylene- inducing peptide
Ni-NTA	Nickel-Nitrilotriacetic acid
NLP	NEP1-like protein
NLS	Nuclear localisation signal
Nd	Niederzenz
nM	Nanomolar
nm	Nanometre
NO	Nitric oxide
Nudix	Nucleoside diphosphate linked to X
O.D.	Optical density
O2$^-$	Superoxide
Oy	Oystese
P	Proline
P. brassicae	*Phytophthora brassicae*

P. cactorum	*Phytophthora cactorum*
P. capsici	*Phytophthora capsici*
P. infestans	*Phytophthora infestans*
P. phaseoli	*Phytophthora phaseoli*
P. parasitica	*Peronospora parasitica*
P. ramorum	*Phytophthora ramorum*
P. sojae	*Phytophthora sojae*
P. ultimum	*Phythium ultimum*
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate buffer saline
PcF	*Phytophthora cactorum*-Fragaria protein
PCR	Polymerase chain reaction
pEG	pEarleyGate
Pex	*Phytophthora* extracellular protein
pI	Isoelectric point
PIP	*Phytophthora* inhibited protease
pv.	Pathovar
PR	Pathogenesis-related
PRR	Pattern recognition receptor
PTI	PAMP-triggered immunity
PVX	Potato virus X
Q	Glutamine
R	Arginine
R-gene	Resistance gene
RB	Right border
Rif	Rifampicin

RLK	Receptor-like kinase
RLP	Receptor-like protein
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RP	Right border primer
rpm	Rotations per minute
RPP	Recognition of *Peronospora parasitica*
rRNA	Ribosomal RNA
RT-PCR	Reverse transcription-PCR
s	second
sq	square
SA	Salicylic acid
sdH$_2$O	Sterile distilled water
SDS	Sodium-dodecyl-sulphate
SDS-PAGE	Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis
SERK	Somatic embryogenesis receptor kinase
SIGnAL	Salk Institute genomic analysis laboratory
SMART	Simple modular architecture research tool
SOC	Super optimal broth with catabolite repression
SNP	Single nucleotide polymorphism
SP	Signal peptide
Spec	Spectinomycin
spp	Species
T-DNA	Transfer DNA
TAE	Tris-Acetate-EDTA

TBLASTN	Protein-nucleotide 6-frame translation
TEMED	Tetramethylethylenediamine
TLR	Toll-like receptor
UV-Vis	Ultra violet- visible
V	Valine or Voltage
v/v	Volume/volume
W	Tryptophan
Ws	Wassilewskija
w/v	Weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
X-gluc	5-bromo-4-chloro-3-indolyl-ß-D-glucuronide
Zeo	Zeocin
µg	Microgram
µl	Microliter
µM	Micromolar
µm	Micrometre
5-AS	5-aminosalicylic acid

1. *INTRODUCTION*

1.1. Global food security and sustainable crop production

According to United Nations World Food Programme, there are currently approximately 805 million people who do not have access to enough nourishing food (United Nations, http://www.wfp.org/hunger). This number will increase with the increasing population of the world that is expected to reach 9 billion by 2050 (Ronald, 2011). In return, this will add to the need of supplying healthy, reliable and sufficient food for everyone. Besides the role of international policies and collaborations, the situation requires a great boost in agricultural production. However, due to human-sourced reasons and environmental stress such as; pollutions, urbanisation, water shortage and global warming, the amount of land suitable for crop production is highly limited. In addition, an immense amount of product is lost to plant diseases caused by pests and plant pathogens, making healthy and sustainable crop production a major challenge (Boyd et al., 2012). Plant diseases not only decrease arable land area and lessen the crop yield but also cause major economic loss for the producers. Due to these serious socioeconomic impacts, there is an increasing global effort to help reduce the damages caused by plant diseases by focusing on researches on plant genetics and plant-pathogen interactions (Wulff et al., 2011). These studies play a critical role in coping with the diseases by

uncovering the underlying mechanisms thus finding more targeted, stronger and radical solutions like breeding disease-resistant crops, instead of using chemical control measures that greatly harms the environment.

1.2. Plant – pathogen interactions

Plant pathology (phytopathology) is an interdisciplinary area of study that investigates the plant pathogens and environmental conditions that are affecting the healthy plants by revealing the details of the disease mechanisms, which involves pathogen identification, disease cycles, comprised molecular actors, management methods and economic impacts (Agrios, 1997). The microbial pathogens that cause many devastating diseases include bacteria, viruses, nematodes, parasitic plants, fungi and oomycetes; all affecting plants from field to fork (Jones and Dangl, 2006).

1.3. Oomycete pathogens

Among above-mentioned phytopathogens, oomycetes form a unique, distinct group of eukaryotic microorganisms (Kamoun, 2003), which includes saprophytic members but more importantly a number of notorious plant pathogens that are known to cause some of the most ruinous, detrimental and socioeconomically important plant diseases on a wide spectrum of crops such as tomatoes, potatoes, lettuces, cucumbers, grapevines, hops, beets, Brussels sprouts, cabbages, cauliflower, kales, mustards, parsnips, turnips, watercress and many more (Margulis and Schwartz, 2000; Stassen and Van der Ackerveken, 2011). Oomycetes were classified under Fungi kingdom at early stages, as they showed resemblance with their filamentous morphology and

lifestyle such as developing similar structures during infection. Later on, further studies revealed that in fact, they highly differ from Fungi i.e. Oomycetes have diploid nuclei at their vegetative state and their cell wall mostly contains cellulose whereas Fungi are haploid or dikaryotic at vegetative state and their cell wall contains chitin (Erwin et al., 1983); and they differ with their metabolic pathways as well as their ultrastructures i.e. some oomycetes produce motile biflagellate zoospores (Judelson and Blanco, 2005; Hardham et al., 1994). Then, supporting studies suggested that they are biologically much closer to some photosynthetic microorganisms like Phaeophyceae (brown algae) and diatoms (Kwon-Chung, 1994; Slusarenko and Schlaich, 2003). Therefore the Oomycota (literally meaning egg fungi) class was taxonomically placed within the Stramenopiles (Heterokonta) phylum of the supergroup Chromalveolates (Keeling et al., 2005), which is a major line of eukaryotic microorganisms (Baldauf et al., 2000). The class Oomycota is divided into several orders one of which is the order Peronosporales that contains the families Albuginaceae, Peronosporaceae and Pythiaceae. These families include the necrotrophic pathogens (killing the host first, then feeding on the content) (Schmidt and Panstruga, 2011) such as *Phythium* species (e.g. *P. ultimum*, root rot) (Lévesque et al., 2010), hemibiotrophic pathogens (colonizes on living tissue for some time and then continues to live on dead tissue) such as *Phytophthora* species (e.g. *Phytophthora infestans*, late potato blight, causal agent of the Great Hunger or more commonly the Irish potato famine 1845- 1852; *Phytophthora sojae*, causal agent of soybean root and stem rot; *Phytophthora ramorum* causal agent of sudden oak death) (Haas et al., 2009; Raffaele et al.,

2010; Tyler et al., 2006), and obligate biotrophic species (needs the host alive, can't survive on dead tissue) (McDowell, 2011; Kemen and Jones, 2012) such as *Albugo candida* (white rust) (Kemen et al., 2011; Links et al., 2011) and narrow host range downy mildews (e.g. Lettuce downy mildew: *Bremia lactucae* and *Arabidopsis thaliana* downy mildew: *Hyaloperonospora arabidopsidis*) (Stassen et al., 2012; Holub, 2008).

1.3.1. Downy mildews

Downy mildew (downy: soft, fuzzy coating; mildew: fungal growth) is the common name for the obligate biotrophic oomycete pathogens. They are known to affect many plants including grapevines, lettuce, tomato, potato, brassicas, tobacco, onions, hops, cucurbit, carrots and basil disturbing the sustainable production, as well as ornamental plants, altering the appearance. At first glance the disease might be confused with powdery mildew or botrytis (grey mould), however with a closer look it can be understood that these are quite distinct. Downy mildews appear as discoloured -usually yellowish- oily spots on the leaf surface and white to bluish-white fluffy growth under the leaves. Sometimes this growth can be seen on the curds and buds of the plants as well (Slusarenko and Schlaich, 2003). This growth can lead to necrosis (death of the tissue) and even turn into a systemic infection, affecting the whole body. Downy mildews are mostly observed at humid (relative humidity 85% at leaf surface) and cool (15-23°C) environments, making early spring and late fall their favourite seasons. Although downy mildews have a narrow host range, they affect economically significant plants; therefore they require an effective control method. Altering the environmental conditions, changing the irrigation method,

providing a good air circulation and using chemicals are some ways to control the disease; however these usually are insufficient and environmentally unfriendly management methods. In order to achieve a much more efficient and radical control over downy mildews, a thorough understanding of the disease is required. Despite its attractive characteristics and economic importance, downy mildews did not receive the attention it deserved for a long time. However, with the improving technologies and genomic tools, the researchers started showing increasing awareness towards this unique group of pathogens. Still, since obligate biotrophs are difficult to work with as they require a living host to survive, the molecular methods are highly limited. For this reason, instead of higher crops, a model pathosystem has been widely used to understand the underlying molecular mechanism which is "the model plant *Arabidopsis thaliana* and its natural downy mildew pathogen *Hyaloperonospora arabidopsidis"* due to its many advantages.

1.4. *Arabidopsis thaliana-Hyaloperonospora arabidopsidis* pathosystem

1.4.1. *Arabidopsis thaliana as* a model plant

Arabidopsis thaliana has been used widely as a model plant for molecular and genetic researches due to its many useful features (Meinke et al., 1998). It is an annual, small angiosperm in the mustard family (Cruciferae or Brassicaceae) and can be found all around North America, Europe and Asia. Its life cycle can be as short as 6 weeks from seed to seed, depending on the environmental factors. The seeds can grow naturally, under fluorescent light or

in a greenhouse; on soil, hydroponic pots and petri dishes. Besides the ease with the rapid cultivation and its small size, its genome consists of only 5 chromosomes of 125 Megabase, containing 25.498 genes. A research community was formed, and as a result of a multinational collaborative work, *Arabidopsis* Genome Initiative (AGI) released the whole sequenced genome to public access in 2000 (AGI, 2000). This advance not only accelerated the researches on plant life but also the plant-pathogen interactions; as *Arabidopsis* is a natural host to many pathogens, one of which is the obligate biotrophic downy mildew: *Hyaloperonospora arabidopsidis (Hpa)*. The overall goal of the studies on this model system is to apply the findings on how the pathogen operates and how the plant defends itself to economically worthy crops starting with close relatives of *Arabidopsis* like Brassicaceae.

1.4.2. *Arabidopsis* accessions

Numerous different accessions (also referred as ecotypes) of *Arabidopsis* have been collected from over 750 natural populations and are named according to collection sites i.e. Col: Columbia, Oy: Oystese, Nd: Niederzenz. Wild types and mutant lines are named as Col-x; where x signifies the strain or the mutation in a given gene. Categorized lines are used for researches and can be obtained from widely-known stock centres: Nottingham *Arabidopsis* Stock Centre (NASC, Nottingham/UK) and the *Arabidopsis* Biological Resource Centre (ABRC, Ohio/USA). Landsberg *erecta* (Ler-0) and Columbia are frequently used ecotypes for research purposes, and were both used for the comparative analysis for the genome sequencing of *A. thaliana* (AGI, 2000). Another widely used background is Wassilewskija (Ws-0) and the mutant Ws-

*eds*1 (*enhanced disease susceptibility 1),* a mutation on *EDS1* gene that is required for the expression of quite a few naturally polymorphic resistance genes necessary for race-specific resistance to the *Hpa* (Parker et al., 1996).

1.4.3. Nomenclature of *H. arabidopsidis*

Hpa was formerly named as *Peronospora parasitica* and it was first changed to *Hyaloperonospora parasitica* after a study on the 5.8S rRNA gene and the internal transcribed spacer sequences (ITS1 and ITS2) in 2002 (Constantinescu and Fatehi, 2002). Further studies revealed that a *H. parasitica* pathogen is specific to Arabidopsis plants thus; its name was once again changed, also adding *–sidis* end (Gäumann, 1918), and became *Hyaloperonospora arabidopsidis* (Göker et al., 2004). There are various isolates of *Hpa*, and they are named according to the collection site of the isolate, forming the two prefix letters and the susceptible *Arabidopsis* accession, forming the two suffix letters i.e. Emoy: Em (East Malling) - Oy (Oystese), Edco: Ed (Edinburgh) - Co (Columbia), Hiks: Hi (Hilliers) - Ks (Keswick) (Holub, 2006; Holub, 2008). *Hpa*-Emoy2 is particularly important among these isolates, since its genome sequencing was completed and released in 2010 (Baxter et al., 2010)

1.4.4. *H. arabidopsidis* disease cycle

The life cycle of *Hpa* consists of both asexual and sexual reproduction (Fig. 1.1). When an asexual conidiospore of *Hpa* lands on the surface of a susceptible plant, it develops a short germ tube if necessary, and forms an appressorium when reaches over epidermal cell junction (Koch and Slusarenko,

1990). This is followed by formation of a penetration hypha which allows getting in between the walls of neighbouring epidermal cells. This hypha continues to branch out into the intercellular space until its tips reach to substomatal openings, but on the way it forms plentiful structures called haustoria (sing. haustorium) that invaginate into the host epidermal and mesophyll cells. Haustoria are pear-shaped pouches where the pathogen obtains nutrients and releases molecules to establish the pathogenicity, without breaking into the plant cells (Voegele and Mendgen, 2003; Catanzariti et al., 2007). The hyphal tips start advancing into early conidiophores within 1-2 weeks, which come out of stomata and spread out with branches carrying spherical asexual conidiospores (Koch and Slusarenko, 1990). These newly produced spores can start new rounds of infections.

On the other side, at the same time, sexual reproduction takes place in the cotyledon or the leaf, approximately a week after the landing of the conidiospores. Oogonia (female sexual organ) and antheridia (male sexual organ) are formed as a result of differentiated connecting hyphae. Oogonia contain oospheres that are fertilized via the antheridia by the use of a fertilization tube extending through its outer wall. The fertilized oosphere matures to become an oospore in the oogonium. Mature oospores endure the winter by staying dormant within the leaf debris in the soil and can restart the infection when favourable conditions reappear. In this case, the oospore germinates and enters the plant's root tips via a germ tube leading to formation of intercellulary expanding hyphae (Slusarenko and Schlaich, 2003).
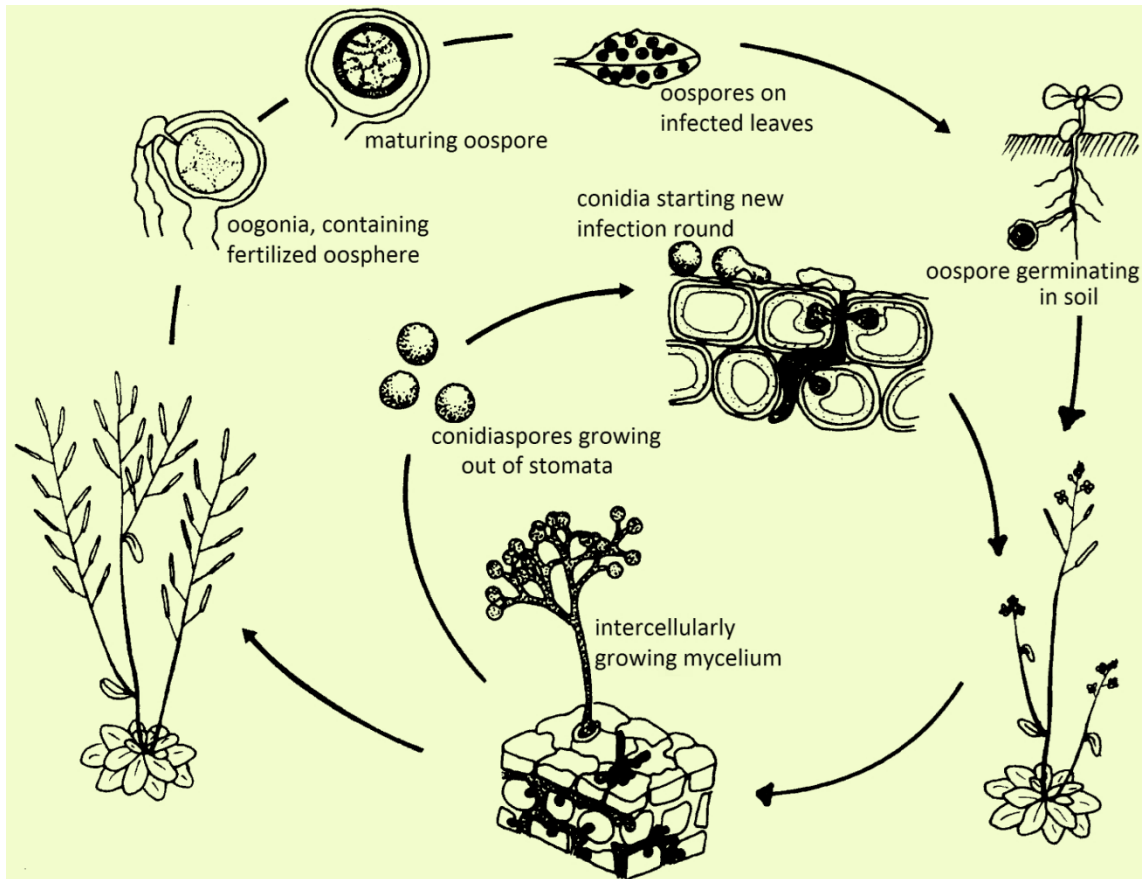
Figure 1.1. Life cycle of *Hpa*. (Slusarenko and Schlaich, 2003). Asexual reproduction: Infections initiated by oospores germinating in soil. Colonization happens through intercellularly growing mycelium. The hyphae forms haustoria and within 1-2 weeks, conidiophores containing asexual conidiospores emerge from stomata to initiate new rounds of infection. Sexual reproduction: About a week following the landing of the conidiospores, oogonia and antheridia are differentiated from connecting hyphae. Oospheres are fertilized by the antheridia. Fertilized oosphere becomes an oospore in the oogonium. Mature oospores can stay dormant in the soil, and re-initiate new cycles of infection. The diagram is not drawn to scale.

1.4.5. The intimate relationship

A common way of founding a relationship with the host for oomycetes and many fungal species is forming structures called haustoria, as mentioned before, which are appendages of the extending hyphae that are invaginated into the host cells while remaining surrounded by host-derived membrane. While

this extension increase the surface area of the pathogen that is in closer contact with the host, the haustorial cell wall and the plant plasma membrane are still not in direct contact due to the extrahaustorial matrix (EHM), the region stuck in between (Mims et al., 2004) (Fig. 1.2). This intimate interface serve as pathogen's site for taking up nutrients and water from the host and releasing molecules to interfere with the plant's defence mechanism and manipulate the cellular processes for a successful infection.
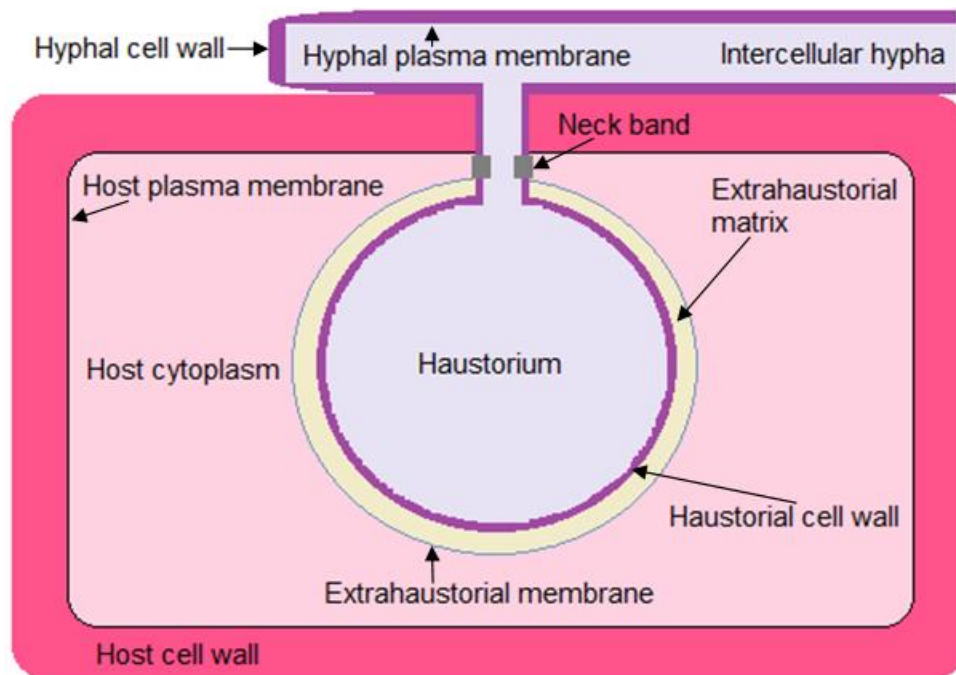
Figure 1.2. The haustorium. Intercellularly growing hypha invaginates into the host cell staying surrounded by a host-derived membrane. Extrahaustorial matrix fills the space between the haustorial cell wall and plant plasma membrane preventing a direct contact. EHM enables the pathogen's nutrient and water intake, and it is also the site where pathogen releases molecules into in order to establish infection. The figure is not drawn to scale.

1.5. Plant innate immunity

Plants are challenged by pathogens continuously (Faulkner and Robatzek, 2012). Despite being naturally resistant to majority of the diseases, they need to defend themselves against others (Lapin and van den Ackerveken, 2013). Since plants lack adaptive immune system and mobile defence cells, they have to rely solely on their innate immunity and systemic signals that are sent from the infection sites. They take advantage of their cuticles, the wax layer and their rigid cell walls (made of complex polysaccharides e.g. cellulose and pectin) and use them as a physical blockade against pathogen entry. The cuticle and the cell wall, depending on the age of the plant, can prevent the pathogen's penetration and the waxy structure on the leaf surface can inhibit reproduction of the pathogen by making the environment unfavourable. However, pathogens speedily evolve new strategies to overcome these obstacles, like applying mechanical pressure and secreting enzymes to degrade those barriers. Degraded plant cell walls can release phenolic and toxic molecules that are detrimental for the pathogens (Hématy et al., 2009). However, invading pathogens can usually overcome this counter-attack by penetrating the host cell and releasing molecules to establish the infection. In return, plants co-evolve with them and improve recognition methods for the molecules released by the damage as well as for the molecules released by the pathogens once they are inside the plant.

It is now accepted that plant immunity has two major divisions (Knoth and Eulgem, 2008; Staal and Dixelius, 2009). First is the recognition of Pathogen-associated Molecular Patterns (PAMPs) that initiates so-called PAMP-triggered

immunity (PTI). This stage of defence is usually overcome by effectors that are virulence molecules secreted by pathogens to enhance infection (suppressors of PTI), and recognition of effectors initiates the second perception process: Effector-triggered immunity (ETI). For plant immunity, Jones and Dangl (2006) suggested a zigzag model (Fig. 1.3) where the PAMP recognition initiates the PTI, and after that releasing of the effectors leads to Effector-triggered susceptibility (ETS). At that moment, the amplitude of the defence falls until the recognition of the effectors by resistance genes (*R*-genes), which starts the ETI and re-amplifies the defence, this time much faster and greater than the PTI. This can be followed by discharging of new set of effectors to avoid ETI, in return causing another moment of ETS, and they can be recognized again, causing ETI. This can go on depending on how advanced the pathogen and the host are in this continuous arms-race for subsistence (Dodds and Rathjen, 2010). According to this model, Jones and Dangl (2006) defines the basal defence (or basal disease resistance) of plants as a phenomenon during which the plants are recognizing non-self from self; activated by virulent pathogens on susceptible hosts. While at first it seems as basal defence should be PTI minus ETS, but in addition they include a weak ETI triggered by weak recognition of effectors, thus describing the term as PTI - ETS + weak ETI.
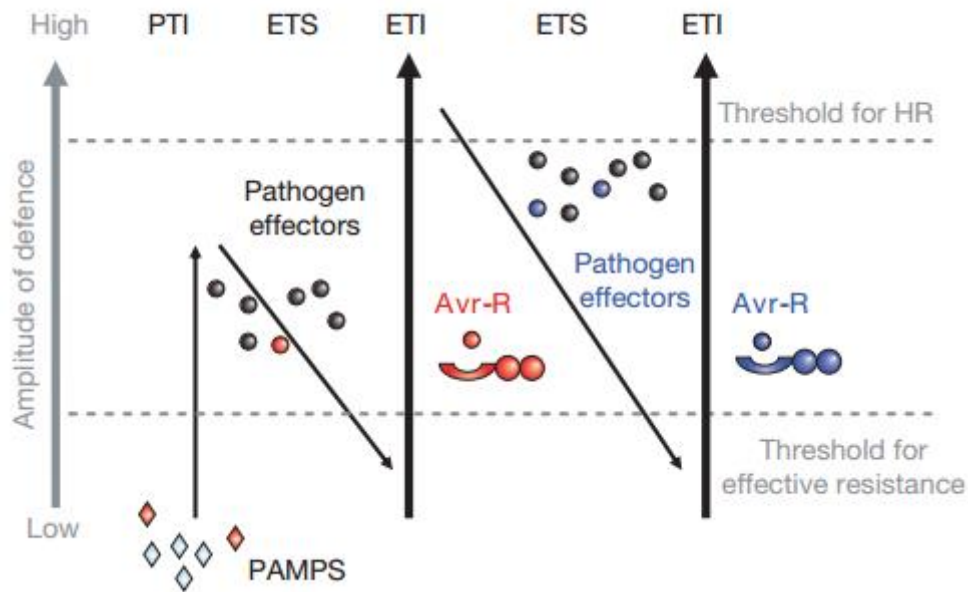
Figure 1.3. The *zigzag* model of the plant immune system (Jones and Dangl, 2006). Plants recognize Pathogen-associated Molecular Patterns (PAMPs) by Pattern Recognition Receptors (PRRs) which triggers PAMP-Triggered Immunity (PTI). Pathogens that are able to overcome PTI, secrete effectors which results in Effector-Triggered Susceptibility (ETS). When an effector is recognized by a receptor, Effector-Triggered Immunity (ETI) is activated, which is a faster and stronger response than PTI and often results in hypersensitive response (HR). ETI and ETS cycle can continue depending on how evolved both the pathogen and the plant are.

1.5.1. PAMP – triggered immunity (PTI)

The plants are capable of recognizing elicitors (molecules that induce defence responses) called PAMPs (or Microbe-associated molecular patterns, MAMPs as non-pathogenic microbes can also activate the immune system), which are conserved requisite molecules characteristic to microbe classes that can be peptides, oligosaccharides or lipids; e.g. bacterial flagellin, fungal chitin, oomycete β-glucan, Pep-13 and CBEL (Robatzek et al., 2006; Wan et al., 2008; Daxberger et al., 2007; Brunner et al., 2002; Larroque et al., 2011). There are

also Damage-associated Molecular Patterns (DAMPs), which are endogenous molecules (e.g. cell wall fragments) produced by pathogen attack. Being conserved makes PAMPs disabled for adaptive evolution and unable to escape recognition (Medzhitov, 2007). Therefore, they are perceived by receptors named as Pattern Recognition Receptors (PRRs) that reside in the plant plasma membrane and act as first line of defence against pathogen attack. This recognition at the cell surface leads to a set of changes in the plant collectively named as PAMP- triggered immunity (Jones and Dangl, 2006). In PTI, Cytosolic calcium (Ca^{++}) levels increase and this elevation mediate the production of Reactive Oxygen Species (ROS) i.e. hydrogen peroxide (H_2O_2) superoxide (O_2^-), and nitric oxide (NO), which creates a hostile environment for the pathogen and hardens the cell wall (Jones and Dangl, 2006). In addition, deposition of the polysaccharide 'callose', a β-(1,3)-glucan polymer found in cell wall appositions, between the plasma membrane and the cell wall is initiated in order to thicken the cell wall and obstruct the entry (Luna et al., 2011). Moreover, the stomata closes up and a signalling cascade is initiated (Mitogen-activated protein kinase, MAPK) in order to activate defence genes. Furthermore, production of plant hormones such as salicylic acid (SA), jasmonic acid (JA) and ethylene is also triggered to alert distant cells (Schwessinger and Zipfel, 2008; Tsuda and Katagiri, 2010; Muthamilarasan and Prasad, 2013). PTI can occasionally be sufficient for ceasing further colonization of the pathogen especially when the pathogen in question is a non-adapted one (non-host resistance); however, in most of the cases the pathogen puts the 'effectors' forward to surmount this resistance (Bozkurt et al., 2012).

1.5.1.1. Pattern recognition receptors (PRRs)

Both animals' and plants' immune systems make use of PRRs to sense danger and create the very first defence layer against pathogen attack. These receptors and the signalling pathways they are involved are quite similar in mammals, invertebrates and plants. For example, PRRs in plants show a noteworthy functional and structural resemblance to Drosophila TOLL and mammalian Toll-like receptors (TLR) (Dardick and Ronald, 2006).

Studied plant PRRs seem to be either transmembrane receptor-like kinases (RLKs) formed by an extracellular domain with Leucine-rich repeats (LRRs), a transmembrane domain and an internal kinase domain, or transmembrane receptor-like proteins (RLPs) that lacks internal signalling domain (replaced by a cytoplasmic tail with no enzymatic role) (Greeff et al., 2012; Monaghan and Zipfel, 2012, Tör et al., 2009).

In Arabidopsis, the receptor-like kinase family is a much bigger class (600 members) than receptor-like proteins (Kamoun, 2006). A famous receptor kinase Flagellin Sensing 2 (FLS2) of *A. thaliana* binds directly to bacterial flagellin and recognizes an N-terminal 22-amino acid-long motif in flagellin (epitope flg22) through its LRR domain and initiates PTI (Gómez-Gómez and Boller, 2000). Other widely known PRRs include the LRR-kinases EFR (recognizes the epitope elf18/26 of Elongation factor Tu, EF-Tu) (Zipfel et al., 2006) and rice Xa21 (recognizes a sulphonated protein 'Ax21' of the bacteria *Xanthomonas oryzae* pv. *oryzae* that enables a race-specific resistance to the *X. oryzae*) (Song et al., 1995; Lee et al., 2006). Besides the ones with LRR

domain, there are also PRRs that contain an N-terminal carbohydrate-binding domain (Lys-M). Chitin elicitor receptor kinase 1 (CERK1, recognizes fungal chitin in *A. thaliana*) and RLP CEBiP (recognizes chitin in rice) are examples for this category (Miya et al., 2007).

1.5.1.1.1. RD and non-RD kinases

For kinase receptors, regulation depends on the activation loop phosphorylation leading to a structural reorientation that allows substrate binding or makes phosphotransfer easier. Most of these kinases contain a conserved arginine (R) that precedes an aspartate (D) in subdomain VI that is essential for the catalytic activity, thus are called RD kinases (Johnson et al., 1996). The R residue is positively charged and is in the centre of positively charged residue cluster inhibiting catalysis due to the negatively charged D residue in the active site (Dardick and Ronald, 2006). When the activation loop which is close to the RD motif is phosphorylated, the inhibition is removed and negatively charged phosho-amino acids are produced, which neutralizes the R residue. This charge neutralization results in structural changes, facilitating phosphotransfer and activating the kinase (Krupa et al., 2004). These RD-kinases can be found in bacteria, fungi, plant and animals. On the other hand, there are kinases that lack this conserved R domain and are called non-RD kinases. In this case, the activation loop is not autophosphorylated; instead are essentially active or controlled via other mechanisms (Dardick et al., 2012, Krupa et al., 2004; Ronald and Shirasu, 2012). Majority of the plant RLKs known to date are non-RD kinases and most of them get help from other kinases for activation as well as RLPs.

1.5.1.1.2. Brassinosteroid insensitive 1- associated kinase 1 (*BAK1*)

Researches revealed that some PRRs interact with other LRR-receptor kinases one of which is the 'Brassinosteroid insensitive 1- associated kinase 1' (BAK1) in order to trigger PTI (Liebrand et al., 2014; Chinchilla et al., 2009). BAK1 is also known as SERK3 as it is also one of the five members of a somatic embryogenesis receptor kinase family. BAK1 does not perceive an elicitor on its own but interacts with other RLKs upon ligand binding and acts as co-receptor. For example after flg22 gets in contact with an elicitor, FLS2 forms a complex with BAK1, which results in a fast phosphorylation of both kinases, and initiates the intracellular signalling pathway (Chinchilla et al., 2007; Gust and Felix, 2014). Many other examples of such interaction entitle BAK1 as a key regulator of PTI. Therefore, in researches for defence activation, *BAK1* mutant plants are widely used (Li et al., 2002; Chung et al., 2012).

1.5.2. Effectors and Effector-triggered immunity (ETI)

In order to establish pathogenicity, adapted pathogens present weapons called effectors, to overcome PTI (Bozkurt et al., 2012). They alter and manipulate host cell processes to their own advantages by suppressing or delaying the responses and avoiding further recognition, thus advance the infection (Birch et al., 2008; Hogenhout et al., 2009; de Jonge et al., 2011). Unlike conserved PAMPs, effectors are highly diverse within and between species (Stam et al., 2014). It can be said that there is a co-evolutionary arms-

race between the pathogens effectors and plant immune receptors that results in a wide repertoire of rapidly evolving genes (Chisholm et al., 2006). When effectors are secreted and they start interfering with the PTI, it results in the vulnerability of the plant, called as effector-triggered susceptibility (Jones and Dangl, 2006). If the effector is able to avoid recognition this means the pathogen is virulent. In this case, the pathogen can continue the invasion by suppressing the defence responses via manipulating the signalling, and acting as or disrupting the host proteins (Jiang and Tyler, 2012). However, plants have evolved to detect majority of these effectors directly or indirectly by intracellular receptors, widely known as products of *Resistance* (*R*) genes, which belong to Nucleotide-binding LRR (NB-LRR) family (van der Hoorn and Kamoun, 2008; Bernoux et al., 2011) and the recognized effectors are now called as Avirulence (Avr) proteins (Jones and Dangl, 2006). This recognition causes a much more rapid and stronger response than PTI, defined as the Effector-triggered immunity (ETI), a term one may recall as the gene-for-gene resistance (Gassmann and Bhattacharjee, 2012), that can stop further pathogen growth. ETI frequently results in switching on new signalling pathways, cell wall and plasma membrane damage, modification of host proteins, DNA laddering, a much greater production of ROS, a local cell death at the infection site named as the hypersensitive response (HR) -a reaction similar to animal apoptosis- and disease resistance (Dodds et al., 2009; Collmer et al., 2000).

1.5.2.1. Oomycete effectors

Oomycete pathogens also secrete a wide range of complex effectors to manipulate the host defence mechanisms and establish infection and colonisation. As knowing the molecular mechanism of how the effectors disturb the host's life is extremely important to understand the disease mechanisms and develop control measures, there is an increasing interest towards effectors involved in the plant-oomycete interactions. With the improving technologies and availability of genomic and bioinformatic tools, substantial progress has been made in documenting oomycete effectors; yet there is still plenty to uncover.

It has been confirmed (Torto et al., 2003) that the effectors have to be secreted from the pathogen, a process which is directed by an N-terminal signal peptide (SP) that is targeting the effector protein to the endoplasmic reticulum to be processed and secreted out (Coates and Beynon, 2010). These secreted effectors are divided into two major groups regarding their target sites: apoplastic effectors are secreted to the plant extracellular space where they encounter with cell-surface receptors or other extracellular targets, and cytoplasmic effectors are translocated inside the plant cell (Stassen and Van den Ackerveken, 2011) (Fig.1.4).
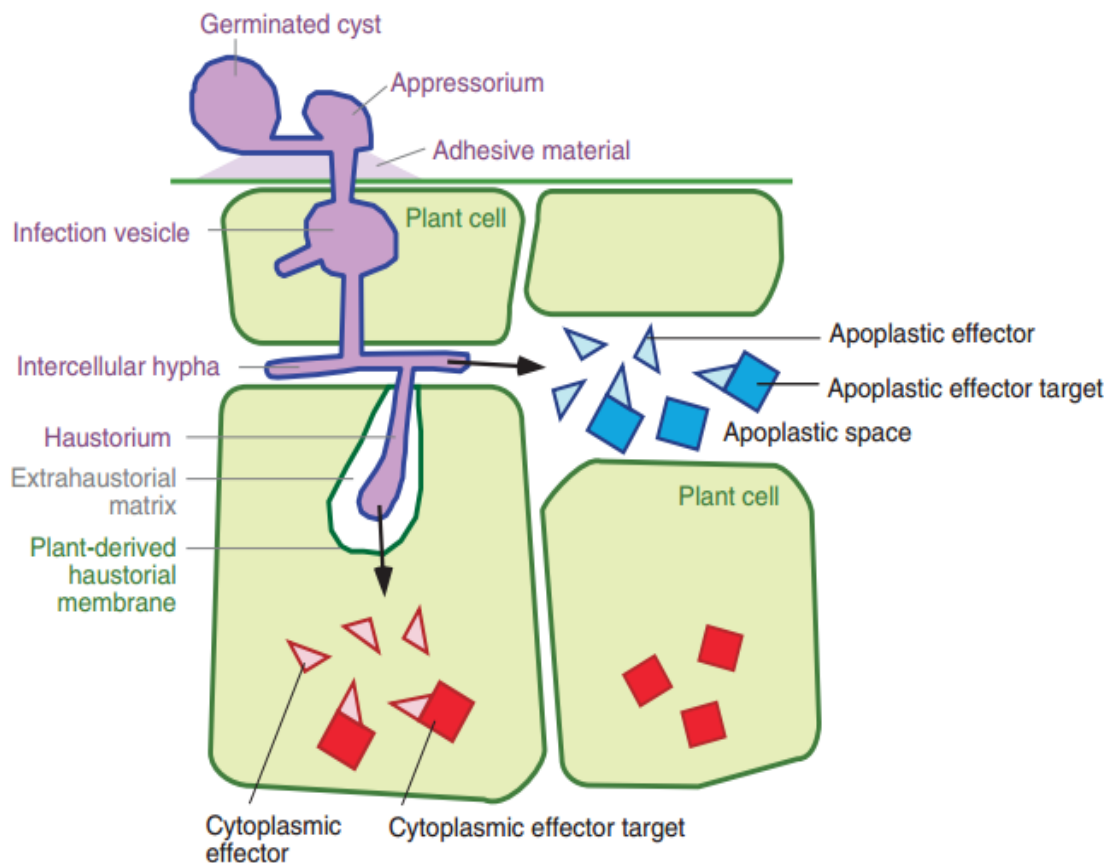
Figure 1.4. The secretion and recognition of apoplastic and cytoplasmic effectors (Kamoun, 2006). Oomycete apoplastic effectors (blue triangles) are secreted into apoplastic space where they interact with extracellular targets (blue squares) or surface receptors. Cytoplasmic effectors (red triangles) are translocated inside host cell where they interact with cytoplasmic targets (red squares). Plant structures are shown in green and pathogen structures are shown in purple. The figure is not drawn to scale.

1.5.2.1.1. Apoplastic oomycete effectors

Whilst occupying the plant extracellular space, the apoplast, oomycete pathogens secrete an array of effectors from their intercellulary extending hyphae. These effectors function in the host-pathogen interface by interacting with the defence related molecules thus, manage guarding against host defence

mechanisms, mediate invasion and contribute to immune responses. So far identified apoplastic oomycete effectors are grouped as either small cysteine rich proteins, enzyme inhibitors or NEP1-like proteins (Stassen and Van der Ackerveken, 2011). Apoplastic effectors are an attention attracting area of study as they play an important role in plant-pathogen interactions. However, not enough known about them since not having conserved motifs in their structures and the possibility of them being secreted into the apoplast but triggering intracellular reactions, make their identification and traceability relatively more difficult.

1.5.2.1.1.1. Small cysteine-rich proteins

It has been shown that some secreted proteins, which are inducing host defence responses when infiltrated into plants, share the feature of containing an even number of cysteine residues and being smaller than 150 amino acid-long. Many avirulence proteins belong to this group. For instance, the causal agent of barley and rye scald - ascomycete *Rhynchosporium secalis*- contains the gene *nip1* and the causal agent of tomato leaf mould –fungus *Cladosporium fulvum*- has the genes *Avr2*, *Avr4*, and *Avr9* that are encoding these secreted proteins (Laugé and De Wit, 1998; van't Slot and Knogge, 2002). The importance of these proteins is that the cysteine pairs form disulphide bridges which provide strength and stability against the proteases in the host apoplast while inducing defence responses (Kamoun et al., 1999a; Luderer et al., 2002). For oomycetes, some elicitins of *Phytophthora* species, which are extracellular proteins inducing defence responses including HR in *Nicotiana* species, are

small cysteine rich proteins (Ricci et al., 1989; Kamoun et al., 1997a; Ponchet et al., 1999; Sasabe et al., 2000). For instance, Kamoun et al. (1998) identified INF1 of *P. infestans* by showing its avirulent functions on *N. benthamiana*. In addition to INF1, many more elicitin-like proteins were identified in *P. infestans* and apparently the genes encoding these proteins all contained a 98-amino acid long elicitin domain with a core of 6 conserved cysteines. Moreover, several *inf* genes contained both N-terminal and C-terminal elicitin domains. Amino acids proline, alanine, threonine and serine were found recurrently within this C-terminal, and due to the features, numbers and orientation of these amino acids led to formation of *O*-linked glycosylation site bundles (Kamoun et al., 1997b) with a tail shaped extension. This structure at the C-terminal can bind the protein to the cell wall while the N-terminal remains open to interactions with the plant-sourced molecules suggesting that these INF proteins might have a role in the infection process by being cell surface or cell wall associated glycoproteins.

Further investigation revealed that class I elicitins were able to bind sterols e.g. ergosterol, and this was critical since *Phytophthora* species obtain the sterols from other sources as they are unable to make their own (Boissy et al., 1999; Mikes et al, 1997; Mikes et al., 1998; Vauthrin et al., 1999). Detailed studies (Nespoulous et al., 1999; Osman et al., 2001a) showed that elicitin family protein may also have a role in lipid interactions and metabolism since elicitin-like proteins of *P. capsici* that were showing notable resemblance to INF5 and INF6, were acting as phospholipases. Moreover, additional work

suggested the elicitin-sterol interactions were vital for binding to a plasma membrane receptor and triggering HR (Osman et al., 2001b).

On the other hand, another elicitin protein called PcF (*Phytophthora cactorum*-Fragaria protein, a phytotoxin) of the PcF toxin family from *P. cactorum* with 6 cysteine residues forming disulfide pattern was shown to induce weakening, HR and necrosis on strawberry and tomato leaves (Nicastro et al., 2009). PcF showed homology to an allergenic plant pollen protein (Ole-e6) and it was also proposed that PcF could be acting as plant signalling protein. Other proteins quite similar to this toxin family were also identified from *P. infestans*. One of which, is the secreted cysteine-rich SCR74, was shown to have an up-regulated expression 60-fold at the early days of infection (Liu et al., 2005).

Lastly, Ppat12, 14, 23 and 24 are secreted cysteine-rich proteins, found in *Hpa* that are exceedingly expressed during infection; however their exact function remains unclear besides the fact that they are thought to have a role in biotrophy (Bittner-Eddy et al., 2003; Baxter et al., 2010).

1.5.2.1.1.2. Enzyme inhibitors

Hydrolytic enzymes such as proteases, chitinases and glucanases are Pathogenesis-related (PR) proteins that plants make use of against pathogens (Boller and Métraux, 1988; Mauch and Staehelin, 1989; Abramovitch and Martin, 2004; Punja 2004). In return, fungal, bacterial and oomycete pathogens have advanced defence mechanisms by producing inhibitors of these enzymes (Misas-Villamil and van der Hoorn, 2008; Peresen et al., 2012). Oomycetes are

naturally resistant against plant chitinases as they have a tiny amount of chitin in their cell wall, yet they have to defend themselves against proteases and glucanases since their cell wall contains β-1,3/1,6 glucans (Erwin et al, 1983). For instance, *P. sojae* has evolved GIP1 and GIP2 glucanase inhibitors against endo-β-1,3 glucanase EGaseA of soybean, which show structural resemblance to serine proteases (trypsin class) while not being involved in proteolysis due to mutations in their catalytic parts (Rose et al., 2002). These GIPs appear to be involved in preventing the β-1,3/1,6 glucan degradation in the cell wall. In addition, they also inhibit the deposition of defence activating oligosaccharides by endoglucanases. While GIP1 is inhibiting the EGaseA, it has no activity over EGaseB -another endoglucanase produced by soybeans-, which is explained by positive selection on β-1,3 endoglucanases of soybean triggered by co-evolving glucanase inhibitors of *P. sojae* (Bishop et al., 2004).

On the other hand, against plant proteases, oomycete pathogens have evolved serine and cysteine protease inhibitors. An extracellular protease inhibitor, EPI1 has been characterized in *P. infestans* that is quite similar to Kazal family of serine protease inhibitors with its two domains (Tian et al., 2004). Studies showed that EPI1 inhibits a serine protease subtilisin A, and additionally inhibits and interacts with PR protein P69B subtilisin-like serine protease found in tomato apoplast (Tian and Kamoun, 2005; Tian et al., 2004). EPI1 and P69B are expressed and up-regulated in accordance, thus forming another example of plant defence – pathogen counter defence, as well as for the similarity between virulence acts of oomycete pathogens and mammalian parasites (apicomplexan *Toxoplasma gondii and Neospora caninum*) (Lindh et

al., 2001; Bruno et al., 2004). Another defined serine protease inhibitor is EPI10, which functions just like EPI1 with its three Kazal-like domains (Tian et al., 2005).

Finally, Tian et al (2007) also described secreted proteins EPIC1 to EPIC4 from *P. infestans* similar to cystatin-like protease inhibitor domains. Among the genes encoding these proteins, rapidly evolving epiC1 and epiC2 genes did not have orthologs in *P. ramorum* and *P. sojae*; and were shown to be up-regulated through the infection of tomato plants, proposing an involvement in plant-pathogen interactions. In addition, EPIC2B apparently targets papain-like extracellular cysteine protease (*Phytophthora* Inhibited Protease 1, PIP1), which is a PR protein similar to a tomato apoplastic cysteine protease involved in resistance (Rcr3), and another protease C14, in addition to other tomato apoplastic proteases (Song et al., 2009; Wawra et al., 2012).

1.5.2.1.1.3. Nep1-like proteins (NLP)

A broad range of plant-associated microbes such as bacteria, ascomycetes, basidiomycetes, and oomycetes are known to secrete a diverse range of necrosis and ethylene-inducing peptide 1 (Nep1)-like proteins (Gizjen and Nürnberger, 2006). In spite of the broad diversity, the members of this family show great similarity in sequences. These canonical 24-25 kDa proteins function as cytotoxins in dicotyledonous plants, causing tissue necrosis and activation of defence responses. The conserved nature among phylogenies and their crucial functions make them a distinct family of proteins. These proteins also displayed structural homology to pore-forming toxins of sea anemones

(Ottmann et al., 2009). Based on this and the *in planta* assays, it was advocated that NLPs have a role in plasma membrane disruption and cytolysis (Ottmann et al., 2009). NLPs seem to need a specific target site on the extracytoplasmic side of the plant cell membrane (Qutob et al., 2006) and members with the role of cell death induction are widely found in *Phytophthora* and *Pythium* species (Pemberton and Salmond, 2004). However, identified NLPs from *Hpa* (HaNLPs) appeared to be non-cytotoxic as they were not causing necrosis due to suppressed necrosis-inducing activity with a surface-exposed region, but only having a role in inducing immune responses in *A. thaliana*. To help determine the exact roles of HaNLPs, some were ectopically expressed in the host which led to defence activation, resulting in growth reduction that meant induced immune responses. Further investigations suggested that these proteins might be acting as MAMPs (Cabral et al., 2012).

In addition to all extracellular effector proteins, there are also proteins with an Arginine-Glycine-Aspartate (RGD) tripeptide that is known to alter the adhesion between the cell wall and the plasma membrane, thereby disturb the integrity (Schindler et al., 1989; Canut et al., 1998). An example of these proteins is the IPI-O1 of *P. infestans* that interacts with *A. thaliana* lectin receptor kinase (LecRK-I.9) through its RGD domain, thus interfering with the cell wall – plasma membrane integrity (Senchou et al., 2004; Gouget et al., 2006). Further studies also revealed that LecRK-I.9 has a role in *A. thaliana*'s resistance against *P. brassicae* (Bouwmeester et al., 2011).

1.5.2.1.2. Cytoplasmic oomycete effectors

Studies discovered that many effectors appear to find their subcellular targets and act within the host cell (Ellis and Dodds, 2011). They are usually unravelled through their avirulent characters which means, when they get in contact with the resistance genes and induce defence responses such as HR on the host. So far identified cytoplasmic effectors were shown to have conserved motifs following the signal peptide that enables their translocation into the host cell (Grouffaud et al., 2010). According to these motifs, cytoplasmic effectors are grouped as RxLR (Arginine-x-Leucine-Arginine) effector family, Crinkler (CRN) effector family and CHxC effector family (Kale and Tyler, 2011; Kemen et al., 2011; Pais et al., 2013).

1.5.2.1.2.1. RxLR effector family

For avirulence proteins, it has been revealed (Rehmany et al., 2005) that after the signal peptide, there is a highly conserved motif within the first 40 amino acid in their sequences 'RxLR' and this is frequently chased by less conserved acidic motif 'EER' (Glutamate-Glutamate-Arginine) within 30 amino acids towards the C-terminal. The discovery of existence and positional specificity of these motifs have helped developing bioinformatic methods to predict and identify potential cytoplasmic effectors. Soon after, it was discovered that these motifs were shown to be actually required for host cell translocation, thus made cytoplasmic RxLR effectors a major study topic for researchers (Whisson et al., 2007; Dou et al., 2008). RxLR motif also shows resemblance to the RxLxE/D/Q (PEXEL) motif that is required for host-targeting

mechanism of proteins from malaria parasites (*Plasmodium spp.*), which can be another proof of the similarity between plant and animal eukaryotic pathogens for effector delivery systems (Hiller et al., 2004; Haldar et al., 2006).

A number of *R*-genes that act against oomycetes of *Phytophthora spp.* and *H. arabidopsidis* that mostly belong to NB-LRR family have been cloned as well as race-specific *Avr* genes. For instance following avirulence on potato *R*-genes (*R3a*), *Avr3a* gene of *P. infestans* was identified from candidate effectors (Armstrong et al., 2005), by showing that the R3a-mediated HR was triggered through interaction with Avr3a allele.

Another RxLR effector, Avr3b of *P. sojae* was also identified, and found to carry a Nudix hydrolase (Nucleoside Diphosphate linked to X, a phosphohydrolase) domain which was suppressing immune responses by acting as Nudix hydrolases (Dong et al., 2011).

On the other side, the soybeans carrying the *R-gene Rps1b* led to the discovery of *Avr1b-1* and *Avr1b-2* of *P. sojae* through avirulence (Shan et al., 2004).

Furthermore, many diverse *A. thaliana R-genes* against *Hpa* (Recognition of *Peronospora parasitica*, *RPP*) were identified and cloned. For instance, three closely related *RPP1-WsA*, *RPP1-WsB* and *RPP-WsC* were cloned from Ws-0 accession, each of which provided resistance against different *Hpa* isolates (Botella et al., 1998). For Nd accession, a locus containing *RPP1-Nd* is responsible for resistance and its matching *Avr* gene is the highly polymorphic *ATR1^{NdWsB}* (*Arabidopsis thaliana Recognized1*) that

encodes a 311 amino acid-long protein triggering HR. Interestingly, ATR1[NdWsB] did not show many polymorphisms in the signal peptide and RxLR region, C-terminal ends were subjected to positive selection proposing that this part of the protein is more under pressure for evolving (Win et al., 2008). This ATR1[NdWsB] was cloned on Nd-0, using a population created through crossing between isolates *Hpa*-Maks9 (virulent) and *Hpa*-Emoy2 (avirulent) (Rehmany et al., 2003; Rehmany et al., 2005). It was shown that, Maks9 allele of ATR1[NdWsB] was recognized solely by RPP1-WsB, and Emoy2 allele was both recognized by RPP1-Nd and RPP1-WsB, whereas Cala2 allele of ATR1[NdWsB] was recognized by neither of them (Rehmany et al., 2005).

Another *Avr* gene *ATR13*, encoding a 187 amino acid-long protein, was characterized for inducing HR by interacting with RPP13 of *A. thaliana* (Allen et al., 2004). For instance, accession Col-0 showed susceptibility for *Hpa*-Maks9 isolate, whereas Nd-0 was resistant due to RPP13-Nd (Bittner-Eddy et al., 1999). When Col-5 (glabrous Col-0) was transformed with a clone containing *RPP13-Nd*, the plants became resistant to Maks9 isolate, thus supporting the previous findings (Bittner-Eddy et al., 2000). In addition, both Col-5 and Col-5[RPP13-Nd] plants were subjected to gene bombardment with the cloned *ATR13* from different *Hpa* isolates, which revealed numerous alleles of ATR13 (Allen et al., 2004). Moreover, sequential analysis of the avirulent and virulent alleles showed that ATR13 contains a heptad leucine/isoleucine repeat motif in addition to the signal peptide and the RxLR motif, and it shows great number of polymorphisms in consistence with the polymorphisms of the *RPP13* locus in different accessions, pointing at diversifying selection (Allen et al., 2004; Rose

et al., 2004). Looking at the functionality of different alleles of both ATR1 and ATR13 from different *Hpa* isolates, despite the polymorphisms, supports the fact that effector activity is managed towards the C-terminal regions of these proteins.

Last but not least, another NB-LRR family protein RPP5, defines *Hpa* resistance to *A. thaliana*. The *Avr* protein ATR5, cloned from Emoy2, is shown to induce immune responses in the plants expressing the functional RPP5 allele of Ler-0. Interestingly, while this ATR5 is containing a signal peptide, a canonical EER motif and an RGD motif, it lacks the RxLR motif suggesting that ATR5 is a variant of the canonical RxLR effectors (Bailey et al., 2011).

1.5.2.1.2.2. Crinkler (CRN) effector family

Another class of cytoplasmic effectors from candidate secreted proteins of *P. infestans* were defined through plants that were infected with recombinant Potato virus X (PVX). Torto et al. (2003) developed an algorithm 'the PexFinder' (*Phytophthora ex*tracellular protein) that identifies extracellular proteins from complementary DNA (cDNA) sub-sequence data sets. After the data mining the candidates were tested with *in planta* virus-based high-throughput functional expression assays performed using an *Agrobacterium* tumefaciens binary vector carrying the PVX. Among those, eight of the clones caused necrotic appearances at the inoculation points and additionally, both necrosis and crinkling on leaves all around the plants were observed. When the partial sequences and restriction fragments of these clones were analysed, they pointed towards two cDNAs: Two clones had crn1 and six clones had crn2

named after crinkling and necrosis (Torto et al., 2003), both belonging to a big compound protein family in *Phytophthora*. Assays with both PVX::crn1 and PVX::crn2 showed that they induce implications of necrosis in *N. benthamiana*. Additional assays also specified that throughout the infection of host tomato plants, both *crn* genes are expressed in *P. infestans* and *crn2* is responsible for inducing immune responses. Therefore, the phenotypic signatures of leaf crinkling, HR and necrosis were tied to the expression of both genes and their induction of defensive reactions. Following these findings, sequences of both cDNAs were analysed. According to BLASTP searches, *Crn1* was encoding a 431 amino acid long protein and *crn2* was encoding a 456 amino acid long protein; and neither showed any resemblance to proteins from other organism or contain any of the known conserved motifs. However, they were similar to each other and TBLASTN searches pointed at many more CRN-like proteins in *Phytophthora*. When these proteins from *P. infestans, P. sojae,* and *P. medicaginis* were aligned, homology at the N-terminal region was observed. Afterwards, CRNs were also found in other oomycete pathogens such as; *P. ultimum* (Lévesque et al., 2010), *P. phaseoli* (Kunjeti et al., 2012), *P. ramorum* (Haas et al., 2009), *H. arabidopsidis* (Baxter et al., 2010) and *B. lactucae* (Stassen et al., 2012).

Like RxLR effectors, CRN effectors also contain a highly conserved motif in the amino acid sequences, the 'LxLFLAK domain' (Leucine-x-Leucine-Phenylalanine-Leucine-Arginine-Lysine) of nearly 50 amino acids at the N-terminal. In addition, a motif 'Aspartate-Tryptophan-Leucine' (DWL) spots the beginning of an assorted C-terminal region ending with a highly conserved

domain consisting of Histidine-Valine-Leucine-Valine-X-X-Proline (HVLVxxP) (Haas et al., 2009).

In order to dig through the functionality of these effectors, Schornack et al. (2010) designed constructs using the N-terminal of CRN2 and CRN16 of *P. infestans* with the C-terminal of the of AVR3a. The constructs were cloned into *P. capsici,* and the expression of the construct triggered ETI in *N. benthamiana* leaves resulting in transgenic *P. capsici*'s avirulence. On the other hand, no avirulence was observed when LxLFLAK motif was mutated to LxAAAA and fused with AVR3a's C-terminal. Therefore, it was suggested that N-terminal regions of CRN family proteins are crucial for translocation due to the LxLFLAK motif. Moreover, localisation experiments done with *in planta* expressions of CRN-GFP fusions (*P. infestans* CRN2, CRN8, CRN15, CRN16 and *Aphanomyces euteiches* CRN5) showed that the constructs targeted the plant cell nucleus after translocation. It was understood that the direction followed depended on a host protein involved in nuclear trafficking machinery and binding the nuclear localisation signal motifs (NLS): importin-a (Schornack et al., 2010).

A conclusive note hasn't been put on the timing of the expression of the CRN effectors. While some CRNs were expressed by *P. infestans* throughout the infection, for *P. sojae*, signs of expressions were seen more in the later stages of infection (Haas et al., 2009; Ye et al., 2011). Finally, a considerably larger number of CRN effectors are secreted by oomycetes that do not form haustoria with respect to RxLR effectors (Wawra et al., 2012).

1.5.2.1.2.3. CHxC effector family

Kemen et al. (2011) revealed another class of cytoplasmic effectors formed by defence-suppressing effectors of *Albugo laibachii* (causes white rust of *A. thaliana*) that share a CHxC (Cysteine-Histidine-x-Cysteine) motif located within 50 amino acids of the signal peptide cleavage site. Due to adoption of biotrophic and hemibiotrophic lifestyles, some oomycetes went through a gene gain and loss process and in the case of *Albugo species;* they are thought to have evolved these CHxC effectors (Kemen et al., 2011). Moreover, while *A. candida* has a significantly smaller range of PR-proteins including elicitins, CRN and RxLR effectors in comparison with *Hpa* and no reported NLPs, 40 CHxC proteins were detected in *A. candida*, with no homology in *Phytophthora* or *Hyaloperonospora* species (Links et al., 2011). It was also shown that the CHxC motif is required for translocation; however, these effectors require supporting investigations (Kemen et al., 2011; Kale, 2012).

1.6. Identifying novel effectors

With the technological breakthroughs made especially in the last decade, substantial amount of data has been gathered regarding plant-pathogen interactions. Even though a significant number of researchers are motivated and have collaborated to shed light on this subject; there are many plant receptors, pathogen molecules and pathways waiting to be discovered.

Using the available sequence data, it is possible to generate and mine through Expressed Sequence Tag (EST) libraries, which are short cDNA sub-sequences resulting from one-shot sequencing of cloned cDNA that can be

used for detecting gene transcripts, the proteins they encode and the functions thereof (Krajaejun et al., 2011; Cabral et al., 2011; Nagaraj et al., 2006). In addition, it is possible to investigate genetic diversity of those genes within and between species. This method can also be followed to discover new effector genes. Candidate sequences can be determined applying certain criteria to the available data sets and processing through bioinformatic tools, programmes and databases (e.g. SignalP, PexFinder, SMART, NCBI, Geneious, Dnastar), and then those candidates can be challenged by high-throughput *in vitro* and *in planta* expression assays combined with proteomic tools for structural and functional analysis, and can be viewed for their effector-like activities in a top-down manner. Moreover, a bottom-up method can also be followed by firstly extracting the proteins from the source, analyzing their activities via functional assays, and then scrutinizing their structures.

Based on the fact that oomycetes are the lead actors of many detrimental plant diseases; getting a deeper insight into the molecular interactions for developing environmentally friendly, targeted, consistent and permanent control measures is extremely important. With this inspiration, using the model pathosystem of *Arabidopsis* and its downy mildew pathogen *Hpa*, we pursued the above-mentioned methods in order to identify novel apoplastic effectors, considering the huge gap about apoplastic effectors in the existing knowledge on this subject.

AIMS and OBJECTIVES

The aim of this study was to investigate apoplastic effectors from the *A. thaliana* downy mildew pathogen*, H. arabidopsidis.* This was achieved by following these objectives:

- Identifying putative apoplastic effectors from *Hpa* using bioinformatics,

- Validating and analysing the time and the intensity of the candidate genes' expression during infection,

- Determining whether the candidate gene sequences are conserved across isolates or they show polymorphisms,

- Investigating the candidate genes' capability of triggering host immune responses via *in vitro* and *in planta* gene expression,

- Collecting the apoplastic fluid of healthy and infected plants, and analysing the protein content via MALDI-TOF,

- Screening whether the healthy and infected apoplastic fluids are able to induce defence responses using GUS reporter system and measuring ROS production,

- Simplifying the complex apoplastic fluids and identifying the active fractions and,

- Observing interaction phenotypes of *A. thaliana* plants deficient in apoplastic LRR proteins selected from the MALDI-TOF results.

.2. MATERIALS and METHODS

2.1. MATERIALS

2.1.1. Chemicals, enzymes, kits and other consumables

Unless otherwise indicated all the chemicals, enzymes, kits and other consumables were mainly purchased from Invitrogen™ Life Technologies (Paisley, UK), Sigma-Aldrich® (Gillingham, UK), Bioline Reagents Limited (London, UK). London, UK), New England Biolabs (NEB, Hitchin, UK), Qiagen GmbH(Hilden, Germany), Roche Products Limited (Welwyn Hatfield, UK), Melford (Ipswich, UK), Merck Chemicals (Nottingham, UK), and VWR Chemicals International Ltd (Lutterworth, UK).

2.1.2. Plant materials

Arabidopsis thaliana accessions (wild type and mutants) Ws-*eds*1 (Parker et al., 1996),Col-0, Col-5, Col-*rpp4* (van der Biezen et al., 2002), Col-*bak1-5/bkk1-1* (Roux et al., 2011), Landsberg *erecta* (Ler-0), Col-0 carrying RLK-GUS or PR-GUS constructs were available in the laboratory stocks and were used throughout the study as well as *Nicotiana benthamiana* and *Nicotiana tabacum*. T-DNA insertional mutant lines or any other mutant lines were ordered from Nottingham Arabidopsis Stock Centre (NASC - Nottingham, UK) when necessary.

2.1.3. Pathogen isolates

All *H. arabidopsidis* isolates (Emoy2, Cala2, Noks1, Emco5, Edco1, Goco1, Maks9, Hiks1, and Emwa1) used in this study were readily available in the -80ºC stocks of the laboratory.

2.1.4. Bacterial isolates

Erwinia amylovora, used mostly as positive control, was available in the -80ºC stocks of the laboratory. Electrocompetent (ElectroSHOX) and chemically competent (BL21 [DE3] pLysE) *Escherichia coli* (*E. coli*) cells were purchased from Bioline (London, UK) for transformation purposes. Electrocompetent *Agrobacterium tumefaciens* (strain GV3101) -used for stable and transient transformation of the plants- was also available in the laboratory and new electrocompetent cells were prepared in the laboratory when required.

2.1.5. Vectors

Vectors pCR8™/GW/TOPO®, pDONR™/Zeo, pEarleyGate100 (pEG100) (Earley et al., 2006), pET28a, pET29-GW, pET32-GW, pDest15™ pDest17™, and pEXP1-DEST used for cloning, sub-cloning, sequencing and expression purposes were either available in the laboratory stocks or purchased from Invitrogen™ Life Technologies. The vectors pFLAG-ATS™ and the clone pFLAG-ATS-EPIC used for expression trials were kindly provided by Prof Sophien Kamoun's group (The Sainsbury Laboratory, Norwich, UK).

2.2. METHODS

2.2.1. Plant growth conditions

All *Arabidopsis thaliana* accessions and *Nicotiana* spp (*N. benthamiana,* *N. tabacum*) seeds were sown on soil-filled (Levington F2+S Seed and modular compost + sand) and tap water-soaked modular trays (40 modules, 4 cm x 4 cm x 3 cm). Approximately 30 seeds were dispersed on each module and sprayed with tap water and the trays were covered with a suitable transparent plastic lid, allowing air circulation. Trays were incubated at 20°C with 14 h dark-10 h light photoperiod (fluorescent light OSRAM L36W/840 - OSRAM Lumilux cool white).

2.2.2. Seed production, threshing and storage

For seed production, 4-5 week old plants were transplanted into big pots (8 cm x 8 cm x 8 cm) individually and were grown covered with bread bags (with holes allowing air circulation) until majority of the siliques were brownish. The plants were then harvested and kept in paper bags and were allowed to dry for couple of weeks. Seeds were sifted several times until the plant debris was completely removed. Seeds were collected in 1.5 ml Eppendorf tubes and stored at -20°C for further use.

2.2.3. Inoculation of the plants with the pathogen and maintaining the infection cycle

All isolates were maintained in Ws-*eds1* (Parker et al., 1996), and 7-day old seedlings were inoculated to produce large quantities of inoculums. To

initiate the infection cycle, frozen stocks of the isolates were used and for further cycling, spores were collected from the previously infected material. Infected material (cotyledons covered with sporangiophores) was put in chilled sterile distilled H_2O (sdH_2O) and were vortexed to release the spores and the suspension was filtered through Miracloth (Calbiochem®, Merck KGaA, Darmstadt, Germany) to remove the contaminants. A damp tissue paper was placed in the tray to provide humidity and the healthy seedlings were then spray-infected with the spore suspensions (concentration of $5x10^4$ spores/ml; measured with a haemocytometer -Labor Optik Ltd., Lancing, UK). The trays were covered with transparent lids and were sealed with tape to maintain the high humidity. Infected plants were kept at 16^oC with 14 h dark-10 h light photoperiod (fluorescent light L36W/840 – OSRAM Lumilux cool white). Same method was followed for other accessions and isolates as well.

2.2.4. Isolation of plant DNA and RNA

Qiagen's DNeasy Plant Mini Kit, RNeasy Plant Mini Kit were used to isolate DNA and RNA, respectively, from both infected and healthy materials, following manufacturer's instructions. Grinding the plant samples in liquid N_2 was preferred. Isolated DNA samples were kept at 4^oC and RNA samples were stored at -20^oC till further use.

2.2.5. Elimination of DNA contaminants from RNA samples

Ambion's Turbo DNase I enzyme was used to remove DNA from RNA samples following manufacturer's instructions. Concentration adjustments were made using RNase-free water (Qiagen) whenever necessary. The enzymatic

reaction was cleaned up using Qiagen's RNeasy Plant Mini Kit following the instructions for RNA Clean-up.

2.2.6. Purification of plasmid DNA

The bacterial culture containing the desired plasmid was streaked on an LB plate (selective) and was incubated overnight. Next day a single colony was picked and inoculated into 5 ml LB (selective) and this was incubated overnight with 200 rpm agitation. Subsequently, plasmid DNA was isolated from samples using Qiagen's Qiaprep Spin Miniprep Kit, and the instructions enclosed were followed. Extracted plasmids were stored at -20°C.

2.2.7. Primer design and storage

Mostly, primers were designed manually using Geneious (R6) (Biomatters, Kearse et al., 2012), regarding the melting and annealing temperatures, GC content, length (both the primer and the template's), and possibility of forming hairpins, self and cross dimers (Table 2.1a-d). For T-DNA insertional mutant lines, primer sequences were obtained using the web site of Salk Institute (http://signal.salk.edu/tdnaprimers.2.html). All primers were purchased from Sigma-Aldrich® (Gillingham, UK). The concentrations of the primers were adjusted to 50 µM or 100 µM with sterile chilled dH$_2$O and were diluted to 10 µM before use. The stocks were stored at -20°C for further use.

Table 2.1a. Gene-specific forward and reverse flanking primers designed for candidate genes

Gene ID	Primers
804480	Forward 5'-ATGAGCGGTTGGCCACGTCGTCCCC-3' Reverse 5'-CTATTCAGTGATCGCGATGAGGGCC-3'
806249	Forward 5'-ATGGAGCTGCTCGCAACGTCATTGA-3' Reverse 5'-TTACTCAGCCACTACTTCCAAAGTC-3'
814231	Forward 5'-ATGAGCACCCCCTCTCCGTTCCGTG-3' Reverse 5'-CTACTGGGAAACAGGCGTCAGCATC-3'
814014	Forward 5'-ATGCGCCTGATCGGTCTTATGTTCT-3' Reverse 5'-CTATTTGTTAGGTAGCGGGTGATCC-3'
813915	Forward 5'-ATGAGTTCGGTGGTCTTGAAGGCAG-3' Reverse 5'-TTACGAGGCGACGTCCGACGTTTGT-3'
801660	Forward 5'-ATGGCCACGTCGCTGGACCTCGCTG-3' Reverse 5'-TCAGACTCTCAAATCAATGACCATG-3'

Table 2.1b. Primers designed for inserting candidate genes into pET28a

Gene ID	Primers
804480	Forward 5'-AGAATTCGAAGATGAGGAGCTGCCAGTGACG-3' Reverse 5'-ACTCGAGTTCAGTGATCGCGATGAGGGCCA-3'
806249	Forward 5'-AGAATTCGCGCCGTTCACGTCCCCCTTTGAA-3' Reverse 5'-ACTCGAGCTCAGCCACTACTTCCAAAGTCTC-3'
814231	Forward 5'-AGAATTCGACAACTACGTCTCGGTCTGCCGTG-3' Reverse 5'-ACTCGAGCTGGGAAACAGGCGTCAGCATCG-3'
814014	Forward 5'-AGAATTCAGCACCACCGGTGCCGAGGGACGCC-3' Reverse 5'-ACTCGAGTTTGTTAGGTAGCGGGTGATCCTT-3'
813915	Forward 5'-AGAATTCTTCTTGGTCGTCGTCAATTATGG-3' Reverse 5'-ACTCGAGCGAGGCGACGTCCGACGTTTGT-3'
801660	Forward 5'-AGAATTCGTGGTACTGCGCTACGTGATCAT-3' Reverse 5'-ACTCGAGGACTCTCAAATCAATGACCATGA-3'

Table 2.1c. Primers designed to clone candidate genes via Gateway® method

Gene ID	Primers
804480	Forward 5'-ACAAGTTTGTACAAAAAAGCAGGCTTTGAAGATGAGGAGCTGCCAGT-3' Reverse 5'-CCACTTTGTACAAGAAAGCTGGGTATTATTCAGTGATCGCGATGAGGG-3'
806249	Forward 5'-ACAAGTTTGTACAAAAAAGCAGGCTTT GCGCCGTTCACGTCCCCCTT-3' Reverse 5'-CCACTTTGTACAAGAAAGCTGGGTATTACTCAGCCACTACTTCCAAAGT-3'
814231	Forward 5'-ACAAGTTTGTACAAAAAAGCAGGCTTTGACAACTACGTCTCGGTCTG-3' Reverse 5'-CCACTTTGTACAAGAAAGCTGGGTATTACTGGGAAACAGGCGTCAGCAT-3'
814014	Forward 5'-ACAAGTTTGTACAAAAAAGCAGGCTTTAGCACCACCGGTGCCGAGGG-3' Reverse 5'-CCACTTTGTACAAGAAAGCTGGGTACTATTTGTTAGGTAGCGGGT-3'
813915	Forward 5'-ACAAGTTTGTACAAAAAAGCAGGCTTTTTCTTGGTCGTCGTCAATTAT-3' Reverse 5'-CCACTTTGTACAAGAAAGCTGGGTATTACGAGGCGACGTCCGACG-3'
801660	Forward 5'-ACAAGTTTGTACAAAAAAGCAGGCTTTGTGGTACTGCGCTACGTGATCA-3' Reverse 5'-CCACTTTGTACAAGAAAGCTGGGTATTAGACTCTCAAATCAATGACCATGA-3'

Table 2.1d. Primers designed to clone synthesized genes into pFLAG-ATS

Gene ID	Primers
813024	Forward 5'-CAAAGCTTTGGGATACACCGGGTGT-3' Reverse 5'-TTGAATTCTTAATGCTGCCACTGG-3'
806256	Forward 5'-CAAAGCTTACCGAATATGCCGGTG-3' Reverse 5'-TTGAATTCTTATGCTGATCCTGAC-3'
814014	Forward 5'-CAAAGCTTAGCACCACCGGTGCAGA-3' Reverse 5'-TTGAATTCTTATTTGTTCGGCAGC-3'

2.2.8. Polymerase chain reaction (PCR)

The length of the template, the intention of further use of the desired product and the melting temperature of the primers defined the PCR conditions; volume of reaction and the enzyme mixes (BioMix™ Red – Bioline and Elongase™ - Invitrogen) to be used to amplify templates. Denaturation and extension temperatures and length, were determined according to the manufacturers and the annealing temperature varied according to the melting temperature of the primers. Duration of extension step was adjusted as 1 min for 1 kilo base of product. The reactions were prepared in 0.2 ml PCR tubes and run on Thermal cycler 2720 of Applied Biosystems® (Thermo Fisher Scientific, Waltham, USA). If the template was a plasmid DNA, the sample was diluted down to 1/500, 1/1000 or 1/2000 with sdH_2O before use. For most of the PCRs, the touchdown method was applied in which the annealing temperature was decreased 1^oC at every cycle for the first 8-10 cycles and remaining cycles (usually 25) were carried on with the lowest annealing temperature reached. Candidate genes were amplified with Elongase® enzyme mix using flanking primers with following conditions: Initial denaturation at 94^oC for 30 s, 10 cycles of denaturation at 94^oC for 30 s, annealing at 65^oC (-1^oC each cycle) for 30 s and elongation at 68^oC for 90 s, then 25 cycles of denaturation at 94^oC for 30 s, annealing at 56^oC for 30 s and elongation at 68^oC for 90 s.

In addition, candidate genes were amplified with Elongase® enzyme mix using primers designed for insertion into pET28a vector with the following conditions: Initial denaturation at 94^oC for 30 s, 8 cycles of denaturation at 94^oC for 30 s, annealing at 66^oC (-1^oC each cycle) for 30 s and elongation at 68^oC for

90 s, then 25 cycles of denaturation at 94°C for 30 s, annealing at 59°C for 30 s and elongation at 68°C for 90 s. For Gateway® cloning (see section 2.2.13.2) the PCR was done to add AttB1/2 sequences with following conditions: Initial denaturation at 94°C for 30 s, 10 cycles of denaturation at 94°C for 30 s, annealing at 66°C (-1°C each cycle) for 30 s and elongation at 68°C for 100 s, then 25 cycles of denaturation at 94°C for 30 s, annealing at 57°C for 30 s and elongation at 68°C for 100 s.

2.2.9. Reverse transcription (RT) – PCR

Bioline's MyTaq™ One-Step RT-PCR kit was used to synthesize and amplify cDNA sequences. Firstly, the template RNA samples were cleaned up and concentrations were adjusted with RNase-free water (Qiagen). The reaction mix was prepared with 10 µl MyTaq One-Step Mix, 0.8 µl forward primer, 0.8 µl reverse primer, 2 µl template RNA, 0.4 µl, RNase inhibitor, 5.8 µl DEPC-treated water, and 0.2 µl reverse transcriptase. The length of the reverse transcription and the PCR conditions (touchdown) were determined according to the gene-specific primers used, the length of the template cDNA, and the amount of product needed. Adjusted conditions for RT-PCR were as follows: 30 min reverse transcription at 45°C, polymerase activation at 95°C for 1 min, 10 cycles of denaturation at 95°C for 30 s, annealing at 65°C (-1°C each cycle) for 20 s and extension at 72°C for 90 s, then 25 cycles of denaturation at 95°C for 1 min, annealing at 56°C for 20 s and extension at 72°C for 90 s.

2.2.10. PCR purification

PCR products were cleaned up using Qiagen's QIAquick PCR Purification Kit or MinElute PCR Purification Kit depending on the size of the product and the volume to obtain after clean-up. Manufacturer's instructions were followed.

2.2.11. Agarose gel electrophoresis

Nucleic acids and PCR products were visualized via electrophoresis on 1.5% or 2% agarose gels, which were prepared by dissolving the required amount of agarose (UltraPure™, Invitrogen) in 1X Tris-Acetate-EDTA buffer (TAE, diluted from 50x stock) by heating to near-boiling in a microwave oven and 10000X GelRed™ (Biotium Inc., Hayward, CA, USA) were added to a final concentration of 0.05µl/ml after slight cooling down. Avoiding bubbles, prepared gels were poured in a cast and an appropriate comb was placed on in order to create wells. After allowing gel to set for 30-40 min, samples were loaded and the gels were run at an appropriate voltage (25 V, 50 V, or 75 V). HyperLadder™ IV (Bioline) was used as size marker and the samples were mixed with 6X loading buffer (30% glycerol, 0.25% bromophenol blue) when necessary (some PCR enzyme mixes already contained loading buffer). Volume of the sample to load, the duration and the voltage of the electrophoresis varied depending on the samples' sizes and the purposes of the experiment. The gels were visualized with BioSpectrum® - MultiSpectral Imaging System (UVP, Upland, CA, USA).

2.2.12. Gel extraction

When purification of a Nucleic acid or a PCR product from an agarose gel was necessary, QIAquick Gel Extraction Kit (Qiagen) was used following producer's instructions. Gels containing the fragment were cut using sharp blades or scalpels under UV light, taking necessary precautions.

2.2.13. Gene cloning

2.2.13.1. TOPO® TA cloning (Invitrogen)

In order to clone PCR-amplified DNA samples into subcloning and expression and sequencing vectors, TOPO® TA Cloning method was employed. Extracted DNA (from Emoy2-infected Ws-*eds*1 plants) was amplified via PCR using gene-specific primers and PCR products were cleaned-up. Purified samples were then cloned into pCR™8/GW/TOPO® vector (Spectinomycin resistance) using pCR™8/GW/TOPO® TA Cloning® Kit which enables TOPO® Cloning of Taq polymerase amplified PCR products into an entry vector for further Gateway® Cloning, following manufacturer's instructions. Constructs were then transformed into electrocompetent *E. coli* cells and plasmids were isolated from successfully transformed colonies.

2.2.13.2. Gateway® cloning (Invitrogen)

For further *in vitro* and *in planta* protein expressions, a rapid and efficient Gateway® cloning method was also followed. Gateway® method involved two reaction steps: BP and LR Cloning. BP Cloning enabled insertion of the PCR product into a donor vector (pDONR™/Zeo) using BP Clonase® II Enzyme Mix

and LR Cloning was performed to transfer the gene inserted into the donor vector (now called the cloning vector) to an expression vector using LR Clonase® II Enzyme Mix. Since addition of AttB1 (5'-ACAAGTTTGTACAAAAAAGCAGGCT-3') and the AttB2 (5'-ACCCAGCTTTCTTGTACAAAGTGGT-3') sequences to 5' and 3' ends of the desired sequences, respectively, was necessary, new primers were designed and templates were amplified via PCR with these new primers prior to the Gateway® reactions. For the BP reaction, 100 ng of cleaned-up PCR products were mixed with 1 µl pDONR™/Zeo vector and the volume of the reactions were brought up to 8 µl with Elution buffer (Qiagen, 10 mM Tris-Cl, pH 8.5). The BP Clonase® II enzyme mix was thawed on ice for 2 min and was vortexed briefly. 2 µl of the enzyme was added to the mixture and the reaction was incubated at 25°C for 1 h. The reaction was then stopped by addition of 1 µl of 2 µg /µl Proteinase K solution and incubating at 37°C for 10 min. 1 µl of this reaction mix was used to transform electrocompetent bacteria and transformation was checked via colony PCR. Plasmids from positive colonies were extracted and sequence verified. Sequence verified plasmids were used as cloning vectors in the LR reaction. For the LR reaction 1 µl of the cloning vector was mixed with 1 µl of expression vector and the reaction volume was adjusted to 8 µl with Elution buffer. The rest of the reaction was same as the BP reaction only this time using LR Clonase® II enzyme mix instead. The end products of LR reaction were also transformed into bacteria and the plasmids inserted with the gene of interest were recovered for further investigations including protein expression or plant transformation.

2.2.13.3. Cloning via restriction enzyme digestion and ligation

The plasmids containing constructed pCR™8/GW/TOPO® vector were processed with an LR reaction to transfer the genes into pEG100 vector. Candidate genes were inserted into pET28a (C-terminal Histidine-tag, Kanamycin resistance) for expression purposes through restriction and ligation. To accomplish this, the insert from the construct in pEG100 were amplified with new forward and reverse primers designed to remove signal peptide and stop codon sequences and to add *Eco*RI and *Xho*I restriction enzyme recognition sequences to 5' and 3' ends, respectively. The PCR products were cleaned up and digested first with *Eco*RI along with the vector pET28a. Eluted products were mixed with the enzyme, enzyme's buffer and sdH$_2$O according to the manufacturer's instructions. The reaction mix was incubated at 37°C for up to 2 h and the reaction was inactivated with incubation at 65°C for 15 min. The digests were cleaned up for the second digestion and *Xho*I enzyme, enzyme's buffer and sdH$_2$O were added. The reactions were incubated and inactivated as before. After both treatments the digests were visualized on 1.5% agarose gel to observe the restriction comparing with uncut controls. The digestions were repeated with restriction enzymes from different suppliers (Invitrogen, NEB, Roche). Double digestion was also tried if both the enzymes were working efficiently with the same buffer. Digested products were quantified and each insert was mixed with the vector with 2:1 ratio for ligation and T4 DNA Ligase enzyme (Invitrogen), enzyme buffer and sdH$_2$O was added to the mix, and incubated following the method for cohesive ends (both rapid and overnight) suggested in the product manual. The ligations were cleaned up by passing

through sepharose columns at 2000 rpm for 3 min, and were transformed into electrocompetent *E. coli* cells. Transformed cells were spreaded on selective LB agar (Kan+, X-gal+, IPTG+) and incubated overnight at 37°C. (X-gal suspension was prepared 50 mg/ml in DMSO and IPTG was mixed with sdH$_2$O to meet 35 mg/ml concentration. Forty microliter of each transformant was spreaded on the selective LB agar and the clones containing inserts were selected from blue-white screening.

In order to insert the genes of interest into the Golden Gate compatible pFLAG-ATS (Ampicillin resistance, N-terminal FLAG®-tag) vector; recognition sequences for *Hind*III and *Eco*RI were added to forward and reverse primers respectively, as described above. Same methods for the digestions and ligation were followed.

2.2.14. Electroporation of electrocompetent *E. coli* cells

Following a cloning procedure, the vectors containing the gene of interest were transformed into ElectroSHOX Competent Cells through electroporation. Twenty microliter of *E. coli* electrocompetent cells were mixed with 1 μl of the plasmid and kept on ice for 1 min and then transferred to a chilled electroporation cuvette (1 mm Gap - VWR International BVBA, Leuven, Belgium.) The cuvette was pulsed using a Bio-Rad MicroPulser™ (2.5 kV, 1 pulse) and 200 μl of SOC medium were added, and the mixture was incubated at 37°C, at 200 rpm for at least an hour. Afterwards, 100 μl were spreaded on an appropriate selective LB agar plate and the plates were incubated overnight at 37°C.

2.2.15. Chemical transformation of BL21 cells

One µl of the vector containing the desired gene was stirred into 20 µl of chemically competent BL21 (DE3) pLysE (Bioline) cells and the mixture was incubated on ice for 30 min, followed by a 45s heat shock at 42°C. The mixture was returned to ice for 2 min and 200 µl of SOC medium was added, and was incubated for 1 h at 37°C at 200 rpm. Subsequently, 100 µl of the culture was than spreaded on selective LB Agar plates and the plates were incubated overnight at 37°C.

2.2.16. SOC medium preparation

After transforming bacterial cells, the cultures were recovered in SOC medium. To make up a 200 ml stock of the medium; 4 g tryptone, 1g yeast extract, 0.1 g NaCl, 0.5 ml 1 M KCl, 2 ml 1 M $MgCl_2$, 2 ml 1 M $MgSO_4$ were mixed and the volume was adjusted to 196 ml with dH_2O. The solution was autoclaved (121° C, 15 min). After cooling down, 4 ml of 1 M glucose (filter-sterilized with 0.22 µm Sartorius Minisart NML, Luer lock, sterile) was added and aliquots were kept at -20° C for further use.

2.2.17. Preparation of selective media for antibiotic selection of transformed cells

According to the antibiotic resistance of the vectors used for transformation, selective media were prepared in order to allow growth of the bacterial cells that have successfully taken up the vector. All antibiotics used to prepare these media were purchased in powder form and suspensions were

prepared depending on the working concentration using appropriate solvent, filter-sterilized (0.22 µm), divided into aliquots and stored at -20ºC. Antibiotics were added to the sterile media to meet 1 µl/ml concentration after cooling down.

2.2.18. Colony PCR

Colony PCR was performed for confirmation of insert DNA in plasmid constructs. Depending on the number of the colonies growing on the plate, 5-10 were selected for screening. The colony PCR was done using Biomix™ Red. A master mix was prepared (10 µl Biomix™, 1 µl forward primer, 1µl reverse primer, 8 µl sdH$_2$O for each colony) and distributed into PCR tubes. For each colony, a sample was taken with a pipette tip (10 µl) and well-mixed into the PCR tube. Same tip was used to streak a new plate to grow the colonies for further uses. The PCR was carried out with the following conditions: Initial denaturation 94ºC for 5 min, 10 cycles of touchdown PCR (denaturation at 94ºC for 30 s, annealing –at the temperature between 65º and 56ºC, depending on primers - for 30 s, extension at 72ºC for 1 min/1 kb, 25 regular cycles (denaturation at 94ºC for 30 s, annealing – at the lowest annealing temperature of touchdown PCR- for 30 s, extension at 72ºC for 1 min/1 kb).

2.2.19. Sequencing inserts in plasmids

In order to confirm the sequences after cloning, plasmids isolated from positive colonies (determined after colony PCR) were sequenced with sequencing PCR, using BigDye® Terminator v3.1 Cycle Sequencing Kit (Invitrogen). Relatively shorter flanking and internal (forward and reverse)

primers were designed according to the length of the target sequence and for each reaction single primer was used. The reaction mix was prepared according to the manufacturer and the PCR conditions were as follows: Initial denaturation at 96°C 2 min and 35 cycles (denaturation at 96°C for 10 s, annealing at 50°C for 10 s, extension at 60°C for 3 min). The reactions were processed and the results were analysed and compared to the reference sequences using Geneious (R6) (Kearse et al., 2012).

2.2.20. Storage of transformed bacterial cultures

Glycerol stocks of the transformed cultures were made up for further use and storage. A sequence verified single colony was picked from the LB agar plate and was grown in 5 ml of selective LB at 250 rpm agitation at appropriate temperature overnight (37°C for *E. coli* and 28°C for *A. tumefaciens*). The culture was then divided into 1.5 ml Eppendorf tubes (700 μl each) and mixed with equal amount of 50% sterile glycerol. The mixes were frozen in liquid nitrogen and stored at -80°C.

2.2.21. Quantification of nucleic acids and cell cultures

NanoDrop 2000c UV-Vis Spectrophotometer (Thermo Fisher Scientific, Waltham, USA), was used to quantify nucleic acids and cell suspensions following manufacturer's instructions.

2.2.22. *In vitro* protein expression

Transformed BL21 (DE3) pLysE cells were induced for *in vitro* protein production. The temperature, duration, the concentration of the cells to be reached before induction, final concentration of IPTG, addition of L-arabinose, and the antibiotics used were the parameters varied at every trial in a controlled manner in order to optimize the process. A single colony from the transformed cells was picked and grown in selective LB. For the experiments in which His-tagged expression vectors were used, induced cells were centrifuged at maximum speed for 15 min and the pellet was obtained. The pellet was treated under native conditions (Imidazole based) or denaturing conditions (Urea and guanidine based) described in Qiagen's Ni-NTA kit or with BugBuster® Protein Extraction Reagent (Merck Millipore Ltd., Watford, UK). Treated pellets were passed through Ni-NTA Spin Columns (Qiagen) designed for purifying His-tagged proteins.

In the experiments with the vector pFLAG-ATS, transformed BL21 (DE3) pLysE cells were induced when their concentration reached 0.6 at $O.D._{600}$ (Optical density measured at 600 λ) with a final concentration of 0.4 mM IPTG. Induction lasted overnight at 37°C with 300 rpm agitation. Induced culture was then centrifuged at maximum speed for 15 min but this time the supernatant was collected and was filter-sterilized (0.22 µm) to eliminate *E. coli* cells and the proteins were precipitated with ANTI-FLAG® M2 Magnetic Beads (Sigma-Aldrich), following the product manual for the batch format. The results were then visualized via SDS-PAGE.

2.2.23. Electrocompetent cell preparation (*Agrobacterium tumefaciens*, strain GV3101)

A selective LB agar plate (50 mg/ml Rifampicin [Rif[+]] and 50 mg/ml Gentamicin [Gen[+]]) was streaked with the strain GV3101 from -80°C stock and the plate was incubated at 28°C for 1-3 days. 20 ml LB media (Rif[+], Gen[+]) was inoculated with a single colony from the plate and was cultivated overnight at 28°C (200rpm). The next day, 10 ml from the culture was transferred into 100 ml LB (Rif[+], Gen[+]) and the cells were allowed to grow at 28°C with agitation at 200rpm until the $O.D._{600}$ = 1-1.5. The culture was poured into cold centrifuge tubes (50 ml) and the tubes were incubated on ice for 20 min. Chilled cultures were centrifuged at 4°C for 15 min at 4000xg. Media was poured off completely and the cells were resuspended gently with 50 ml cold sdH_2O. Cell suspensions were centrifuged again at 4°C for 15 min at 4000xg. Pellets were resuspended with 50 ml 10% cold glycerol solution (made from 80% sterile stock) and centrifuged at 4°C for 15 min at 4000xg. This step was repeated for 4 times. Last pellets were resuspended in 5 ml 10% glycerol, combined and centrifuged again for 10 min at 4°C 3000xg. Final pellet was resuspended in 1 ml 10% glycerol, divided to aliquots and frozen in liquid nitrogen and stored at -80°C for further use.

2.2.24. *Agrobacterium tumefaciens*-mediated transformation of *Nicotiana* plants

Genes of interest cloned into the plasmid pCR™8/GW/TOPO® were transferred into pEG100 (Kanamycin resistance) expression vector with LR

reaction. pEG100 vectors, now containing the gene of interest, were electroporated into previously prepared electrocompetent *Agrobacterium tumefaciens* (strain GV3101) cells. Transformed cells were spread on selective LB (Rif+, Gen+, Kan+) plates and incubated at 28°C overnight. Insertion was confirmed with colony PCR. Selected colonies were cultivated overnight in 10 ml LB (Rif+, Gen+, Kan+) at 28°C and following day the cultures were centrifuged and resuspended in sterile 10 mM $MgCl_2$ and diluted to O.D.$_{600}$= 0.5 and O.D.$_{600}$= 0.25. The suspensions were then pressure injected using 1 ml syringes into 6-7 week-old *N. benthamiana* and *N. tabacum* leaves (~3 sq in area) and injected plants were kept at normal growth conditions. Leaves were observed for appearance of chlorosis for up to 10 days.

2.2.25. *Agrobacterium tumefaciens*-mediated transformation of *A. thaliana*: Floral dipping

A modified version of Clough & Bent's (1998) protocol for floral dipping was followed. *Agrobacterium* cells carrying the genes of interest were grown in 10 ml LB (Rif+, Gen+, Kan+) overnight at 28°C with agitation at 200 rpm. Following day, the cultures were poured in 250 ml LB (Rif+, Gen+, Kan+) and cultivated overnight with same conditions. The cultures were centrifuged at 4100 rpm for 15 min and the pellets were resuspended in 250 ml 5% sucrose (w/v). Subsequently, 50 µl Silwet L-77 (Momentive Performance Materials Inc., Stanlow, UK) was added to the resuspension as surfactant and poured in a beaker. Col-0 plants (grown until having plentiful immature flower buds and preferably few number of seed pods) were inverted and dipped in the solution

for 15-20 s with gentle stirring. Plants were then bagged to create a humid environment and incubated in the dark at room temperature overnight in a tilted position. The next day, the bags were changed with bread bags allowing air circulation and plants were moved to normal growth conditions and kept for seed collection.

2.2.26. Selection of the transformed seedlings using herbicide BASTA® (glufosinate)

Seeds were collected from Agrobacterium treated pants and sown on soil which was soaked thoroughly in 1 ml/L BASTA® solution (non-selective herbicide, Bayer CropScience AG, Monheim am Rhein, Germany) prepared in tap water, evenly and avoiding the edges. Soil trays were covered with aluminium foil and plastic bags, and stored at $4^{\circ}C$ for 7 days. After the stratification, the foil was replaced with tray lids and the trays were moved to normal growth conditions. As soon as surviving seedlings appearing greener, bigger, taller were discernible (usually within a week) they were transplanted into clean soil. Transplanted seedlings were grown until seed production stage and collected T_1 lines were challenged with BASTA® once again to obtain T_2 lines. Transgenic lines were screened with PCR for confirmation of the transferred DNA.

2.2.27. Collection of intercellular wash fluid (IWF) from infected and healthy plants

A straight-forward method, which was first introduced by Klement (1965), was used to obtain the intercellular washings. 2-week-old soil-grown healthy Ws-*eds1* seedlings were transplanted individually to modular trays and spray-inoculated with *Hpa*-Emoy2 suspension at 4-5-week-old stage. Inoculated plants were incubated as described before. 10 days after inoculation, leaves were cut off and gently washed with sdH$_2$O several times to clear away the conidia and conidiophores. Leaves were then transferred in a beaker (250 ml) containing chilled sdH$_2$O and vacuum-infiltrated for up to 1 min. Infiltrated leaves were then placed on a tissue paper to remove the remaining excess water on the leaf surfaces. Subsequently, the leaves were placed in plunger-removed 10 ml syringes (Terumo) avoiding breaking and the barrels were placed in 50 ml collection tubes. The tubes were centrifuged for 5 min at 3000 rpm at 4°C. The collections were combined and filter-sterilized (0.22 µm) to remove any possible spore and bacterial contamination. IWF was also collected from healthy leaves as control using the same procedure. Both samples were quantified with Bradford assay and visualized on 12.5% SDS-PAGE. Samples were stored in -80°C for further use.

2.2.28. Quantification of proteins using dye-based Bradford Assay

A modified version of the protocol introduced by Bradford (1976) for a rapid spectrophotometric determination of protein concentration in a solution

was used, which is based on the colour change observed when Coomassie Brilliant Blue G-250 (dark red to blue) binds to the proteins. Bovine Serum Albumin (BSA) was used as standard and by plotting known amount of BSA against the corresponding absorbance (at λ= 590 nm), a standard curve was obtained to derive the protein concentration in unknown samples. Standards were prepared from 10X BSA stock (1 mg/ml) with sdH$_2$O, filter-sterilized (0.22 µm) and stored at -20°C. The stock was diluted to 1X (0.1 mg/ml) when required. For each set of measurements, new standards were prepared in disposable semi-micro cuvettes (Fisherbrand) (Table 2.2) and mixed well.

Table 2.2. Preparation of BSA standards

Total µg of BSA	dH$_2$O (µl)	1X BSA (µl)	Bradford reagent (µl)
0	800	0	200
2	780	20	200
4	760	40	200
7	730	70	200
10	700	100	200

The reaction was allowed to develop for 2-5 min (no longer than 10 min). Using a spectrophotometer (NanoDrop 2000c), the O.D. measurements of the standards were taken at λ=590 nm, using 1 ml dH$_2$0 as blank. A standard curve (linear trendline) of absorbance vs. total µg of BSA was plotted (Microsoft Office 2010, Excel). The equation and R^2 value of the trend line were displayed. The samples to be quantified were prepared mixing 790 µl dH$_2$O, 10 µl of the sample and 200 µl of Bradford reagent. The absorbance values were recorded and the concentrations (µg/µl) were calculated using the standard curve's

equation. In some cases, NanoDrop 2000c UV-Vis was also used for direct quantification of proteins at 280 nm absorbance.

2.2.29. β -glucuronidase (GUS) reporter activity

Transgenic Col-0 plants carrying PR-GUS or RLK-GUS fusions were used to determine the activation of defence. Four to five week old plant leaves were pressure injected using 1 ml syringes (Terumo) with samples. RLK-GUS plants were incubated for 1 day, while PR-GUS plants were incubated for 2 days at normal growth conditions. Following the incubation, injected leaves were detached from the plants and immersed in the GUS staining solution (10% Triton X-100, 0.1 M X-Gluc [5-Bromo-4-chloro-3-indolyl β-glucuronide], 50 mM $K_3Fe(CN)_6$, 0.5 M EDTA, 1M $NaPO_4$, 100 mg/ml Chloramphenicol) for 5 h at 37°C. Following day, the solution was exchanged with methanol to fully destain the leaves (repeated when necessary). Insoluble blue dye accumulation was an indication of GUS expression and untransformed leaves were used as controls.

2.2.30. Sodium-Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE gels were run in order to visualize proteins. The percentage of the resolving gel was determined according to the size of the protein samples such as for bigger proteins, lower percentage was preferred. The components were combined in the given order below (Table 2.3) for the resolving gel and the gel was poured between the glasses (1 mm spacer placed

as a cassette). The top of the gel was covered with 0.1% SDS to prevent it from drying and the gel was allowed to set for up to 1 h. When the gel was set, the 0.1% SDS was poured off and the stacking gel was made up and laid over the resolving gel. The comb was placed immediately avoiding bubbles and the gel was allowed to set for up to 1.5 h.

Table 2.3.SDS-PAGE reagents

Components	Resolving Gel			Stacking gel
	10%	**12.5%**	**15%**	**4%**
dH$_2$O	2.535 ml	1.785 ml	1.546 ml	3.05 ml
1.5 M Tris pH 8.8	1.875 ml	1.875 ml	1.875 ml	1.25 ml (1M Tris pH 6.8)
20% (w/v) SDS	37.5 µl	37.5 µl	37.5 µl	50µl
30% Acrylamide (37:5:1)	2.25 ml	3.0 ml	3.5 ml	650 µl
TEMED	3.75 µl	3.75 µl	3.75 µl	5 µl
10% Ammonium persulphate	37.5 µl	37.5 µl	37.5 µl	37.5 µl

The comb was then carefully removed and the cassette was placed in the holder and put into the gel tank (Bio-Rad MiniProtean® Tetra System, Hercules, USA). The tank was filled with running buffer consisting of 25 mM Tris-HCl, 0.192 M Glycine and 0.1% SDS. The samples were quantified and mixed with required amount of 5X reducing sample buffer (60mM Tris-HCl pH 6.8, 25% glycerol, 2% SDS, 10% β-mercaptoethanol, and 0.1% bromophenol blue [optional]) and were boiled for 2 min. After brief centrifuging, the supernatant of the samples were loaded in the wells following the size marker

(HyperPAGE Pre-stained Protein Marker, 10 kDa-190 kDa, Bioline). The gel was run between 25 mA-50 mA until the dye completely ran off the gel. If the samples were not mixed with the dye-containing loading buffer, the pre-stained marker was observed considering the estimated size of the protein of interest. After completion of the electrophoresis, the gel was carefully separated from the glass pieces and was stained (0.1% w/v Coomassie Blue R-250, 45% methanol, 10% acetic acid) for at least 3 h (gently rotating on a shaker). The staining solution was replaced with 10% methanol-10% acetic acid mix to destain. This solution was changed when necessary until the gel was fully destained. Gels were visualized with BioSpectrum® - MultiSpectral Imaging System under white light. Extra precautions were taken at all steps since highly toxic substances were among the ingredients.

2.2.31. High- throughput cumulative quantification of Reactive Oxygen Species

Production of reactive oxygen species (ROS) was used as an indication of defence activation. To do this, 6 leaf disks for each treatment were removed from 4-5 week-old non-bolted *Arabidopsis* plants (6 weeks for *N. tabacum*) with a cork borer (0.5 cm diameter) and placed in a 1.5 ml microcentrifuge tube containing 1 ml sterile 1X Murashige-Skoog (MS) media. The tubes were rocked gently for 30 min to wash the disks. The wells of a 96-well microtiter plate (U-bottom, sterile) were filled with 50 µl sterile 1X MS, excluding edge wells (instead they were filled with sdH$_2$O to avoid edge effect). Each column represented a treatment and each row was for replicates. The washed leaves

were transferred individually into the wells (avoiding damaging the leaves) using forceps and each well was treated with up to 5 µg protein sample. Sterile dH$_2$O and any other buffers involved were used as negative controls and horseradish peroxidase (0.2 M prepared in sodium phosphate buffer, added just before quantification) and/or flg22 (100 nM final concentration) was used as positive control. After applying the treatments, the plate was covered with a transparent film to minimize evaporation and the edges were sealed with Parafilm®M to prevent contamination. The plate was then incubated for 16-20 h at room temperature in the light on a plate shaker with secure agitation at 500 rpm. Prior to the quantification, the substrate 5-Aminosalicylic acid was prepared (1 mg/ml) in sdH$_2$O (pre-heated to 70°C to enable solubilisation). After cooling back to room temperature, the solution's pH was adjusted to 6.0 using 1 M NaOH and 1% H$_2$O$_2$ was added to a final concentration of 1%. On the other hand, the treated solutions from the 96-well plate were transferred into a new transparent plate with a multichannel pipette (avoiding damaging the leaves) as well as the water in the edge wells, and 20 µl of horseradish peroxidase suspension was added to the positive control column. One hundred µl of the substrate was added to each treatment well with a multichannel pipette and absorbance values were measured at 450 nm with a plate-reading spectrophotometer (Biotek EL800 Microplate reader) immediately and after 2 min.

2.2.32. Ion exchange chromatography-based fractionation of IWF samples

Both infected and healthy IWF samples were adjusted to the composition of the start buffer (20 mM Tris-HCl pH 8.0) using PD-10 desalting column (GE Healthcare Life Sciences, Buckinghamshire, UK) and were filter-sterilized with 0.22 µm filters (Minisart® - Sartorius Stedim Biotech) prior to fractionation. A strong anion exchange column (GE Healthcare, HiTrap Q HP 1 ml) was used for obtaining the fractions. Manufacturer's instructions for "purification" to equilibrate the column and "elution with stepwise ionic strength gradients" to collect the fractions were followed. Start buffer (20 mM Tris-HCl, pH 8.0) and elution buffer (1 M NaCl prepared in the start buffer) were made and filter-sterilized (0.22 µm) before use. Additionally, ionic strength buffers (0.1 M, 0.15 M, 0.2 M, 0.25 M, 0.3 M, 0.35 M, 0.4 M, 0.45 M, 0.5 M and 0.55 M) were prepared by combining the start and the elution buffers. Flow-through of the applied sample and the first wash were also collected in addition to the fractions eluted with increasing ionic strength buffers (volume depended on the operating conditions). In order to increase the concentration of the proteins, samples were precipitated using acetone.

2.2.33. Acetone precipitation of proteins

Four volumes of chilled acetone were added to protein suspensions in sterile 50 ml tubes and were incubated at -20º C overnight. Samples were then centrifuged at 4ºC, at 4100 rpm for 15 min. The supernatant was discarded and the pellet was air-dried at room temperature. The dry pellet was then

resuspended in sdH$_2$O or in a buffer consisting of 20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl and 0.125% Triton X-100. The resuspended fractions were quantified using Bradford assay and visualized via 12.5% SDS-PAGE and were stored at -20° C until further use.

2.2.34. Proteomics on intercellular wash fluid

Three biological replicas from Infected and healthy IWF samples were analysed via MALDI-TOF (Matrix-assisted laser desorption/ionization – Time of flight) method by The Sainsbury Lab (Norwich, UK). The outcome of the MALDI-TOF experiments were analysed and pathogen and plant originated genes were separated. Three genes annotated as LRR family proteins (AT1G33610.1, 2 AT3G20820.1 and AT1G49750.1) were chosen to investigate further.

2.2.35. Selecting homozygous lines from the T-DNA mutants

Mutant lines for 3 LRR genes (2 for each gene) with Col-0 background were selected through T-DNA Express: Arabidopsis Gene Mapping tool (SIGnAL: Salk Institute Genomic Analysis Laboratory) and the seeds were ordered from Nottingham Arabidopsis Stock Centre (NASC) (Table 2.4).

Table 2.4.SALK and NASC ID numbers, and insertion positions of the selected mutant lines

Gene ID	SALK T-DNA Line	NASC ID	Insertion
AT1G33610.1	SALK_079946	N579946	chr1 12188805
	SALK_055913C	N660880	chr1 12190630
AT3G20820.1	SALK_119747C	N665050	chr3 7281723
	SALK_047337(CE)	N547337	chr3 7280124
AT1G49750.1	SALK_047896	N859846	chr1 18412697
	SALK_151177C	N661217	chr1 18413084

2.2.35.1. Rapid DNA extraction from plant leaves

The seeds of the mutant lines obtained from NASC were sown and grown as previously described. Plants were screened via PCR, using Extract-N-Amp™ Plant PCR Kit (or REDExtract-N-Amp™, Sigma-Aldrich) for DNA isolation and amplification. A tiny piece from the leaves were cut off and put in a 1.5 ml Eppendorf tube. Subsequently, 75 µl of extraction solution was then added to the tube -making sure the leaf piece was covered-, vortexed briefly and was incubated at 95°C for 10 min. Then, 75 µl of Dilution solution was added and vortexed. This diluted extract was used for the PCR and was stored at 4°C.

2.2.35.2. PCR for checking homozygosis

Extracted DNA samples were amplified via PCR using 3 primers: Gene-specific left and right border primer sequences were given in T-DNA Express

tool of SIGnAL, as well as the middle primer (LBb1.3:5'-ATTTTGCCGATTTCGGAAC-3') that was common for all reactions (Table 2.5). The PCR mix was prepared with 5 μl (RED)Extract-N-Amp™ PCR Ready mix, 2 μl DNA, 0.5 μl left primer, 0.5 μl right primer, 0.5 μl middle primer, and 1.5 μl of sdH$_2$O to a total of 10 μl reaction volume (water replaced DNA in control samples). The reaction was performed with following conditions: Initial denaturation at 94°C for 3 min, 10 cycles of denaturation at 94°C for 30 s, annealing at 68°C (-1°C each cycle) for 30 s and extension at 72°C for 90 s, then 25 cycles of denaturation at 94°C for 30 s, annealing at 57°C for 30 s and extension at 72°C for 90 s. The PCR products were loaded on a 1.5% agarose gel and the bands were evaluated according to the products sizes expected for wild type, heterozygous and homozygous lines (found through T-DNA Express tool and NASC). Selected homozygous lines were transplanted into big pots for seed production. The seeds were then used for pathogenicity assays. Experiments were repeated three times.

Table 2.5. Primers and expected product sizes for mutant lines

SALK ID	Left and Right Border Primers	LB-RB segment size (base)	size (base)
SALK_079946	LP 5'-TCATGGATTCGAAAATATCCG-3' RP 5'-TCGCTATTATGGTGTTCCCAC-3'	1131	469-769
SALK_055913C	LP 5'-TGTTTGTGTTGTGGGAGTGAG-3' RP 5'-GAACCTAACGTTTCCGAGGAC-3'	1164	604-904
SALK_119747C	LP 5'-CAAACCCAACTAAATGCATTCC-3' RP 5'-CGGAAGATTAAACCGGTTAGC-3'	1120	463-763
SALK_047337(CE)	LP 5'-CTCGGAATATTTCATGGCATC-3' RP 5'-TGGCCACAAAAATGAAGAAAC-3'	1072	483-783
SALK_047896	LP 5'-GCATTTGATTGATAATTCACA-3' RP 5'-TGAGTCCGAGGTATTTGTTG-3'	1190	502-802
SALK_151177C	LP 5'-TACACCCATCCCCAACTATCC-3' RP 5'-TGGAGGATTGTTGTCGTTACC-3'	1170	555-855

2.2.36. Screening mutant lines with *Hpa*-Emoy2 and *Hpa*-Cala2 for enhanced susceptibility

Selected homozygous lines, Col-0 and Ws-*eds*1 seeds were sown in modular trays, randomly distributed 3 replicates each and were grown at normal conditions. Seven-day-old plants were then spray inoculated with the pathogen isolate as explained before. Seven days after inoculation, randomly chosen 15 cotyledons from each pot were picked with a forceps and examined under stereo microscope (Wild Heerbrugg M3Z Gais, Switzerland). Sporangiophores covering one leaf of the cotyledon were counted. Average number of the sporangiophores and the standard deviation was calculated considering 3

replicates for each line. Homozygous mutant lines were compared to that of controls to have a better understanding of the role of the knocked-out gene on the susceptibility/resistance towards the pathogen.

2.2.37. Screening mutant lines with *Hpa*-Noks1 for enhanced resistance

The mutant lines' reaction towards isolate *Hpa*-Noks1 was also observed. This time instead of counting the sporangiophores, the number of spores was counted since Col-0 is highly susceptible to Noks1. Control and mutant lines were sown, grown, and infected as indicated above. 7 days after inoculation, 15 cotyledons from each pot were put in a 1.5 ml Eppendorf tube and 250 µl of sdH_2O were added. The tube was vortexed to release the spores and 10 µl of this suspension was put on the haemocytometer. The number of spores was counted under with light microscopy (10X magnification, ML6520 PCM Asbestos Microscope, Meiji Techno UK Ltd., Axbridge, UK) 4 corner squares were counted and the average was calculated. Then the average and standard deviation were calculated considering 3 replicates for each line. Homozygous mutant lines were compared to that of controls with the same purpose. Experiments were repeated three times.

3. A Bioinformatic approach towards identifying novel apoplastic effectors

3.1. Introduction

In recent years with the help of daily improving technologies, significant amount of data has been produced and gathered regarding plant-pathogen interactions. Genomes of oomycete pathogens such as *P. sojae*, *P.ramorum* and *P. infestans* have been sequenced as well as the Emoy2 isolate of *Hpa* (Tyler et al., 2006; Haas et al. 2009; Baxter et al., 2010). The available sequence data allowed generation of Expressed Sequence Tag libraries. ESTs are short cDNA subsequences resulting from single-pass nucleotide reads of cloned cDNA which can be used for gene discovery, detecting gene transcripts and proteins they encode and functions of those proteins, as well as investigating transcript and single nucleotide polymorphisms within and between species (Krajaejun et al., 2011; Cabral et al., 2011; Nagaraj et al., 2006). The EST libraries have also become valuable sources for identifying potential effector genes (Sierra et al., 2010; Torto-Alalibo et al., 2007; Randall et al., 2005). With the motivation of the previous studies and using the available sequence data of *Hpa* isolate Emoy2, mining through the EST library in pursuit of identifying novel apoplastic effectors present itself as a sensible approach.

In order to select candidate genes showing potential of being apoplastic effectors, certain criteria determined by previous studies can be applied to data sets using online bioinformatic platforms and computer programs. First of all, it is known that effectors are secreted molecules, and their secretion is enabled by an N-terminal signal peptide, which directs the effector protein to the endoplasmic reticulum to be handled and secreted out (Torto et al., 2003; Coates and Beynon, 2010). The secreted effectors are divided into two groups according to their target sites in the host as cytoplasmic and apoplastic effectors (Stassen and Van den Ackerveken, 2011). It has been shown that cytoplasmic effectors contain highly conserved motifs in their sequences that are responsible for the translocation of the effector protein into the host plant cell (Grouffaud et al., 2010). These conserved motifs are the RxLR motif chased by EER, the LxLFLAK motif of CRN effectors and the CHxC motif, which are closely following the signal peptide (Kale and Tyler, 2011; Kemen et al., 2011; Pais et al., 2013). For that reason, it is essential to pick the sequences that are predicted to contain a signal peptide and lack the given translocation motifs in order to narrow down the number of candidates for apoplastic effectors. Additionally, it is necessary to examine the sequences for previously annotated functions to eliminate the sequences with potential having intracellular functions. With the genetic and biochemical criteria in hand, in this chapter we aimed to select putative apoplastic effectors taking advantage of the EST library of *Hpa*-Emoy2 using bioinformatic tools.

3.2. Results

3.2.1. Putative secreted apoplastic effector candidates were selected from ESTs using bioinformatic tools

Over 3000 *Hpa*-Emoy2 ESTs were examined to select putative secreted apoplastic effector candidates. The short EST sequences were in translated form, therefore the complete nucleotide sequences and annotated gene ID numbers were obtained through the http://protists.ensembl.org/index.html database. In order to do that the tblastn function was used which provides translated nucleotide sequences using a protein inquiry, and finds regions of similarity between sequences. In this case the aminoacid sequences were aligned against the genomic sequence of *Hpa*-Emoy2. Out of the results, the overlapping gene with the highest score was selected and the full genomic nucleotide sequence of the gene was obtained. The complete sequences were then carried over to Geneious (R6) software and were translated. After having the amino acid sequences in hand the next step was to determine the ones containing a signal peptide as one of the first features of an effector is carrying a signal peptide that enables its secretion from the pathogen. With this purpose the amino acid sequences were examined using the online platform SMART (Simple Modular Architecture Research Tool, smart.embl-heidelberg.de) (Schultz et al., 1998; Letunic et al., 2014) which allows identification and annotation of genetically mobile domains and the analysis of domain architectures. Besides the signal peptide prediction, SMART also provides information about outlier homologues, PFAM domains, internal repeats and

intrinsic protein disorders. After checking the sequences for existence of signal peptide sequences, the ones that were not predicted to contain one were discarded. The remaining sequences were re-examined for previously annotated intracellular functions and were discarded if found any. Subsequently, the presence and the position of the signal peptide and cleavage sites in the selected amino acid sequences were rechecked and confirmed using the SignalP server (http://www.cbs.dtu.dk/services/SignalP/, Bendtsen et al., 2004) for eukaryotic organisms. Then, each remaining sequence was manually examined on the Geneious (R6) software for the motifs known for translocation of the cytoplasmic effectors that were RxLR (and EER), LxLFLAK and CHxC. The sequences carrying these motifs were eliminated. As a result, from the remaining sequences, no intron containing, relatively shorter five candidates were determined as genes *Hpa* 804480, *Hpa* 806249, *Hpa* 814231, *Hpa* 814014 and *Hpa* 813915. Below are shown the nucleotide sequences, amino acid sequences, signal peptide cleavage site predictions and SMART diagrams for the candidates. First candidate *Hpa* 804480 contained 1179 nucleotides (Fig. 3.1a) that translated to 392 amino acids (Fig. 3.1b). SMART predicted that this protein has a similar domain to metal transporter family proteins, yet this protein was uncharacterized (Fig. 3.1c). SignalP software showed the N-terminal signal peptide was 26 amino acid-long (Fig. 3.1d). The second candidate *Hpa* 806249 was 1347 nucleotide-long (Fig. 3.2a) that translated to 448 amino acids (Fig. 3.2b). SMART predicted two low complexity regions. The protein contained a domain similar to a region found in the beta-subunit of glucosidase II, however the protein itself was uncharacterized (Fig.

3.2c). SignalP software positioned the N-terminal signal peptide between amino acids 1 and 18 (Fig. 3.2d). The third candidate *Hpa* 814231 had 798 nucleotides (Fig. 3.3a) and had 265 amino acids when translated (Fig 3.3b). SMART diagram of *Hpa* 814231 showed two low complexity regions, but no domain or function predictions were found (Fig. 3.3c). SignalP software showed the N-terminal signal peptide was 20 amino acid-long (Fig. 3.3d). The fourth candidate *Hpa* 814014 was the shortest with 225 nucleotides (Fig. 3.4a) which translated to 74 amino acids (Fig. 3.4b). SMART diagram of *Hpa* 814014 showed a low complexity region and no domain or function predictions were found (Fig. 3.4c). SignalP software SignalP software positioned the N-terminal signal peptide between amino acids 1 and 19 (Fig. 3.4d). Finally the fifth candidate *Hpa* 813915 was 531 nucleotide (Fig. 3.5a) and 176 amino acid long (Fig. 3.5b) SMART diagram of *Hpa* 813915 showed no domain or function predictions for the protein (Fig. 3.5c). SignalP software showed the N-terminal signal peptide was 16 amino acid-long (Fig. 3.5d)

.

```
  1              10            20            30            40            50            60
A T G A G C G G T T G G C C A C G T C G T C C C C A T T T G C T G C T G C T C T G C G T C G C T A T T C T A A G C C T C
                 70            80            90           100           110           120
A C G C T C C T C T C A A C T G C T G A A G A T G A G G A G C T G C C A G T G A C G C C T G C G G T C C T G A C C A A A
                130           140           150           160           170           180
G A G C G C T A C G G A C A C G T G C C C A G T A C C G C C A A C C A C G A C C A T G A C C A C G A A C C T A T C A T C
                190           200           210           220           230           240
A A G C C T G A A A A C A C C G G A A G A T T C T C G A T C G T A G T G T G G A G C C A A G C G C T G C T G G C C A C G
                250           260           270           280           290           300
G C A C T T G T G G G A A C A G C T C C G G T G C T C G T G C T G C T G G C A G T G C C G C T G G G T A T G G G G A A C
                310           320           330           340           350           360
C A C G A C A A G C A A C A G C C G C T A C T G C G C G T G T T T C T G A G C T T T G C A G C C G G T G G A T T G C T G
                370           380           390           400           410           420
G G C G A C G C G C T G T T T C A T C T G C T C C C C C A C T C T C T T C C T G G G A C A G A T G G G A G T A C C T A T
                430           440           450           460           470           480
A A C C A T A A C C A T G A C C A T A A C C A T G G T A A T G A C G G T C A T G G T C A T T C C C A T T C G G T G G C T
                490           500           510           520           530           540
G A T T T G T A C G T G T G G T T G T G G A C A C T T G C T G G G A T A C T G A G C T T C C T A A T G C T G G A G A A G
                550           560           570           580           590           600
T T T G T A C G A G C A C A G A G T A G T G G A G G T G G C G G A C A C T C G C A T G G A C A T G C T C A T T C T T C C
                610           620           630           640           650           660
G C A T C T G T T C A A G A T G G A T G T G C A T T G T C G G G G A A G A C T A C T A C G G C A C G A A A G C G C T C G
                670           680           690           700           710           720
C A G T T C A A G G A A G G A A A G G A T G A T T C G G A C G C G C C G G A T G A C G A A A C A G T G G T T G T A C C A
                730           740           750           760           770           780
G A G C T G G A G A A G A A G C C G G T T G C C G C T G C A G C T T A T C T G A A T T T G G C G G C A G A T T T C T C G
                790           800           810           820           830           840
C A T A A T T T C A C G G A C G G A T T G G C G A T T G G C G C A A C T T T T G T C C G C G G T A A T G G C T G G A C G
                850           860           870           880           890           900
A C A A C A G T G G C G A T G C T G T T A C A C G A A C T G C C G C A T G A G A T C G G C G A C T T T G C C A T C C T T
                910           920           930           940           950           960
A T C C A G T C G G G C T T C A C G C G A C G T G A A G C G A T G G C C A C G C A G C T A C T G A C T G C C G T T G G A
                970           980           990          1,000         1,010         1,020
G C C A T G A T T G G C A C G G T C G T T G G T T T G T T G G T G G A A G G C G C T G G C G A T T C T A G C T C G G C G
              1,030         1,040         1,050         1,060         1,070         1,080
T G G A T C T C T C C G T T T A C A G C G G G T G G G T T C A T T T A C A T C G C G T G T G C T A G C G T C A T G C C G
              1,090         1,100         1,110         1,120         1,130         1,140
G A A T T A T T G T C A G A C T G C T C A C T C G C T C A G T C G C T C A A A G A G G G G G C T G C C A T G T G T G C T
              1,150         1,160         1,170        1,179
G G A A T T G G A C T C A T G G C C C T C A T C G C G A T C A C T G A A T A G
```

Figure 3.1a. Nucleotide sequence of *Hpa* 804480. The *Hpa* 804480 contained 1179 nucleotides.

Figure 3.1b. Amino acid sequence of *Hpa* 804480. The *Hpa* 804480 gene translated into 392 amino acids.
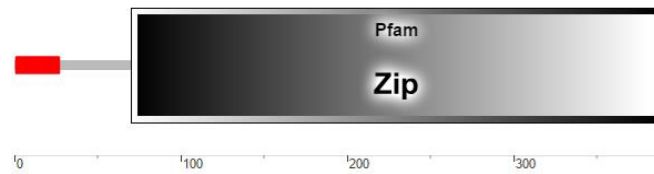
Figure 3.1c. SMART diagram of *Hpa* 806249. Red box represents the signal peptide. The results suggested a similarity to metal transporter family proteins, but the protein itself was uncharacterized.

Figure 3.1d. The signal peptide cleavage site prediction for *Hpa* 804480. SignalP software predicted the cleavage site for *Hpa* 804480 as between amino acid 26 and 27.

```
        1           10              20              30              40              50              60
        A T G G A G C T G C T C G C A A C G T C A T T G A C T C T G A G C G C C G T C T C C A C C G C G C T C G C A G C G C C G
              70              80              90             100             110             120
        T T C A C G T C C C C C T T T G A A T C G T C G T A C G T T G T G C G G C T G T A T G A G A G C C G C G A C C G C T T C
             130             140             150             160             170             180
        A T G C C G C C T C A A G A C G C A G C T G C G T T G C C G T C T C A G C T C A T G A C A A C G G A G T C T G G G C A C
             190             200             210             220             230             240
        A C G T T C G A G T G T T T T C T A C C G T C A C T C G C T C C G T C A G A C C C G G A G A C T G C G G A G C C A C T G
             250             260             270             280             290             300
        G C T A G A A A A G A C T C G A A A G A G G T C G A A A G T G A G C A A G A C A C T G A A G C C T T G G T G G C G T T C
             310             320             330             340             350             360
        C A T C G C G C T G C A G T T C G A G A G C T G G A G T C C C T C T G T A T T G A G T A T A C A G A C G A A G A G A C C
             370             380             390             400             410             420
        T C G T G G T T C T A C C G A A T C T G T G C A A A C T C G T T G A T C T T C C G G T C C C G A A A G C C G T T G A C T
             430             440             450             460             470             480
        G A A G T G G C A A T T G A A G T A A C A G G A G C A G C T G C A C C A G C A C A A G T T G A G G A G G G A G G T G G T
             490             500             510             520             530             540
        G G G T A T G A T G T G G A A G A A G T T G G T A C T T T T G T G A C T G A G A G T A G T C G C T C G G T T C T C T C T
             550             560             570             580             590             600
        T A C G A T G C G T T T G C T G A C C T T G A G A C G C A A G A G C G A G T A A A G G A T C T C G G T C A C G C G T T G
             610             620             630             640             650             660
        C T T A C A C A A A C G T T T A G C A A G A A C G G A C A A G A G G T G C A G G T G C A G T T T A T C T G C A G T G C T
             670             680             690             700             710             720
        A G T G C A C A G G A C G A T G T C G T G G C T G C T G T G C A G T G G C G A G A A G C G G T G G C G A C G G A C G G T
             730             740             750             760             770             780
        A G A C G G G A G A T T G C T G C A T T T C T G G T C G A G T C A C A A G T G T T T T G T A A G T C G C A A G A G T C A
             790             800             810             820             830             840
        G A A G T A G A C A C A G A T A G A C T G G T T A C G A T G C A G T C G C T G C T G C A G C C T C T G C A G A A T G A G
             850             860             870             880             890             900
        C A C A C G T G T G T C A C G C G A G A C G A A G G C T G G T G G A C C T A T G A G T A C T G C T T T G G A C G G T C A
             910             920             930             940             950             960
        G T G C G A C A G T A C C A T C G C G A C G A A G A C G G T C A G C T C A C A G C C G A C T A T A C G C T C G G T G T G
             970             980             990           1,000           1,010           1,020
        T T T G A T G C A G A C G G G A A C C G C G A A T T G G A G C G G T T G G A C T C A G C G C T G G T G T T G G A G C C T
            1,030           1,040           1,050           1,060           1,070           1,080
        T T T G A T G A G A C T C A T G A C G T G T C A C G T C C T G C G T A C T T G G A G T T G T A C A A T T T T G G A A C T
            1,090           1,100           1,110           1,120           1,130           1,140
        T T G T G C A A A A C A T C G G A G A A G C A G G C G C C G C G G A A G G C C A A G G T C T T T C A C T T C T G C A A C
            1,150           1,160           1,170           1,180           1,190           1,200
        C G A G A C G G C T C G A G C C A G C G C C T G A T A A T G G T G A A A G A G A T G C A G A C G T G C G T T T A C A C G
            1,210           1,220           1,230           1,240           1,250           1,260
        G T C A A G G T G T T G T C T C C A G T C T T T T G C G A C C A C C C A C A T T T T C T C A A C G A T G A G C A G G A A
            1,270           1,280           1,290           1,300           1,310           1,320
        A G C G A T G A A C G T T T G A G A T C G T G C A C T G C A T C C C G G C A G T A G A A G A A G C C G A A G C A G T G
            1,330           1,340         1,347
        G A G A C T T T G G A A G T A G T G G C T G A G T A A
```

Figure 3.2a. Nucleotide sequence of *Hpa* 806249. The *Hpa* 806249 contained 1347 nucleotides.

Figure 3.2b. Amino acid sequence of *Hpa* 806249. The *Hpa* 806249 gene translated into 448 amino acids.

Figure 3.2c. SMART diagram of *Hpa* 806249. Red box represents the signal peptide. Pink boxes represent low complexity regions. The protein contained a domain similar to a region found in the beta-subunit of glucosidase II, however the protein itself was uncharacterized.

Figure 3.2d. The signal peptide cleavage site prediction for *Hpa* 806249. SignalP software predicted the cleavage site for *Hpa* 806249 as between amino acid 18 and 19.

```
     1          10              20              30              40              50
A T G A G C A C C C C C T C T C C G T T C C G T G C T G C A G C A T T G G C C G C C C T G G C C G T
         60              70              80              90             100
C G T C T C T G C C G A C A A C T A C G T C T C G G T C T G C C G T G A C G C G A C T T A C A A T G
    110             120             130             140             150
T C G C G A A C T C C A A C G G C G C C A T T T G C T C T G G T G C G G G T G A C G C C C C A G C C
    160             170             180             190             200
G G C A C G A T C T G C C C C A A A G C T G G G G A C A T G G C T G C T G G C G A C T G C C A C C C
    210             220             230             240             250
G T A C T T G C C C T C G T A C G A C G G C T C C C A G T G C A A G G C C A A G G A G G A C G C C G
    260             270             280             290             300
T G T G T G C C A T C G T G A C C G G C A A C A C G T G G G G C T G C G T C T T C C C G T C C G T G
    310             320             330             340             350
G G C T G C C A C G A C A G C G C C A C G A T G C C G G G G G C T C A G C C G C C G A C G T C G G A
    360             370             380             390             400
C G T C A A C A T G T C C A T G G A C C C T C A C A T G C C G G G C C C T A A G A T G A C G G A C G
    410             420             430             440             450
C C C A C A C C A T G C C G G G T G G T G T C C C T G C G A C T C C T G A G A C T C C G T C C G G A
    460             470             480             490             500
A T G A C G C C C T C G G T C C T G C A C G A C C C A A G T G C T G G T G G A A A C C A G T A C A T
    510             520             530             540             550
G C C G A C T A C A A T G G C T C C T G T A A T C C C G G G G T A C G G C G A C A C G C T C C A T G
    560             570             580             590             600
A A A A A A T G C C C C A C G C C G A G G G T A A A G G A G T T G G T G G T T C C C C G G C G T C A
    610             620             630             640             650
A T G A C G A C C C C A G A A G C G A C C C C C A C G G C G A C G A A T C C G G A G A C C A A G A C
    660             670             680             690             700
G A C G A G C T C G G A G A A C A T G A T G A C G C C G C C A A C T G T C C C T A T G A C G C C G C
    710             720             730             740             750
C A A C T G C C C C C A T G A C G A T G C C A G C G A C A A C C C C C G A G A T G C C A G C T G C T
    760             770             780             790             798
A C T C C T G A G A T C G A G A C G G A C C C G A T G C T G A C G C C T G T T T C C C A G T A G
```

Figure 3.3a. Nucleotide sequence of *Hpa* 814231. The *Hpa* 814231 contained 798 nucleotides.

Figure 3.3b. Amino acid sequence of *Hpa* 814231. The *Hpa* 814231 gene translated into 265 amino acids.
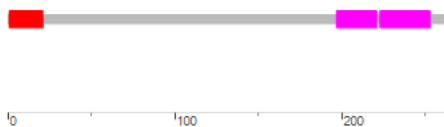
Figure 3.3c. SMART diagram of *Hpa* 814231. Red box represents the signal peptide. Pink boxes represent low complexity regions. No domain or function predictions were found.
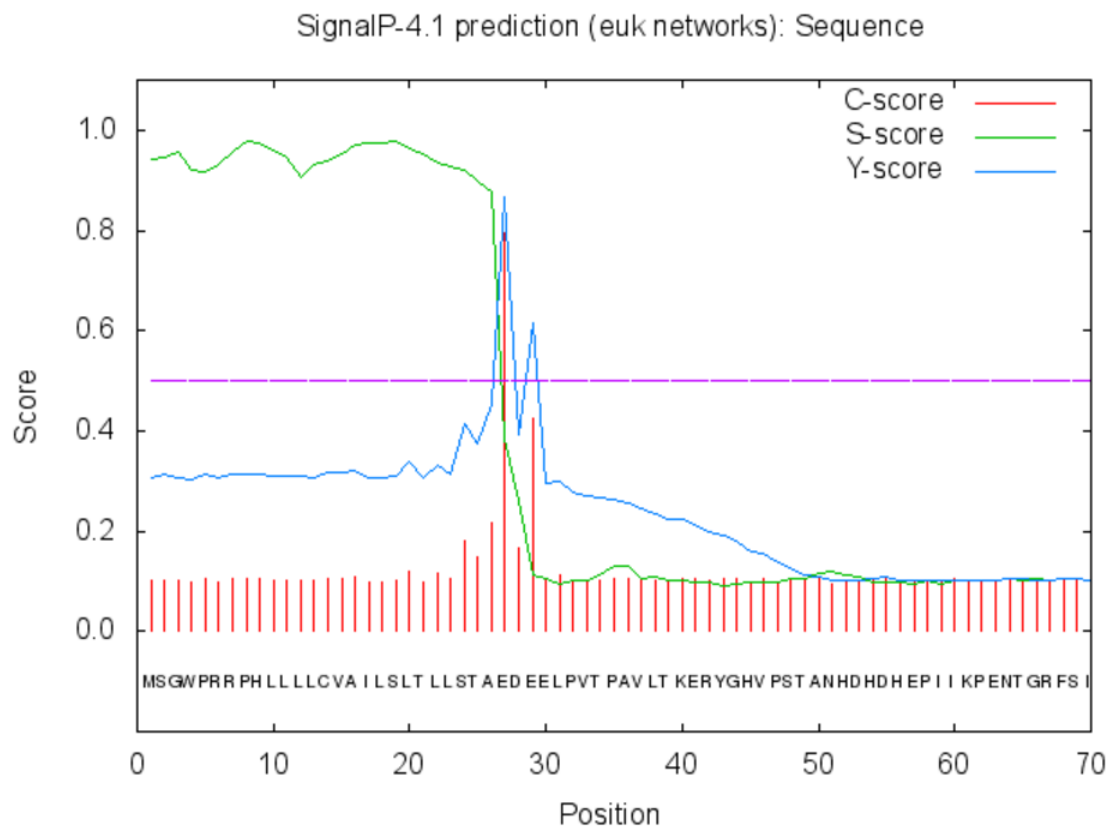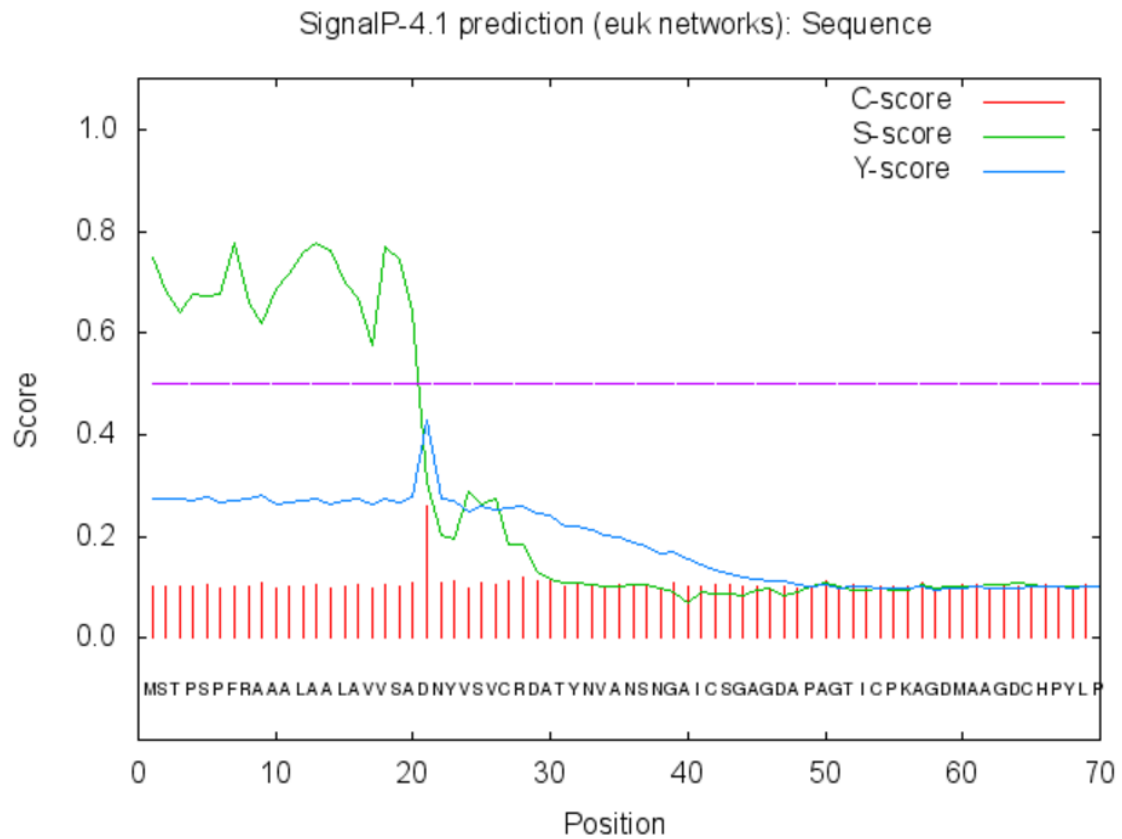
Figure 3.3d. The signal peptide cleavage site prediction for *Hpa* 814231. SignalP software predicted the cleavage site for *Hpa* 814231 as between amino acid 20 and 21.

```
1                    10                    20                    30                    40                    50
A T G C G C C T G A T C G G T C T T A T G T T C T T C A C A G T G G C G A C C C T A G C T G C A A C
              60                    70                    80                    90                    100
C A G C G C G A G C A C C A C C G G T G C C G A G G G A C G C C G T A G C C A A A A A A G A A C T
              110                   120                   130                   140                   150
C T G C A G A A C C T G A A G A C G A A T C A A G T T T G G A G G A T T C T T C G A G T G C G T T G
              160                   170                   180                   190                   200
G C T G C T G C T T A T G T C A A A C T T A T G A A C G G G C A T G T C A G A G G G T T T C G C A A
              210                   220        225
G G A T C A C C C G C T A C C T A A C A A A T A G
```

Figure 3.4a. Nucleotide sequence of *Hpa* 814014. The *Hpa* 814014 contained 225 nucleotides.

Figure 3.4b. Amino acid sequence of *Hpa* 814014. The *Hpa* 814014 gene translated into 74 amino acids.
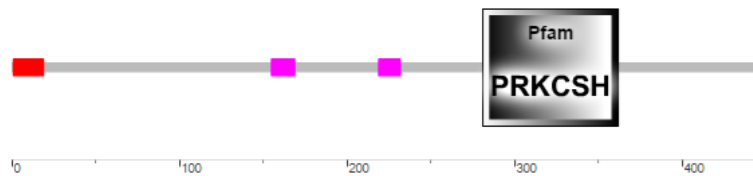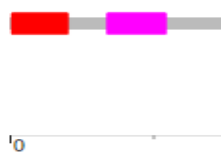
Figure 3.4c. SMART diagram of *Hpa* 814014. Red box represents the signal peptide. Pink box respresents low complexity regions. No domain or function predictions were found.
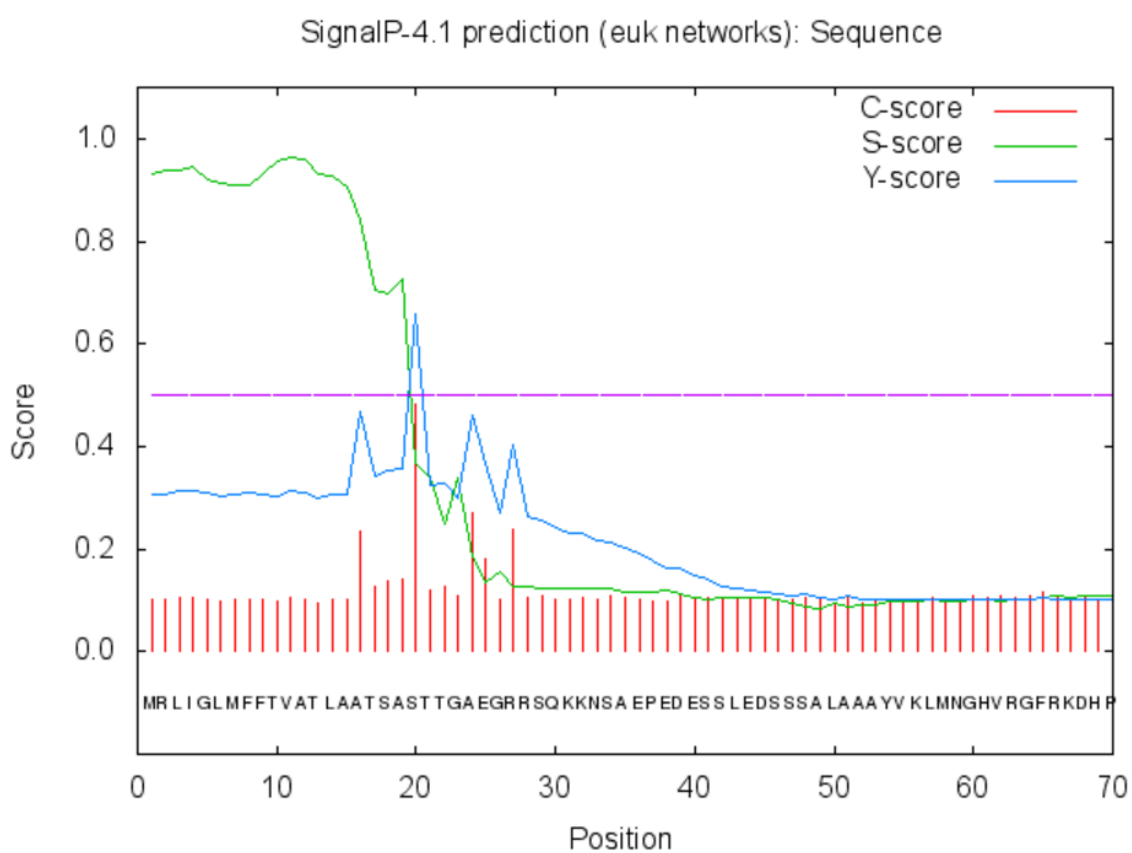
Figure 3.4d. The signal peptide cleavage site prediction for *Hpa* 814014. SignalP software predicted the cleavage site for *Hpa* 814014 as between amino acid 19 and 20.

```
  1              10                  20                  30                  40                  50
A T G A G T T C G G T G G T C T T G A A G G C A G C A G T A T T T C A C G T C T C G G A A T C T T T
                 60                  70                  80                  90                 100
C T T G G T C G T C G T C A A T T A T G G C G T C G G T G G T G G A G A T C G C A C A G C G A G A A
                110                 120                 130                 140                 150
C C A T T G T C A A C A C G A G G A C G T A C G C A A G T G A A C G T G T T C G C A G C G A G G A T
                160                 170                 180                 190                 200
A A C T G C T T G C G C A A T G A T A T C G C T T A T A A G C T C G A T C T G G T T A T G T C A C G
                210                 220                 230                 240                 250
A A G A C C C G A A A G C T C C A C C G A C T C C G C G A C A T C A G T T T G G A T G C G T C G A C
                260                 270                 280                 290                 300
A A A G A G A A G C A T T C A A G C G G A G A T A C A G C A A G A T T G T G A C G C A G A G A G A C
                310                 320                 330                 340                 350
G C G G T G T C T C C A G A C T G G T C C A C A G C A G G A A A T C T T T T T T C A A C A T C C G T
                360                 370                 380                 390                 400
C A A A G T G T T T A A G T C T G C A A G C G T C C T C G C C C G A G A A A T A G A C G A T C G C A
                410                 420                 430                 440                 450
T G G A C G A G T C A A G G A C A G G A A G C G G C G A T T C C A G C A A A G T C G C G A T C A G T
                460                 470                 480                 490                 500
C G G G A A A T C T T C A A G C T C A G G A C C T T G C A T G C G C T C G T C G A G A C T C G A G A T
                510                 520                530 531
G A G A C G A C A A A C G T C G G A C G T C G C C T C G T A A
```

Figure 3.5a. Nucleotide sequence of *Hpa* 813915. The *Hpa* 813915 contained 531 nucleotides.

Hpa 813915 amino acid sequence

```
  1              10                  20                  30                  40                  50
M S S V V L K A A V F H V S E S F L V V V N Y G V G G G D R T A R T I V N T R T Y A S E R V R S E D N C L R
                 60                  70                  80                  90                 100
N D I A Y K L D L V M S R R P E S S T D S A T S V W M R R Q R E A F K R R Y S K I V T Q R D A V S P D W S T
                110                 120                 130                 140                 150                 160
A G N L F S T S V K V F K S A S V L A R E I D D R M D F S R T G S G D S S K V A I S R E I S S S G P C M R S
                170                 177
S R L E M R R Q T S D V A S *
```

Figure 3.5b. Amino acid sequence of *Hpa* 813915. The *Hpa* 813915 gene translated into 176 amino acids.

Figure 3.5c. SMART diagram of *Hpa* 813915. Red box represents the signal peptide. No domain or function predictions were found for the protein.
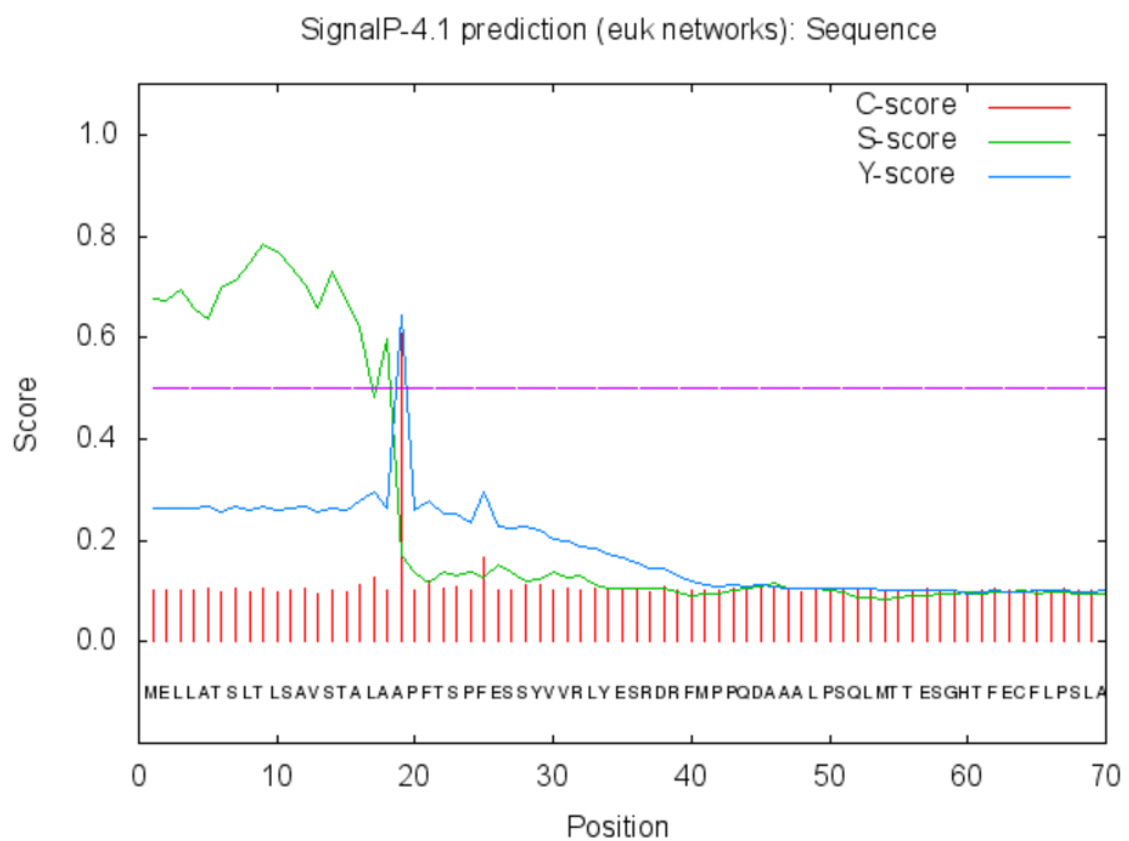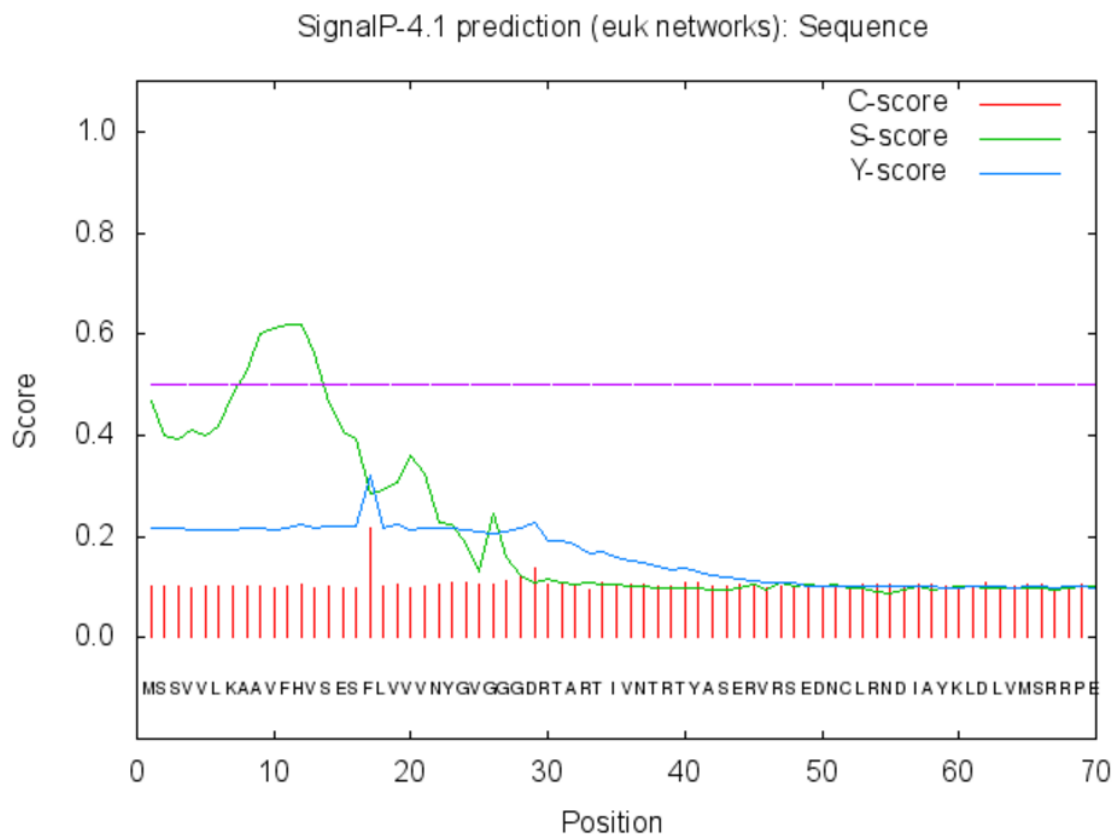
Figure 3.5d. The signal peptide cleavage site prediction for *Hpa* 813915. SignalP software predicted the cleavage site for *Hpa* 813915 as between amino acid 16 and 17.

3.3. Summary

With the aim of choosing candidate apoplastic effectors taking advantage of the available genomic data of *Hpa*-Emoy2, we screened over 3000 short sub-sequences from cDNA (ESTs) database. The criteria for selection was determined according to the known common features of the so-far-identifed effectors. A combination of bioinformatic tools and software was used to select five candidates without a function assigned, all of which contained a signal peptide and lacked the signatures of cytoplasmic effectors such as RxLR (and EER), LxLFLAK or CHxC. It was also essential to choose relatively shorter

sequences for easier manipulation in further assays. The selected candidate genes were then subjected to numerous assays and molecular characterization techniques to validate their expression and investigate their roles in host defence activation/suppression.

4. Molecular characterization and evaluation of the candidate genes

4.1. Introduction

In the previous chapter, five candidates were selected from the EST library of *Hpa*-Emoy2 applying certain criteria and following a bioinformatic pathway. In this chapter, our aim is to investigate these candidate genes in terms of expression patterns, *in vitro* expression assays, their capabilities of triggering plant defence responses with *in planta* assays and observing their variations across other *Hpa* isolates.

As *Hpa* is an adapted pathogen of *A.thaliana*, we took advantage of this extensively used pathosystem (Slusarenko and Schlaich, 2003) to scrutinize the molecular interactions between the pathogen and the plant. There are various examples of effector discovery via a bioinformatic approach. EST-mining has led to the discovery of the Serine protease-inhibitors EPI1 and EPI10 (Tian et al., 2004), cysteine protease inhibitors EPIC1 and EPIC2 (Tian et al., 2007) and small cysteine-rich proteins, PcF-like SCR74 and SCR91 (Orsomando et al., 2011). In all of these studies, candidate genes were validated for their expressions and were subjected to subsequent *in vitro* and *in planta* functionality assays. Based on that, here we pursued a similar fashion and

isolated DNA infected Ws-*eds1* seedlings and amplified the candidate genes via PCR, in order to show that the candidates actually existed in the *Hpa*-Emoy2 isolate and to use for cloning purposes. Following that, candidate genes were cloned for *in vitro* and *in planta* expression and functionality assays in order to assess whether they were able to trigger defence responses.

The expression levels of the candidate genes were observed as well, over a 7-day period at 24 h intervals by performing RT-PCR. As a reference *actin* gene was also amplified. A*ctin* has been a useful tool as a control for being expressed in an increasing mode as a response to pathogen invasion during a biotrophic infection, due to rearrangements of the host cytoskeleton (Jin et al., 1999; Kamoun et al., 1998).

On the other hand, in order to observe whether the candidate genes were causing HR reaction, we tried *Agrobacterium*-mediated transient expression method. In this system a culture of the bacteria *Agrobacterium tumefaciens*, carrying the gene of interest is infiltrated into plant leaves, such as tobacco (Zupan and Zambryski, 1995; Li et al., 2009). *A. tumefaciens* is known for its ability to transfer genes when it is used to infect a plant. What happens underneath is that the bacteria transfer T-DNA into the plants cells, a segment that is a part of the bacterial tumour-inducing plasmid; and when the bacteria is engineered to carry the gene of interest, it is delivered alongside. The gene of interest is then expressed for only a short amount of time, since it is not integrated into the plant genome, and not replicated. In previous studies, it was shown that the RxLR effector AVR3a was able to trigger HR response and suppress INF1-induced cell death in *N. benthamiana* plants via

Agrobacterium-mediated transformation (Bos et al., 2006). Additionally, Liu et al. (2011) used bioinformatically-selected, expression–validated cytoplasmic effector candidates and observed their role in HR induction, following the same method. In this chapter we also tried to transiently express our candidate genes in *Nicotiana benthamiana* and *Nicotiana tabacum* plants. For its known ability to trigger HR, *Erwinia amylovora* was used as positive control (Kim and Beer, 1998).

Furthermore, we wanted to create transgenic lines carrying the candidate genes in order to observe whether inserting the genes into resistant Col-0 genome was resulting in any difference in resistance or susceptibility of the plant towards the pathogen when challenged by the pathogen. With this purpose a method proposed by Clough and Bent (1998) was used. The technique is an *Agrobacterium*-mediated transformation based on simple dipping of developing floral tips (female gametes) of the plant into a solution containing the *Agrobacterium* cells carrying the gene of interest.

It was important to express the candidate genes *in vitro* to be able to use the proteins they encode in subsequent functionality assays to assess whether they carried the characteristics of an effector. Previously, the recognition of the INF1 elicitor protein was studied by inserting the encoding gene into a FLAG-tagged expression vector (Kamoun et al., 1998). Similarly, the cytoplasmic effectors ATR1 and ATR13 went into an expression vector for further assays (Sohn et al., 2007) Another effector's encoding gene, Avr1b, was inserted into expression vector fused with GFP for translocation assays (Dou et al., 2008). Moreover, Gateway® cloning method was used to express the Avr3a gene

(Bos et al., 2010). Therefore, in this chapter we also wanted to clone the candidate genes into expression vectors, induce the expression of the genes *in vitro*, and purify the proteins they encode to be able to use in functionality and localisation assays.

Last but not least, the candidates were examined to see whether they showed variations across other *Hpa* isolates or they were conserved. According to the studies, avirulent members of RxLR family effectors under pressure to evolve and they show variations (Win et al., 2008; Allen et al., 2004; Rose et al., 2004; Haas et al., 2009). Conversely, LysM effectors and apoplastic Nep1-like proteins were shown to be conserved molecules (Gizjen and Nürnberger, 2006). High levels of variations and mutations can be an indication of a pathogen's effort to avoiding recognition. On the other hand, a conserved nature may suggest that an effector can be perceived as a PAMP by PRRs, or can be used to design resistant plants. (Thomma et al., 2011; Bart et al., 2012). Therefore in this chapter, the candidate genes were amplified in 8 other isolates of *Hpa* and were sequenced referencing to *Hpa*-Emoy2, to observe mutations and to have a better understanding on their recognition.The results of the assays are presented below.

4.2. Results

4.2.1. The candidate genes were amplified successfully

In order to amplify the genes of interest, flanking primers were designed manually using Geneious (R6) (Table 2.1a). DNA was isolated from 2-week-old *Hpa*-Emoy2-inoculated Ws-*eds*1 seedlings and a touchdown PCR was

performed with Elongase® enzyme mix using designed primers. The amplicons were then run on 2% agarose gel at 50 V for 30 min, which showed intact bands in line with corresponding size indicated by the DNA ladder thus proving the successful amplification of the candidate genes (Fig 4.1).
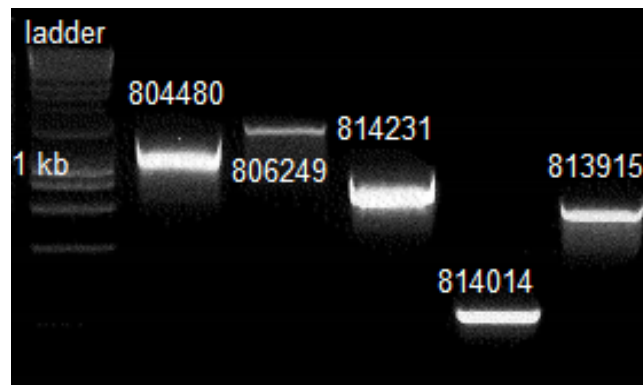
Figure 4.1. Amplification of the candidate genes. DNA was isolated from 2-week-old infected seedlings and the candidate genes were amplified with gene specific primers via PCR. The bands corresponded to the expected size indicated by the ladder. *Hpa* 804480: 1179 bases, *Hpa* 806249: 1347 bases, *Hpa* 814231: 798 bases, *Hpa* 814014: 225 bases and *Hpa* 813915: 531 bases.

4.2.2. Expression pattern analysis reveals differences in the expression times and levels of the candidate genes

Seven-day-old Ws-*eds*1 seedlings were inoculated with *Hpa*-Emoy2 to observe the daily change of the expression levels of the candidate genes after infection. Seedlings were collected from healthy plants prior to inoculation as negative controls. Then seedlings from the infected plants were collected every 24 h after inoculation for 7 days. Samples were immediately frozen in liquid nitrogen and were kept at -80°C and were stored until the last sample was collected. Then,

RNA was isolated and DNA was eliminated from each sample. Cleaned-up samples were amplified via RT-PCR (touchdown) using flanking primers, during which reverse transcription and PCR amplification was done at single step. *Actin* gene was also amplified using the RNA samples with its gene-specific primers for comparison. For *Hpa* 806249 and *Hpa* 813915, the expression started 2 days after inoculation (dai) and the level of expression seemed to stay the same. For *Hpa* 814231, the expression was also starting 2 dai but the level of expression seemed to decrease slightly 6 and 7 dai. The expression of *Hpa* 804480 was starting 3 dai whereas for *Hpa* 814014 the expression was starting in the very first 24 h after infection and slightly increased towards the 7th dai. In case of *actin*, expression initiated on the 3rd dai . On the 4th dai, the level of expression slightly increased and stayed at that level until 7 dai. (Fig. 4.2a-f)
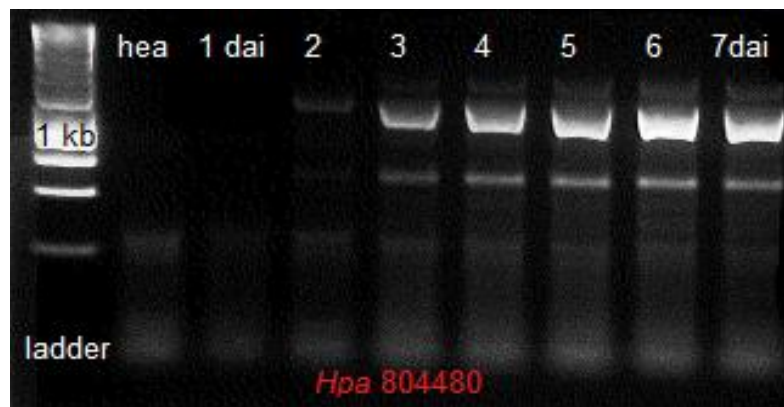
Figure 4.2a. Expression pattern of *Hpa* 804480 (hea: healthy). The gene expression started 3dai and the level of gene expression increased towards 7dai.
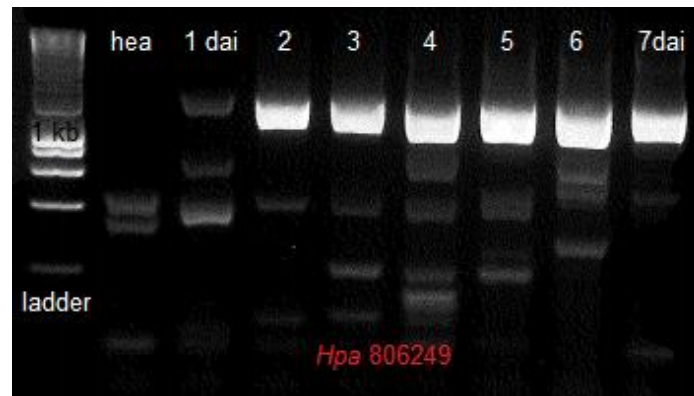
Figure 4.2b. Expression pattern of *Hpa* 804249. The gene expression started 2 dai and expression levels stayed around the same level through 7dai.
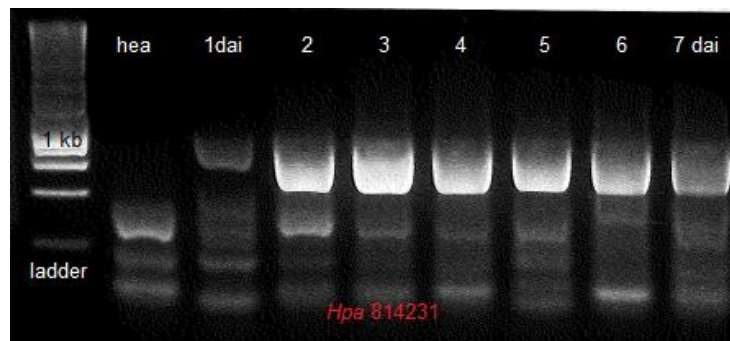
Figure 4.2c. Expression pattern of *Hpa* 814231. The gene expression started 2 dai but the level of expression decreased slightly 6 and 7 dai.
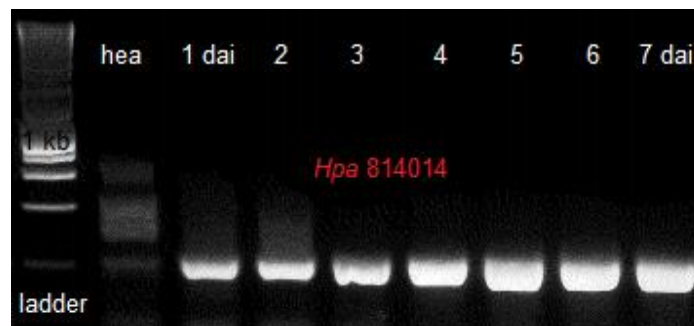
Figure 4.2d. Expression pattern of *Hpa* 814014. The gene expression was starting in the very first 24 h after infection and slightly increased towards the 7th dai
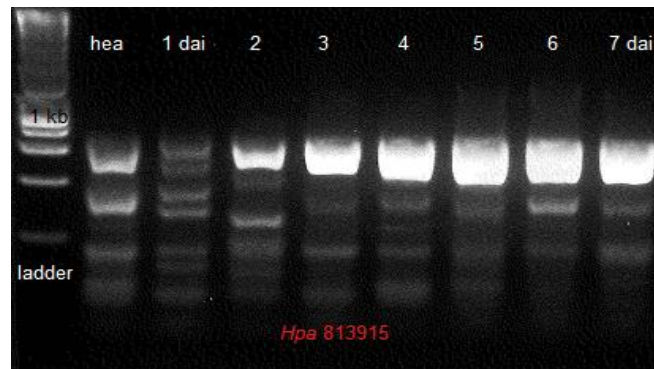
Figure 4.2e. Expression pattern of *Hpa* 813915. The gene expression started 2 dai and expression levels stayed around the same level through 7dai.
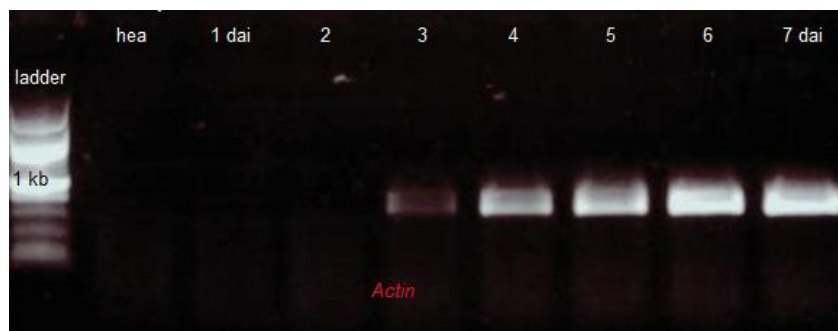
Figure 4.2f. Expression pattern of *Actin*. *Actin*'s expression initiated on the 3rd dai. On the 4th dai, the level of expression slightly increased and stayed at that level until 7 dai.

4.2.3 Transient expression of the candidate genes suggests that candidates are unlikely to cause an HR

The PCR products were cleaned up and cloned into pCR™8/GW/TOPO® vector (Fig. 4.3a-b) following TOPO® TA Cloning method. Cloning was confirmed via colony PCR using Elongase® enzyme mix with flanking gene-specific primers. Plasmids were isolated from surviving colonies and were

sequence verified. Sequencing results were aligned according to forward and reverse reading and were referenced to the original sequence. The orientation of the insertion was confirmed comparing to the expected position of the PCR product within the pCR™8/GW/TOPO® vector as shown in Fig 4.4a-e.
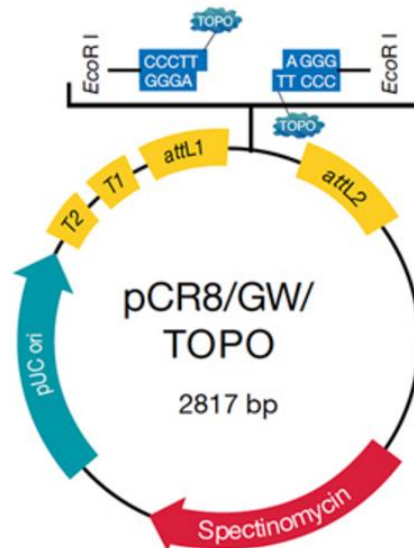
Figure 4.3a. The vector map of pCR8/GW/TOPO. This vector enables TOPO-TA cloning.Selective antibiotic resistance is against spectinomycin. The insert is placed between flanking EcoRI sites. (Source: Thermofisher)
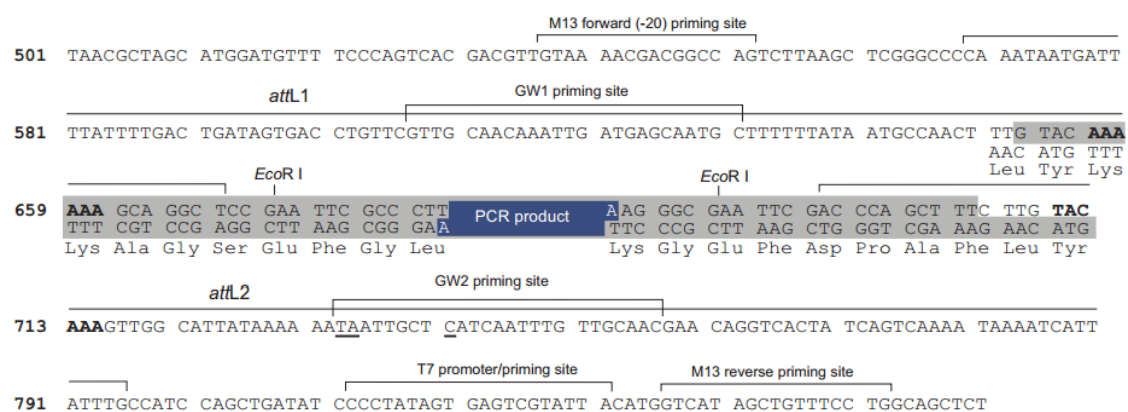
Figure 4.3b. The pCR8/GW/TOPO cloning region. The PCR product to be inserted is flanked by EcoRI recognition sites. (Source: Invitrogen)

Figure 4.4a. Sequence and orientation confirmation of *Hpa* 804480 in pCR8/GW/TOPO plasmid. Multiple forward and reverse sequence reads were aligned and matched with the original sequence and flanking nucleotides were compared to the plasmid map. Sequences underlined with green bars prove the insertion in the correct orientation.

```
   1        10          20          30          40          50          60          70          80
ĀGGCTCCGAĀTTCGCCCTTĀTGGAGCTGCŤCGCAACGTCĀTTGACTCTGĠCGCCGTCTĊCACCGCGCTĊGCAGCGCCGŤ
```

```
          90         100         110         120         130         140         150         160
TCACGTCCCĊCTTTGAATCĠTCGTACGTTĠTGCGGCTGTĀTGAGAGCCGĊGACCGCTTCĀTGCCGCCTCĀAGACGCAGCŤ
```

```
         170         180         190         200         210         220         230         240
GCGTTGCCGŤCTCAGCTCAŤGACAACGGAĠTCTGGGCACĀCGTTCGAGTĠTTTTCTACCĠTCACTCGCTĊCGTCAGACCĊ
```

```
         250         260         270         280         290         300         310         320
GGAGACTGCĠGAGCCACTGĠCTAGAAAAGĀCTCGAAAGAĠGTCGAAAGTĠAGCAAGACAĊTGAAGCCTTĠGTGGCGTTCĊ
```

```
         330         340         350         360         370         380         390         400
ATCGCGCTGĊAGTTCGAGAĠCTGGAGTCCĊTCTGTATTĠGTATACAGAĊGAAGAGACCŤCGTGGTTCTĀCCGAATCTGŤ
```

```
         410         420         430         440         450         460         470         480
GCAAACTCGŤTGATCTTCCĠGTCCCGAAAĠCCGTTGACTĠAAGTGGCAAŤTGAAGTAACĀGGAGCAGCTĠCACCAGCACĀ
```

```
         490         500         510         520         530         540         550         560
AGTTGAGGAĠGGAGGTGGTĠGGTATGATGŤGGAAGAAGTŤGGTACTTTTĠTGACTGAGAĠTAGTCGCTCĠGTTCTCTCTŤ
```

```
         570         580         590         600         610         620         630         640
ACGATGCGTŤTGCTGACCTŤGAGACGCAAĠAGCGAGTAAĀGGATCTCGGŤCACGCGTTGĊTTACACAAAĊGTTTAGCAAĠ
```

```
         650         660         670         680         690         700         710         720
AACGGACAAĠAGGTGCAGGŤGCAGTTTATĊTGCAGTGCTĀGTGCACAGGĀCGATGTCGTĠGCTGCTGTGĊAGTGGCGAGĀ
```

```
         730         740         750         760         770         780         790         800
AGCGGTGGCĠACGGACGGTĀGACGGGAGAŤTGCTGCATTŤCTGGTCGAGŤCACAAGTGTŤTTGTAAGTCĠCAAGAGTCAĠ
```

```
         810         820         830         840         850         860         870         880
AAGTAGACAĊAGATAGACTĠGTTACGATGĊAGTCGCTGCŤGCAGCCTCTĠCAGAATGAGĊACACGTGTGŤCACGCGAGAĊ
```

```
         890         900         910         920         930         940         950         960
GAAGGCTGGŤGGACCTATGĀGTACTGCTTŤGGACGGTCAĠTGCGACAGTĀCCATCGCGAĊGAAGACGGTĊAGCTCACAGĊ
```

```
         970         980         990       1,000       1,010       1,020       1,030       1,040
CGACTATACĠCTCGGTGTGŤTTGATGCAGĀCGGGAACCGĊGAATTGGAGĊGGTTGGACTĊAGCGCTGGTĠTTGGAGCCTŤ
```

```
       1,050       1,060       1,070       1,080       1,090       1,100       1,110       1,120
TTGATGAGAĊTCATGACGTĠTCACGTCCTĠCGTACTTGGĀGTTGTACAAŤTTTGGAACTŤTGTGCAAAAĊATCGGAGAAĠ
```

```
       1,130       1,140       1,150       1,160       1,170       1,180       1,190       1,200
CAGGCGCCGĊGGAAGGCCAĀGGTCTTTCAĊTTCTGCAACĊGAGACGGCTĠGAGCCAGCGĊCTGATAATGĠTGAAAGAGAŤ
```

```
       1,210       1,220       1,230       1,240       1,250       1,260       1,270       1,280
GCAGACGTGĊGTTTACACGĠTCAAGGTGTŤGTCTCCAGTĊTTTTGCGACĊACCCACATTŤTCTCAACGAŤGAGCAGGAAĀ
```

```
       1,290       1,300       1,310       1,320       1,330       1,340       1,350       1,360
GCGATGAAAĊGTTTGAGATĊGTGCACTGCĀTCCCGGCAGŤAGAAGAAGCĊGAAGCAGTGĠGAGACTTTGGĀAGTAGTGGCŤ
```

```
       1,370       1,380       1,390
GAGTAAAAGĠGCGAATTCGĀCCCAGCTTTĊ
```

Figure 4.4b. Sequence and orientation confirmation of *Hpa* 806249 in pCR8/GW/TOPO plasmid. Multiple forward and reverse sequence reads were aligned and matched with the original sequence and flanking nucleotides were compared to the plasmid map. Sequences underlined with green bars prove the insertion in the correct orientation.

```
  1         10        20        30        40        50        60        70        80
GCTCCGAATTCGCCCTTATGAGCACCCCCTCTCCGTTCCGTGCTGCAGCATTGGCCGCCCTGGCCGTCGTCTCTGCCGACA
        90        100       110       120       130       140       150       160
ACTACGTCTCGGTCTGCCGTGACGCGACTTACAATGTCGCGAACTCCAACGGCGCCATTTGCTCTGGTGCGGGTGACGCCC
       170       180       190       200       210       220       230       240
CAGCCGGCACGATCTGCCCCAAAGCTGGGGACATGGCTGCTGGCGACTGCCACCCGTACTTGCCCTCGTACGACGGCTCCC
       250       260       270       280       290       300       310       320
AGTGCAAGGCCAAGGAGGACGCCGTGTGTGCCATCGTGACCGGCAACACGTGGGGCTGCGTCTTCCCGTCCGTGGGCTGCC
       330       340       350       360       370       380       390       400
ACGACAGCGCCACGATGCCGGGGGCTCAGCCGCCGACGTCGGACGTCAACATGTCCATGGACCCTCACATGCCGGGCCCTA
       410       420       430       440       450       460       470       480
AGATGACGGACGCCCACACCATGCCGGGTGGTGTCCCTGCGACTCCTGAGACTCCGTCCGGAATGACGCCCTCGGTCCTGC
       490       500       510       520       530       540       550       560
ACGACCCAAGTGCTGGTGGAAACCAGTACATGCCGACTACAATGGCTCCTGTAATCCCGGGGTACGGCGACACGCTCCATG
       570       580       590       600       610       620       630       640
AAAAAATGCCCCACGCCGAGGGTAAAGGAGTTGGTGGTTCCCCGGCGTCAATGACGACCCCAGAAGCGACCCCCACGGCGA
       650       660       670       680       690       700       710       720
CGAATCCGGAGACCAAGACGACGAGCTCGGAGAACATGATGACGCCGCCAACTGTCCCTATGACGCCGCCAACTGCCCCCA
       730       740       750       760       770       780       790       800       810
TGACGATGCCAGCGACAACCCCCGAGATGCCAGCTGCTACTCCTGAGATCGAGACGGACCCGATGCTGACGCCTGTTTCCC
       820       830   834
AGTAGAAGGGCGAATTCGACCCAG
```

Figure 4.4c. Sequence and orientation confirmation of *Hpa* 814231 in pCR8/GW/TOPO plasmid. Multiple forward and reverse sequence reads were aligned and matched with the original sequence and flanking nucleotides were compared to the plasmid map. Sequences underlined with green bars prove the insertion in the correct orientation.

```
  1         10        20        30        40        50        60        70        80
TTGTACAAAAAAGCAGGCTCCGAATTCGCCCTTATGCGCCTGATCGGTCTTATGTTCTTCACAGTGGCGACCCTAGCTGCA
        90        100       110       120       130       140       150       160
ACCAGCGCGAGCACCACCGGTGCCGAGGGACGCCGTAGCCAAAAAAAGAACTCTGCAGAACCTGAAGACGAATCAAGTTTG
       170       180       190       200       210       220       230       240
GAGGATTCTTCGAGTGCGTTGGCTGCTGCTTATGTCAAACTTATGAACGGGCATGTCAGAGGGTTTCGCAAGGATCACCCG
       250       260       270       280       290   295
CTACCTAACAAATAGAAGGGCGAATTCGACCCAGCTTTCTTGTACAAAGTTG
```

Figure 4.4d. Sequence and orientation confirmation of *Hpa* 814014 in pCR8/GW/TOPO plasmid. Multiple forward and reverse sequence reads were aligned and matched with the original sequence and flanking nucleotides were compared to the plasmid map. Sequences underlined with green bars prove the insertion in the correct orientation.

```
    1          10          20          30          40          50          60          70          80
ÀAAAGCAGGĊTCCGAATTĊGCCCTTATGAGT̈TCGGTGGTĊTTGAAGGCAG̈CAGTATTTCȦCGTCTCGGAȦTCTTTCTTGG̈T
            90         100         110         120         130         140         150         160
CGTCGTCAȦTTATGGCGTĊGGTGGTGGAG̈ATCGCACAGĊGAGAACCATT̈GTCAACACGȦGGACGTACG̈CAAGTGAACG̈TGT
           170         180         190         200         210         220         230         240
TCGCAGCG̈AGGATAACT̈GCTTGCGCAȦTGATATCGCT̈TATAAGCTCGȦTCTGGTTATG̈TCACGAAGAĊCCGAAAGCTĊCAC
           250         260         270         280         290         300         310         320
CGACTĊĊGCGACATCAG̈TTTGGATGCG̈TCGACAAAGȦGAAGCATTCȦAGCGGAGATȦCAGCAAGATT̈GTGACGCAGȦGAGA
           330         340         350         360         370         380         390         400
CGCGGT̈GTCTCCAGAĊTGGTCCACAG̈CAGGAAATCT̈TTTTTCAACȦTCCGTCAAAG̈TGTTTAAGTĊTGCAAGCGTĊCTCGC
           410         420         430         440         450         460         470         480
CCGAGAAATAGACGȦTCGCATGGAĊGAGTCAAGGȦCAGGAAGCGG̈CGATTCCAGĊAAAGTCGCGȦTCAGTCGGGȦAATCTC
           490         500         510         520         530         540         550         560
AAGĊTCAGGACCT̈TGCATGCGCTĊGTCGAGACTĊGAGATGAGAĊGACAAACGTĊGGACGTCGCĊTCGTAAAAGGGCGAATT
           570         578
CGȦCCCAGCTT̈
```

Figure 4.4e. Sequence and orientation confirmation of *Hpa* 813915 in pCR8/GW/TOPO plasmid. Multiple forward and reverse sequence reads were aligned and matched with the original sequence and flanking nucleotides were compared to the plasmid map. Sequences underlined with green bars prove the insertion in the correct orientation.

The genes inserted in pCR™8/GW/TOPO® vector were then transferred into pEG100 vector with LR Reaction. Successfully transformed plasmids were isolated from the surviving colonies and were transformed into electrocompetent *Agrobacterium* cells. Following the confirmation of insertion via another colony PCR on *Agrobacterium* cells, overnight cultures were prepared for all the candidates. Each culture was diluted to both O.D.$_{600}$= 0.5 and 0.25 and then were pressure-injected into tobacco leaves using *E. amylovora* (O.D.$_{600}$=0.1) as positive control as it is known to cause HR in tobacco plants. Empty pEG100 vector (inserted in *Agrobacterium*), 10mM MgCl$_2$ and sdH$_2$O were also pressure-injected as negative controls. A slight chlorosis was observed on all of the injection sites at both concentrations for each candidate 6-7 dai, whereas the positive control started showing signs of HR 1-2 dai (Fig 4.5a-d). These results suggested that the candidates are unlikely to cause an HR, since a more rapid reaction was expected in a non-host transient expression. This could

mean that the candidates may not be sufficient to cause a strong HR on their own or they can be capable of avoiding recognition by the plant receptors.
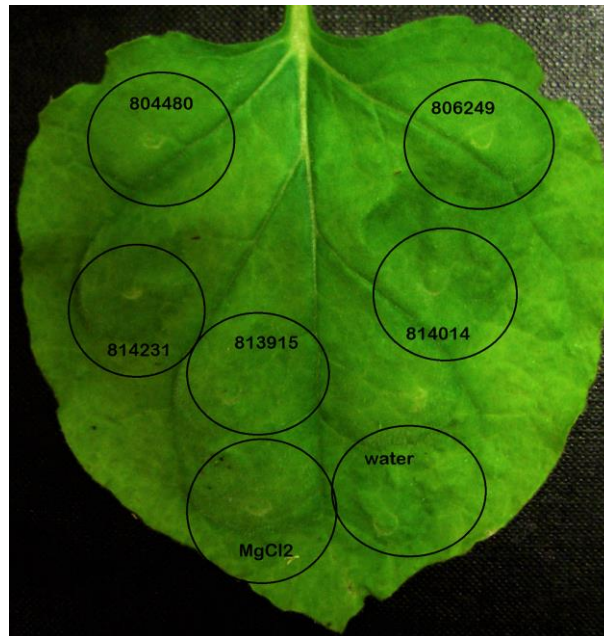
Figure 4.5a. Transient expression of the candidate genes in *N. benthamiana*. Candidate genes (O.D.$_{600}$=0.25) did not cause an HR within 5 dai on *N. benthamiana*.

Figure 4.5b. *E. amylovora* causing HR 1 dai on *N. benthamiana*. A clear chlorosis visible to naked eye was observed with the positive control E. amylovora 1 dai, which was an indication of hypersensitive response.
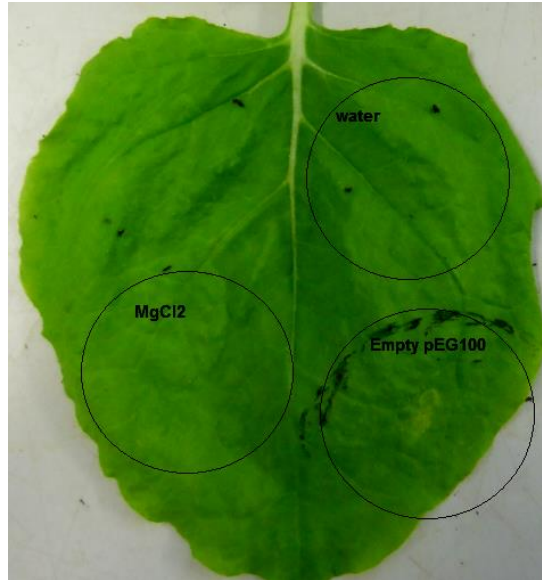
Figure 4.5c. The negative controls on *N. benthamiana* 5 dai. Negative controls water, MgCl$_2$ and empty pEG100 were injected into tobacco leaves. There was no sign of chlorosis even 5dai, as expected.
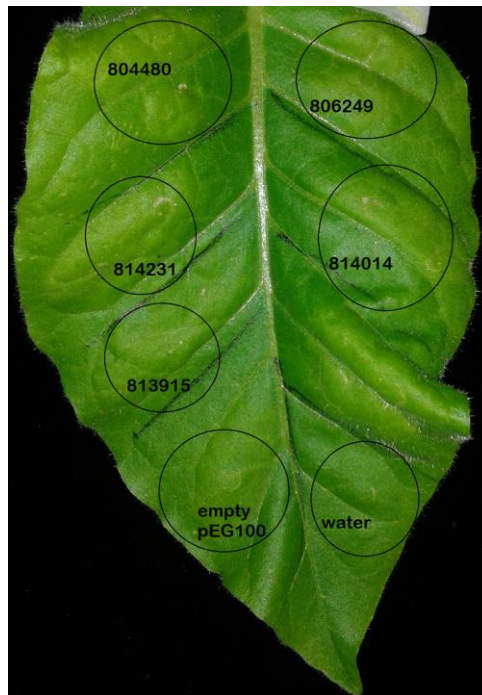
Figure 4.5d. Transient expression of the candidate genes in *N. tabacum*. A slight chlorosis appeared at injection sites of the candidates and empty pEG100 vector 7 dai on *N. tabacum.*

4.2.4. *Agrobacterium-mediated* stable transformation of *Arabidopsis*: Floral dipping

Wild-type Col-0 plants were transformed with the *Agrobacterium* cells carrying the pEG100 vector inserted with the candidate genes, following a modified version of Clough and Bent's (1998) protocol. The seeds of the transformed plants were collected and the transformation was confirmed via RT-PCR (Fig 4.6). The seeds were then grown to obtain the T_2 lines to use in assays for interaction phenotypes; however, it was not possible to grow seedlings without them catching secondary fungal infections despite the controlled environment and repetitions.
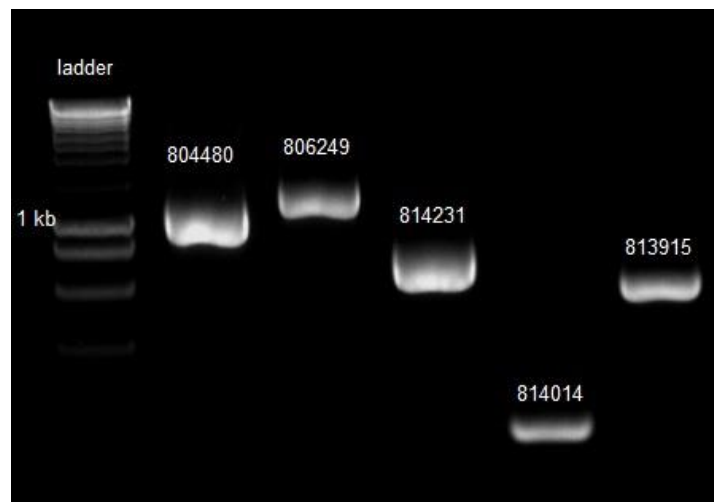
Figure 4.6. Transgenic lines were checked for inserts by RT-PCR. Presence of inserts indicated that the selected Col-0 plants were transformed with the candidate genes successfully.

4.2.5. Several approaches were taken to accomplish *in vitro* expression of the candidate genes

In order to assess whether the candidate genes were carrying effector-like features, it was sensible to express them *in vitro*, thus obtain the proteins they encoded to use in activity assays. Several approaches were taken to achieve this. First, was to insert the genes into an expression vector, pET28a (Fig. 4.7a), via restriction enzyme digestion and ligation method. Out of the multiple cloning site of the vector, *Eco*RI and *Xho*I restriction enzymes were selected (Fig. 4.7b). New forward and reverse primers were designed to remove signal peptide and stop codon sequences and to add *Eco*RI and *Xho*I restriction enzyme recognition sequences to 5' and 3' ends, respectively (Table 2.1b). This was initially performed on 2 of the candidates, *Hpa* 814231 and *Hpa* 814014, to check the viability of the method before carrying on with the longer candidates.
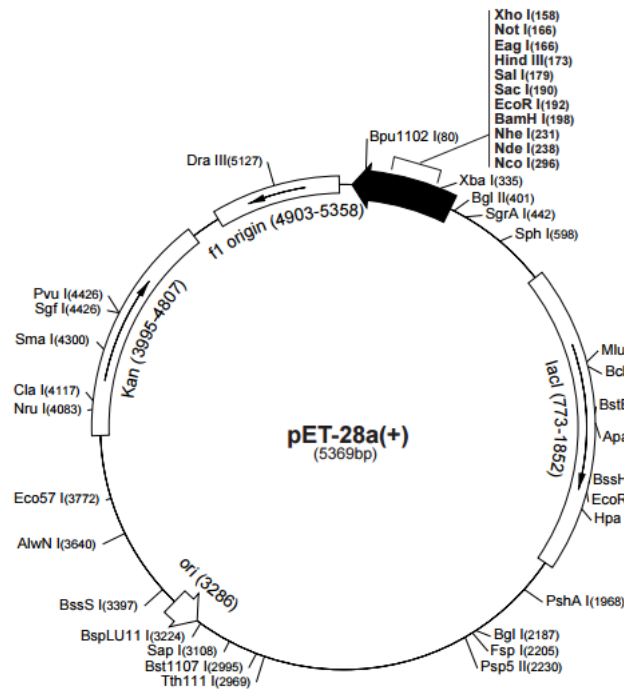
Figure 4.7a. The vector map of pET-28a. The pET-28a vector carry kanamycin resistance. (Source: Novagen)

pET-28a-c(+) cloning/expression region
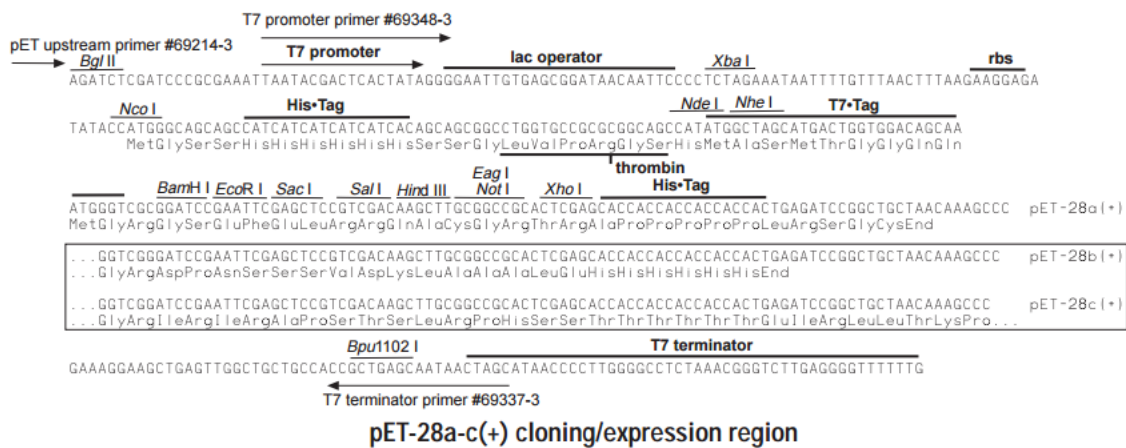
Figure 4.7b. The pET-28a cloning/expression region. The pET-28a vector carries an N-terminal His•Tag as well as an optional C-terminal His•Tag sequence. (Source: Novagen)

The genes of interest were amplified by PCR using the new primers, cleaned up and were first digested with *EcoR*I and then XhoI enzyme, along with the pET28a vector to create sticky ends for subsequent ligation steps. The gel images after digestions are shown below in comparison to untreated genes and the vector (Fig 4.8a-e).
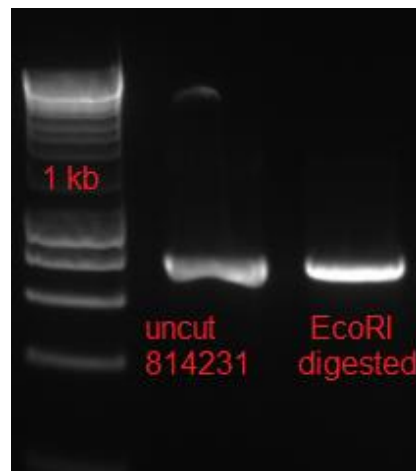
Figure 4.8a. Gel image of uncut *Hpa* 814231 vs *Eco*RI digested *Hpa* 814231. The treated gene fragment supposedly contained EcoRI recognition site added to 5' end.
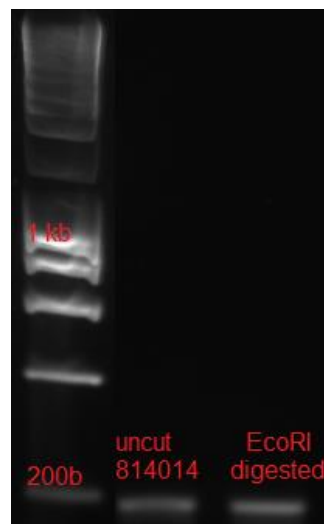
Figure 4.8b. Gel image of uncut *Hpa* 814014 vs *Eco*RI digested *Hpa* 814014. The treated gene fragment supposedly contained EcoRI recognition site added to 5' end.
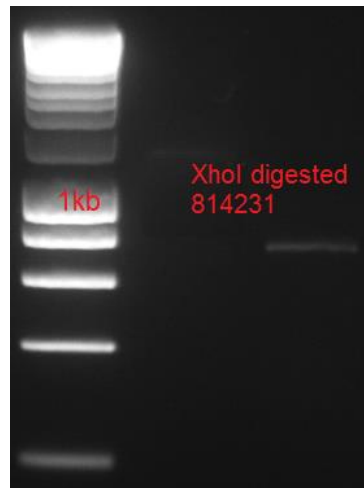
Figure 4.8c. Gel image of *Xho*I digested *Hpa* 814231. The treated gene fragment supposedly contained *Xho*I recognition site added to 3' end.
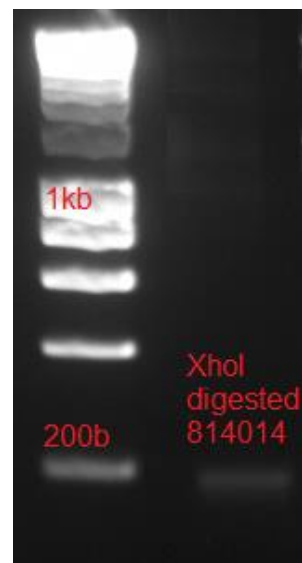
Figure 4.8d. Gel image of *Xho*I digested *Hpa* 814014. The treated gene fragment supposedly contained *Xho*I recognition site added to 3' end.
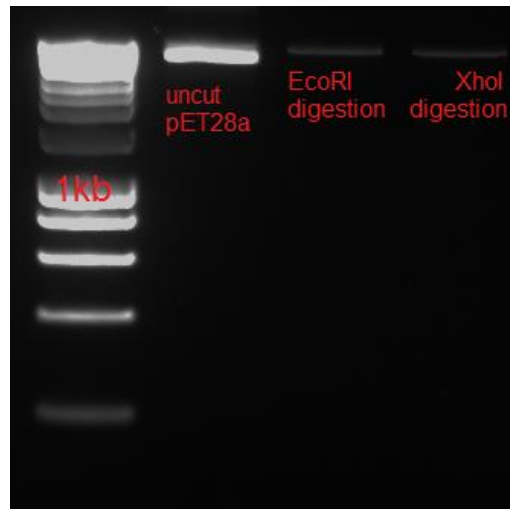
Figure 4.8e. Gel image of uncut pET28a vector vs *Eco*RI and *Xho*I digested digested pET28a. The 2 step restriction enzyme digestion was observed on the agarose gel. While uncut vector was circular, the digested vector was linear with sticky ends.

After ligation, few colonies were observed; however, colony PCR did not reveal successful insertions (Fig. 4.9a-b). The whole process was repeated several times with different enzymes stocks and alterations in digestion temperature and duration, to ensure the result was not due to human error and/or consumable malfunction.

Figure 4.9a. Colony PCR to check the insertion of *Hpa* 814231 into pET28a vector. Colonies were observed on the plate however the PCR did not show any amplification of the gene of interest.
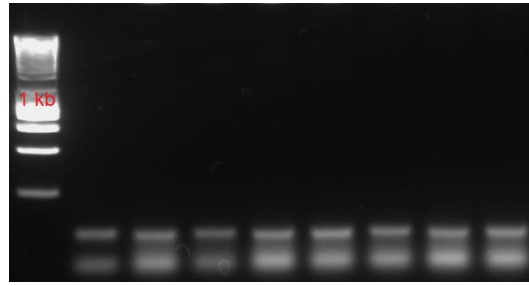
Figure 4.9b. Colony PCR to check the insertion of *Hpa* 814014 into pET28a vector. Colonies were observed on the plate however the PCR did not show any amplification of the gene of interest.

In the second attempt to express the genes, Gateway® cloning system was practiced. In this case instead of the original candidates, 2 cysteine-rich genes *Hpa* 813024 and *Hpa* 806256 that were chosen from MALDI-TOF screening of IWF (see section 5.2.1) were inserted into destination vector pDEST17™ (His-tagged, Ampicillin resistance) (Fig. 4.10). Additionally, one of the candidates, *Hpa* 814014, was also purchased as ready-synthesized (GenScript, USA) and inserted in pDEST17™ as well.

The ordered genes already contained the *att*B1/2 sequences and were cloned into BL21 (DE3) pLysE cells via BP and LR Cloning straight away. The insertions were sequence-verified (Fig. 4.11a-c) and were transformed into BL21 (DE3) pLysE cells.
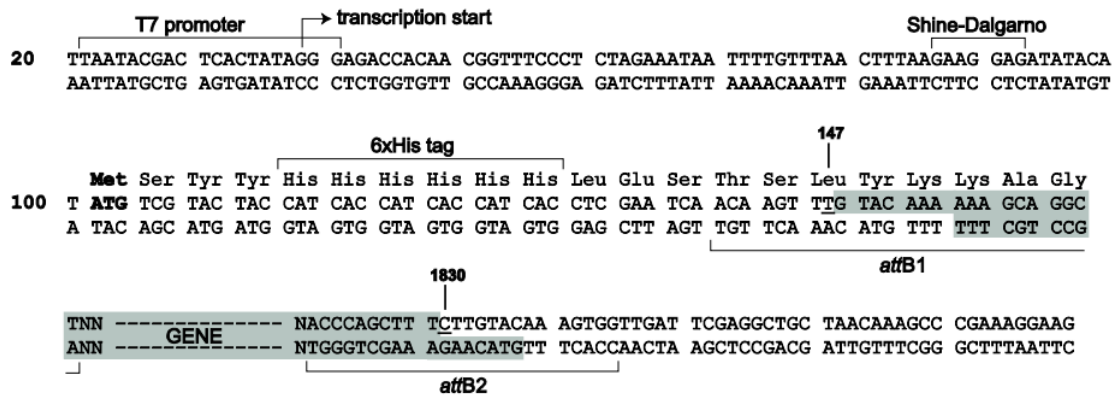
Figure 4.10. The pDEST17™ cloning/expression region. The pDEST17™ carry an N-terminal His•Tag and gene is inserted between *att*B1 and *att*B2 sequences. (Source: Invitrogen)

Figure 4.11a. Sequence and insertion verification of *Hpa* 813024. Green bar annotates the His-tag. Forward and reverse sequencing reads were referenced to the cloning map of pDEST17.

Figure 4.11b. Sequence and insertion verification of *Hpa* 806256. Green bar annotates the His-tag. Forward and reverse sequencing reads were referenced to the cloning map of pDEST17.

```
  1          10             20            30            40            50            60            70            80
CATCACCATCACCATCACCTCGAATCAACAAGTTTGTACAAAAAAGCAGGCTTCAGCACCACCGGTGCAGAAGGTCGTCGT
          90            100           110           120           130           140           150           160
AGCCAGAAAAAAACAGCGGCAGAACCGGAAGATGAAAGCAGCCTGGAAGATAGCAGCAGCGCACTGGCAGCAGCATATGTT
         170           180           190           200           210           220           230           240
AAACTGATGAATGGTCATGTTCGCGGTTTTCGTAAAGATCATCCGCTGCCGAACAAATAACACCCAGCTTTCTTGTACAAA
         250           260           270          279
GTGGTTGATTCGAGGCTGCTAACAAAGCCCGAAAGG
```

Figure 4.11c. Sequence and insertion verification of *Hpa* 814014. Green bar annotates the His-tag. Forward and reverse sequencing reads were referenced to the cloning map of pDEST17.

Furthermore, *Hpa* 814014 and *Hpa* 813024 were inserted into several other destination vectors i.e. GW-adapted pET29 and GW-adapted pET32, and transformed into BL21 (DE3) pLysE cells. Successfully transformed BL21 (DE3) pLysE cells (PCR done on colonies to amplify the genes for confirmation, data not shown) were then induced to produce encoded proteins by several combinations of varying parameters (Table 4.1).

Table 4.1. The parameters changed for the induction of *in vitro* protein production

Parameters	Variations
Induction temperature	16°C, 20 °C, 25 °C, 37 °C
Duration of induction	1 h, 2 h, 3 h, 4 h, 6 h, overnight
Final concentration of IPTG	0.2 mM, 1 mM
Final concentration of L-Arabinose	Zero, 0.2%
O.D.$_{600}$ of the culture before induction	0.4, 0.5, 0.8
Selective antibiotic	Ampicillin, Carbenicillin

After the inductions, induced cells were precipitated and the pellets were processed with different approaches to purify the proteins of interest: Using Ni-NTA Kit (Qiagen) under native conditions (Imidazole based) and under denaturing conditions (Urea and guanidine based); and using BugBuster® Protein Extraction Reagent (Merck Millipore). All treatments were passed through Ni-NTA Spin Columns (Qiagen) designed for purifying His-tagged proteins. The eluted collections were then visualized by SDS-PAGE, along with the samples obtained as flow-through and after wash steps as well as and non-induced cells. Unfortunately, none of the attempts revealed production of the proteins of interest. In order to prove that these outcomes were not due to human error and the followed methods; a positive control, BL21 (DE3) pLysE cells carrying pET28a plasmid inserted with *DspF* gene (a secretion chaperone encoding gene in *E. amylovora*, obtained from laboratory stock), were induced overnight at 20°C with a 0.2 mM IPTG final concentration. After the induction, the cells were processed under native condition using the Ni-NTA kit. The SDS-PAGE gel showed successful production of the DspF protein (Fig. 4.12).
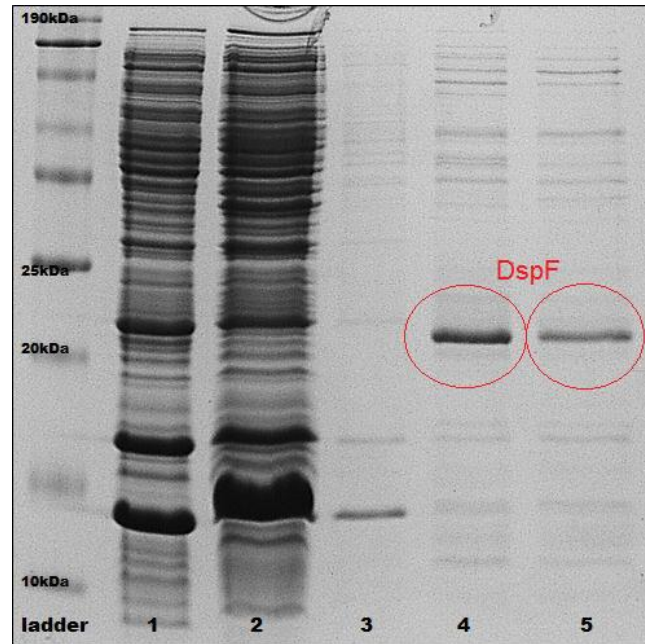
Figure 4.12. Induction of DspF production. Lane-1: non-induced cells, lane-2: flow-through, lane-3: wash and lanes- 4-5: Elutions. BL21 (DE3) pLysE cells carrying pET28a plasmid inserted with *DspF* gene were induced at 20°C with a 0.2 mM IPTG final concentration, overnight. The cells were processed under native condition using the Ni-NTA kit. The SDS-PAGE gel showed successful production of the DspF protein.

Besides all of the above, an additional approach, adapted from Kamoun et al. (1997) was taken to express the genes. In this case, target genes, ready-synthesized *Hpa* 814014 and *Hpa* 813024, were cloned into pFLAG-ATS (Fig. 4.13a-b), a vector that allows the product protein to be secreted out by transformed *E. coli*. The genes of interest were amplified with primers adding *Hind*III (forward) and *Eco*RI (reverse) recognition sites and *excluding* AttB1/2 sequences, signal peptide and the stop codon (Table 2.1d).
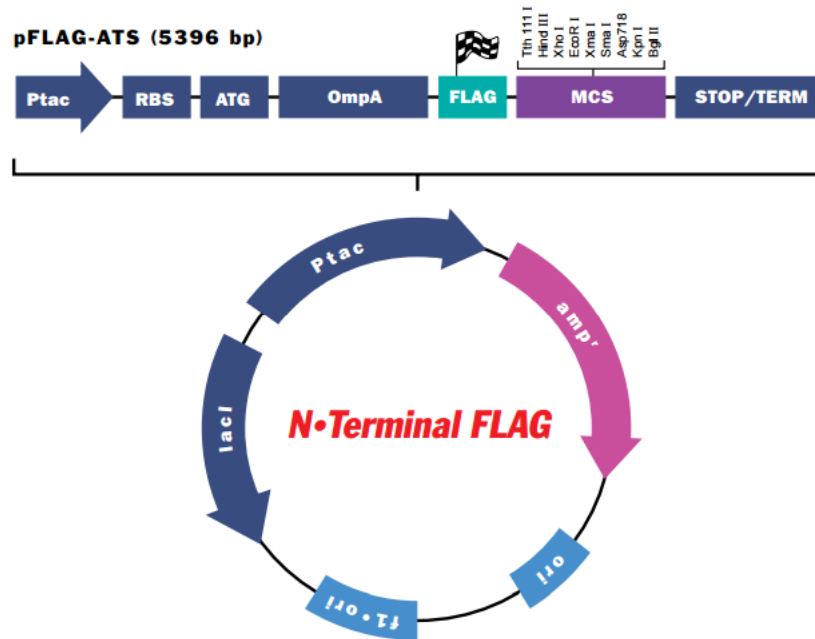
Figure 4.13a. The vector map of pFLAG-ATS. The pFLAG-ATS vector allows the product protein to be secreted out by transformed *E. coli*. (Source: Sigma-Aldrich)
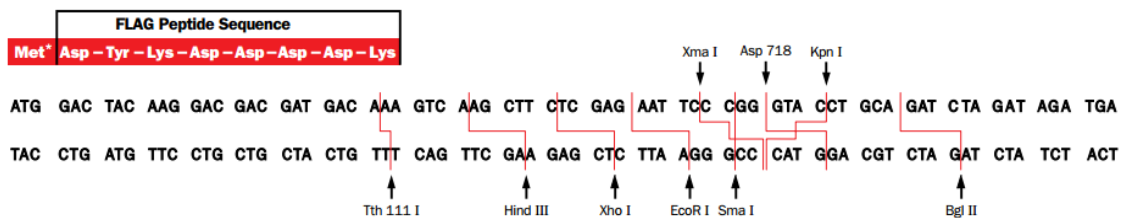
Figure 4.13b. The pFLAG-ATS cloning/expression region. The pFLAG-ATS carry an N-terminal FLAG peptide sequence. Multiple cloning site has 9 restriction enzyme recognition sites. (Source: Sigma-Aldrich)

The amplicons and the vector were double-digested with HindIII and EcoRI enzymes (Fig. 4.14a-b).
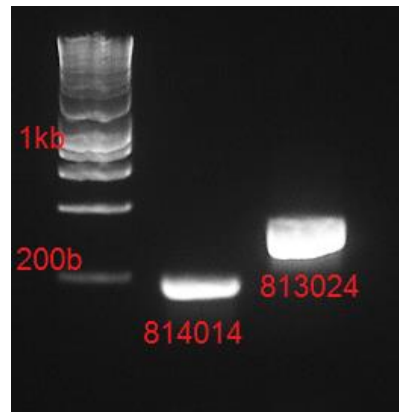
Figure 4.14a. Gel image of *Hpa* 814014 and *Hpa* 813024 after double digestion with *Eco*RI and *Hind*III. The treated gene fragment supposedly contained *Hind*III recognition site added to 5' end and *Eco*RI recognition site added to 3' end.
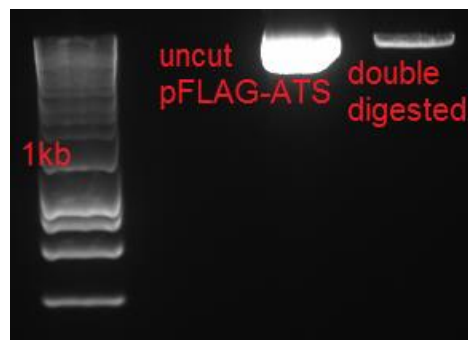
Figure 4.14b. Gel image of uncut pFLAG-ATS vector vs *Eco*RI and *Hind*III double-digested pFLAG-ATS. While uncut vector was circular, the digested vector was linear with sticky ends.

The vector and the products were ligated with T4 DNA Ligase. The ligation products were transformed into BL21 (DE3) pLysE cells and the insertion was confirmed by a colony PCR (Fig. 4.15a-b).
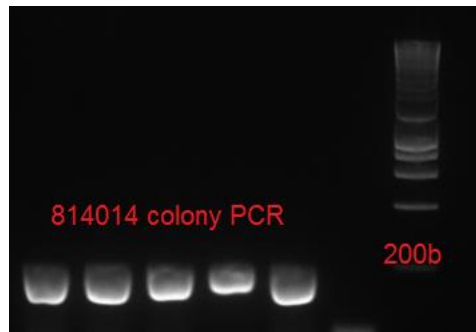
Figure 4.15a. Gel image of colony PCR of *Hpa* 814014 performed on transformed into BL21 (DE3) pLysE cells. The gene-of-interest was amplified in the transformed cells.
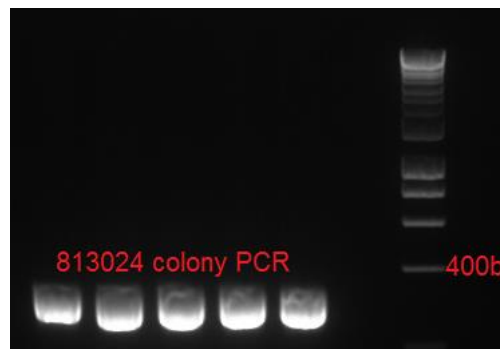
Figure 4.15b. Gel image of colony PCR of *Hpa* 813024 performed on transformed into BL21 (DE3) pLysE cells. The gene-of-interest was amplified in the transformed cells.

The transformed cells were induced at O.D.$_{600}$=0.6 with 0.4 mM IPTG final concentration. Induction was incubated overnight at 37°C. The induced culture was precipitated, this time the supernatant was collected. Supernatant was filter-sterilized to eliminate the *E. coli* cells and the proteins were precipitated with ANTI-FLAG® M2 Magnetic Beads. Unfortunately this approach also did not result in production of target proteins. Therefore, as a positive control, cells transformed with pFLAG-ATS carrying *EPIC* (kindly provided by Prof Sophien Kamoun's group, The Sainsbury Laboratory) were induced with same

conditions. SDS-PAGE showed successful induction of the EPIC protein (Fig. 4.16).
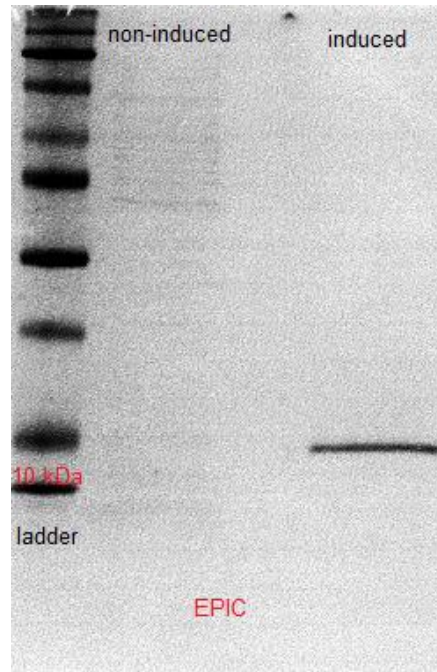
Figure 4.16. Induction of the *EPIC* gene. The BL21 (DE3) pLysE cells, that were transformed with pFLAG-ATS carrying *EPIC* gene, were induced at O.D.$_{600}$=0.6 with 0.4 mM IPTG final concentration. After overnight incubation at 37°C, proteins were precipitated with ANTI-FLAG® M2 Magnetic Beads. The SDS-PAGE showed a successful induction of production of the protein when compared to non-induced cells.

4.2.6. Candidate genes showed few variations across *Hpa* isolates

Avirulent effectors, particularly RxLR family effectors, are under persistent pressure to in an arms race in order to survive (Win et al., 2008; Allen et al., 2004; Haas et al., 2009). Conversely, LysM effectors and apoplastic Nep1-like proteins are widely conserved across pathogens (Gizjen and Nürnberger, 2006; Kamoun, 2006). PAMPs are also described as conserved molecules, therefore

Thomma et al. (2011) proposed that effectors showing a conserved nature across other isolates and pathogens may not be significantly different from PAMPs, and can also be detected by cell surface PRRs.

Being less predisposed to variations and mutations, essentially may increase the possibility of an effector of being recognized by host plant receptor; which in return supposedly would put more pressure to the pathogen to evolve. If the effector protein shows a trend towards keeping its conserved nature despite the evolutionary pressure, these proteins can be used for resistant plant engineering purposes (Bart et al., 2012).

Under the light of this information, here we wanted to analyse our candidate across various *Hpa* isolates to observe whether they were conserved or having variations. With the results of this analysis we hoped to have a better understanding on their recognition by the host plant, the evolutionary trend they follow, as well as to evaluate their potential for further use in generating resistant plants. With this purpose, candidates were sequence-analysed across available *Hpa* isolates, and were compared to the reference sequence of *Hpa*-Emoy2. In general, the results showed most of the candidates were conserved across isolates, and few SNPs resulted in non-synonymous mutations.

Eight isolates of *Hpa* (Cala2, Noks1, Emco5, Edco1, Goco1, Maks9, Hiks1, and Emwa1) were selected from the available laboratory stocks to analyse the polymorphisms of the candidate genes across them. To do that, 7-day-old Ws-*eds1* seedling were inoculated with these isolates. Seven days after inoculation, DNA was isolated from the plant materials and used as template for the PCR

(touchdown with Elongase® enzyme mix) to amplify candidate genes using flanking primers. The products were visualized and on a 2% agarose gel and the corresponding bands were excised and cleaned-up for sequencing. The sequencing results were manually analysed using Geneious (R6) to detect SNPs and translations were compared to the protein sequences of the candidates on *Hpa*-Emoy2 (See appendix 1 for SNPs). It was not possible to obtain conclusive sequencing results for the candidate *Hpa* 806249 despite repetitions.

Nucleotide alignments showed few SNPs across isolates, and most of them were synonymous mutations. Non-synonymous mutations were analysed in detail to assess whether the amino acids were changing into amino acids with different biochemical properties and affecting the general structure of the protein. For *Hpa* 814014, *Hpa*-Cala2 and *Hpa*-Emwa1 both possessed a change from Valine to Leucine at the same position within the signal peptide sequence. In the coding sequence *Hpa*-Emco5, *Hpa*-Edco1 and *Hpa*-Goco1 had a change of Glutamate to Aspartate at the same position (Fig. 4.17a).
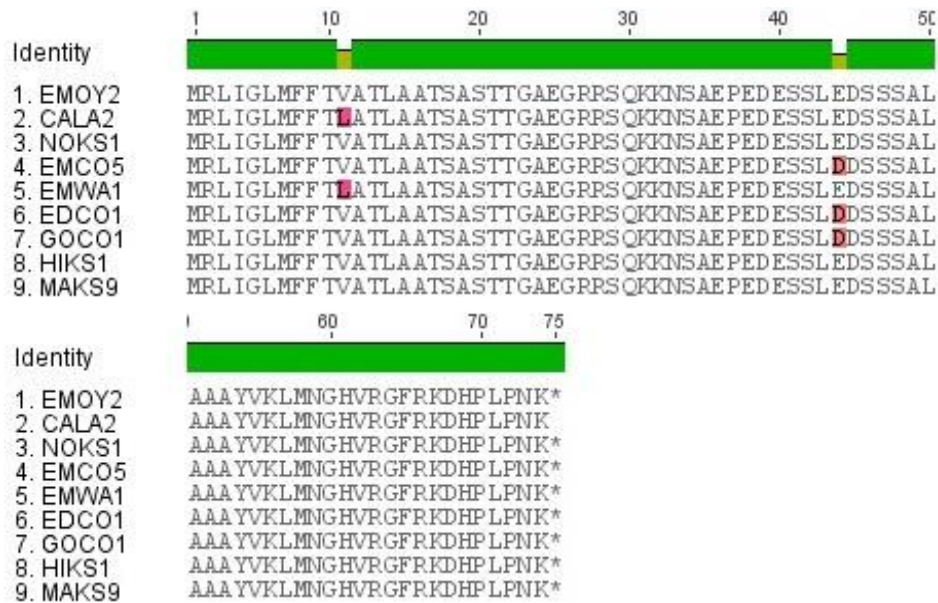
Figure 4.17a. Allelic variation in *Hpa* 814014 gene across isolates. *Hpa*-Cala2 and *Hpa*-Emwa1 showed a change from Valine to Leucine at the same position within the signal peptide sequence. In the coding sequence *Hpa*-Emco5, *Hpa*-Edco1 and *Hpa*-Goco1 had a change of Glutamate to Aspartate at the same position.

In case of *Hpa* 804480, it was possible to see a change of Threonine to Serine within the SP of *Hpa*-Cala2, *Hpa*-Noks1, *Hpa*-Edco1, *Hpa*-Goco1, *Hpa*-Hiks1 and *Hpa*-Maks9 at the same position. *Hpa*-Maks9 showed another difference within the SP, which was from Leucine to Phenylalanine. Same isolates, showed two additional differences in the coding sequence at the same position, where a Proline became Threonine and Lysine became Arginine. *Hpa*-Edco1, *Hpa*-Goco1, *Hpa*-Hiks1 and *Hpa*-Maks9 shared a difference at the same position from Glycine to Asparagine. Moreover, Noks1 carried different amino acids than the reference sequence due to a change from Glutamate to Glycine, Isoleucine to Valine, and Glycine to Serine (Fig. 4.17b).
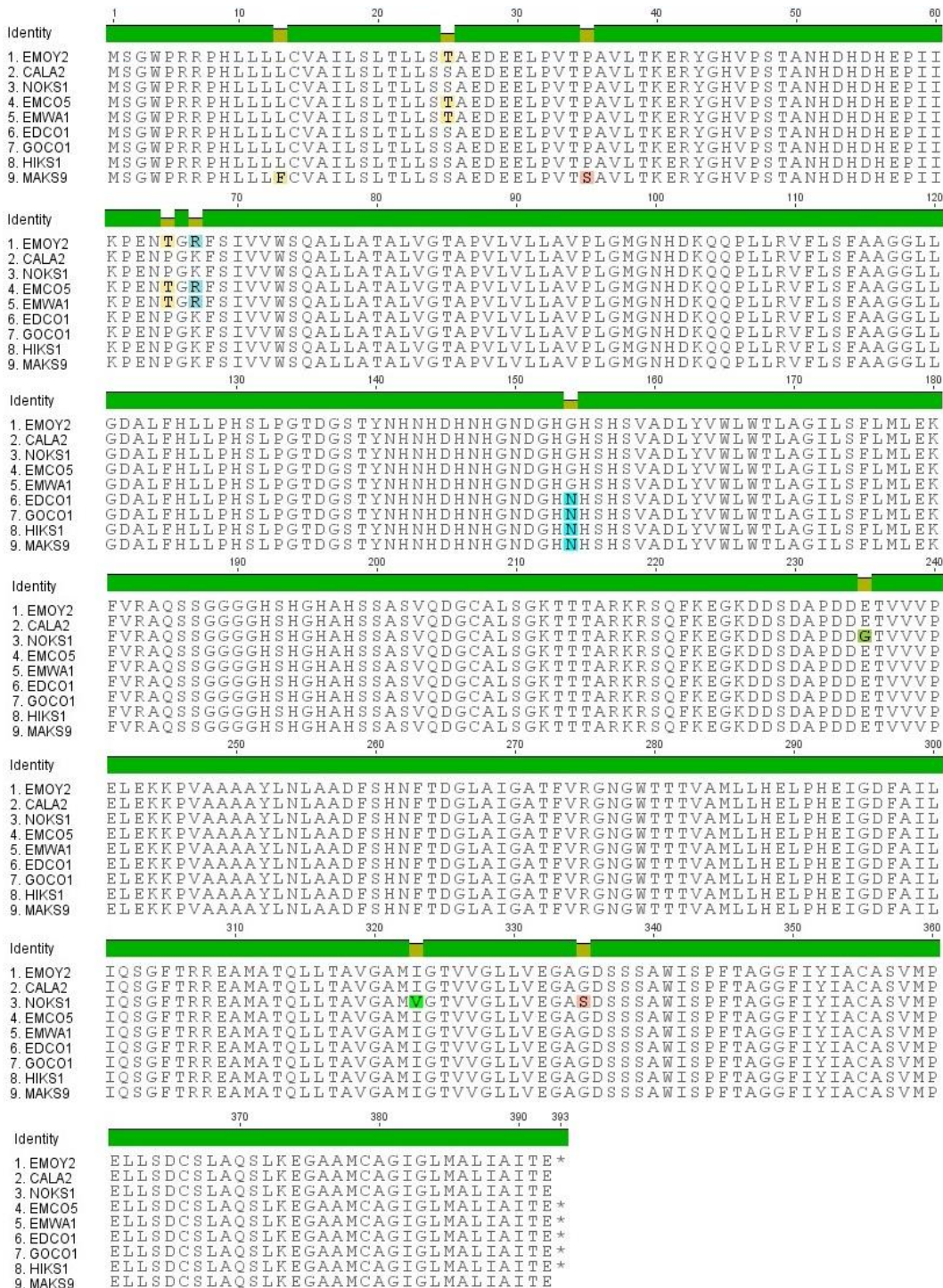
Figure 4.17b. Allelic variation in *Hpa* 804480 gene across isolates. *Hpa*-Cala2, *Hpa*-Noks1, *Hpa*-Edco1, *Hpa*-Goco1, *Hpa*-Hiks1 and *Hpa*-Maks9 had a change from Threonine to Serine within the SP at the same position. *Hpa*-Maks9 had a change from Leucine to Phenylalanine in SP. Proline became Threonine and Lysine became Arginine for the same isolates. *Hpa*-Edco1, *Hpa*-Goco1, *Hpa*-Hiks1 and *Hpa*-Maks9 shared a difference at the same position from Glycine to Asparagine. *Hpa*-Noks1 carried different amino acids than the reference sequence due to a change from Glutamate to Glycine, Isoleucine to Valine, and Glycine to Serine.

In case of *Hpa* 814231, there were also few non synonymous mutations. For instance *Hpa*-Noks1 possessed a Serine instead of Proline within the SP. *Hpa*-Noks1, *Hpa*-Emco5, *Hpa*-Edco1, *Hpa*-Goco1 and *Hpa*-Hiks1 had a Proline instead of a Threonine in the coding sequence. One another mutation carried by *Hpa*-Goco1 was a change of Proline to Serine (Fig. 4.17c).

Lastly, for *Hpa* 813915, *Hpa*-Noks1 had a Glutamate instead of Lysine and Arginine instead of a Threonine within the coding sequence. *Hpa*-Cala2 and *Hpa*-Maks9 shared a change at the same position from Serine to Arginine in the coding sequence. *Hpa*-Cala2 also had a switch from Serine to Proline, and *Hpa*-Maks9 had replaced Serine with a Cysteine towards the C-terminus. An interesting mutation observed was in *Hpa*-Edco1 and *Hpa*-Goco1 isolates, in which a Serine switched to a stop codon towards C-terminal, at the same position (Fig. 4.17d)

According to the results, it can be said that, candidates were mostly conserved across isolates, and few mutations were observed, of some carried by more than one isolate. Nevertheless, in general the amino acids were switched to such amino acids carrying similar biochemical features.
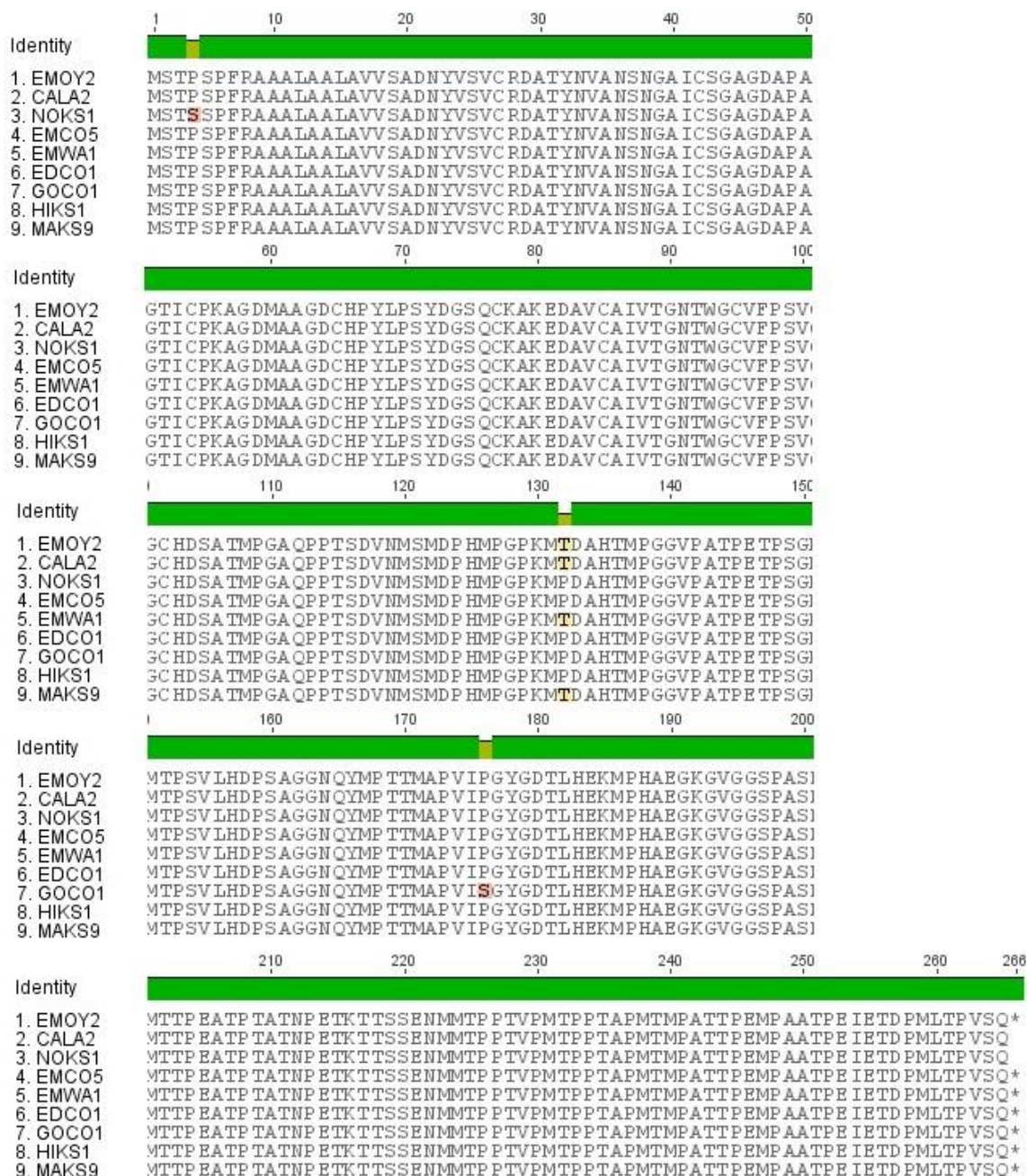
Figure 4.17c. Allelic variation in *Hpa* 814231 gene across isolates. *Hpa*-Noks1 had a Serine instead of Proline within the SP. *Hpa*-Noks1, *Hpa*-Emco5, *Hpa*-Edco1, *Hpa*-Goco1 and *Hpa*-Hiks1 had a Proline instead of a Threonine in the coding sequence. *Hpa*-Goco1 also had a change of Proline to Serine.
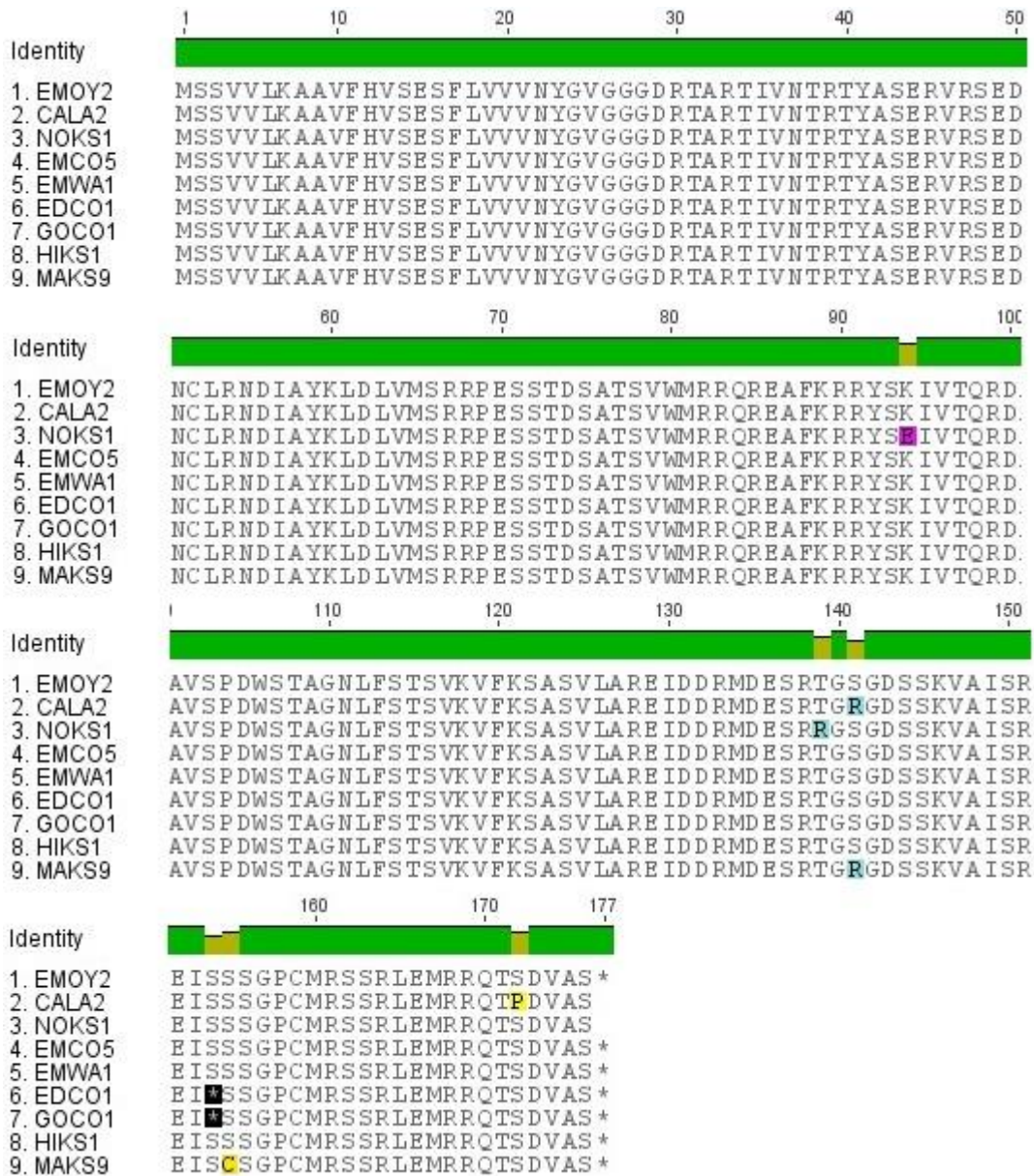
Figure 4.17d. Allelic variation in *Hpa* 813915 gene across isolates. *Hpa*-Noks1 had a Glutamate instead of Lysine and Arginine instead of a Threonine within the coding sequence. *Hpa*-Cala2 and *Hpa*-Maks9 had a change at the same position from Serine to Arginine in the coding sequence. *Hpa*-Cala2 had a switch from Serine to Proline, and *Hpa*-Maks9 had Cyteine instead of Serine towards the C-terminus. In *Hpa*-Edco1 and *Hpa*-Goco1 a Serine switched to a stop codon towards C-terminal, at the same position.

4.3. Summary

In this chapter the purpose was to use the candidate genes in gene expression assays and functionality assays to investigate their roles in defence suppression/activation in host plant. With this intention, DNA was isolated from infected Ws-*eds1* seedlings and the candidate genes were amplified via PCR, and candidates' presence in *Hpa*-Emoy2 isolate were proven. After observing amplification of the candidate genes, they were cloned for subsequent *in vitro* and *in planta* expression assays.

In order to check when and at what level the candidate genes were expressed by the pathogen after inoculation, samples were collected from infected material with 24 h intervals. RT-PCR (touchdown) was on the samples. In the results of RT-PCR, it was possible to observe many non-specific bands on the gels, which could have resulted from operating conditions and non-specific binding of the primers to the cDNAs. This occurrence of non-specific bands was observed in the repetitions of the experiments; therefore we focused on the bands-of-interest. The level and the timing of the expression were different among the candidate genes, suggesting that some encoded proteins by the candidate genes can be put forward by the pathogen right away or as a second layer of attack to suppress the PTI later during the infection. On the other hand, the proteins can be produced in an increasing manner (upregulation) to establish pathogenicity by evading the defence responses or they can be recognized during the infection which may result in a decrease in the expression level (downregulation).

In order to observe whether our candidate genes were causing any HR on the non-host plant: *N.* benthamiana and *N. tabacum*, the genes were transiently expressed using *Agrobacterium*-mediated expression. We expected to see rapid chlorosis as a result of HR, however, none of the candidates showed signs of HR. These results could mean either the candidates genes are not expressed within tobacco (although RT-PCR results indicated that they do express, data not shown), or tobacco plants are not the suitable plants for our candidates to trigger HR, the candidates do not trigger HR at all but instead induce other defence responses such as callose deposition or ion flux.

With the purpose of creating transgenic lines carrying our candidate effector genes, wild-type Col-0 plants were transformed with the candidates following *Agrobacterium*-mediated stable transformation method. The transformation step went smoothly, however, the next generation seeds were repeatedly catching secondary fungal infections. Therefore, susceptibility/resistance assays were not possible to carry out with transformed plants.

In this study, we also attempted to express the candidate genes *in vitro*, so that we can produce the encoded proteins and use them in assays to evaluate their roles in host plant defence activation. With pET28a expression system or Gateway systems, it was not possible to express the candidate genes. However, both pET28-DspF and pFLAG-ATS-EPIC were successfully induced as positive controls, suggestion the methodology for inductions were fine. (The results presented here belonged to the relatively shorter and/or ready-synthesized candidates. Remaining candidates also underwent cloning

and expression attempts with use of additional expression vectors, positive results were not obtained –data not-shown-).

Finally, the candidate genes were analysed for their variations across 8 other *Hpa* isolates. Overall, the results showed that the candidate genes were mostly conserved across isolates. There were few single nucleotide polymorphisms, although most of them were synonymous. The non-synonymous mutations mostly resulted in switches to amino acids with similar biochemical features, suggesting that the function of the protein may not be affected significantly. The conserved trend of the candidate genes also suggested that they might be perceived as PAMPs by the host plant, which also would explain why there was no HR observed.

5. A Proteomic approach towards identifying novel apoplastic effectors

5.1. Introduction

Oomycete plant pathogens initiate and establish pathogenesis by inhabiting in the host plant's intercellular regions (Kamoun, 2006). This apoplastic area is not only the space into which the pathogen secretes effectors, but also where the apoplastic effectors would find their corresponding receptors associated with plant's defence responses (Hammerschmidt, 2010). Previous studies have demonstrated that collecting the intercellular fluid from infected plant materials by vacuum infiltration-centrifugation method and scrutinizing the collection can reveal both pathogen and plant-associated molecules (Klement, 1965; Rathmell and Sequiera, 1974; Rathmell and Sequiera 1975; De Wit and Spikman, 1982; Boller and Métraux, 1988). Based on that, in addition to the bioinformatics approach for finding novel apoplastic effectors, we decided to take another approach, which was inspecting the intercellular washing fluid (IWF) from *Hpa*-Emoy2 infected Ws-*eds*1 plants.

As a sensible methodology to investigate the IWF in terms of contents, an analysis via Matrix-assisted laser desorption/ionization coupled to Time-of-flight (MALDI-TOF), a mass spectrometry technique, was chosen. To better understand the effects of the contents of the IWF on pathogenicity and defence,

certain assays i.e. β -glucuronidase (GUS) reporter system and cumulative quantification of Reactive Oxygen Species (ROS) were performed. The whole IWF was also fractionated using Ion-exchange chromatography method and to use those fractions in functionality assays as well, in order to narrow down the focus. Last but not least, the MALDI-TOF analysis allowed selection of T-DNA insertional mutant lines in order to observe the interaction phenotypes and see whether there was a change in resistance/susceptibility attributes.

Overall in this chapter, we looked deep into the apoplastic fluid, where the apoplastic effector and their targets presumably interact, with the intention of identifying novel apoplastic effectors and their receptors via a proteomics approach, using various assays to investigate the dynamics in terms of pathogenicity, defence, resistance and susceptibility.

5.2. Results

5.2.1. Intercellular washing fluid contains host and pathogen originated molecules

In order to obtain the dynamic source of apoplastic effectors, the fluid in the host intercellular area was collected, following Klement's (1965) method, from both healthy and *Hpa*-Emoy2-infected 5-6 week-old Ws-*eds1* seedlings. Washed plant leaves were vacuum-infiltrated and centrifuged. The collection fluids were filter-sterilized, quantified with Bradford assay and visualized on

12.5% SDS-PAGE. Both infected and healthy IWF samples showed a wide range of proteins (Fig. 5.1).
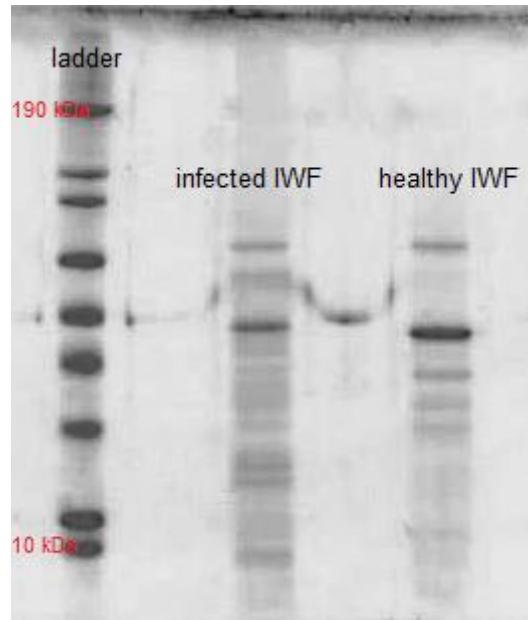
Figure 5.1. SDS-PAGE of infected and healthy IWF samples. A wide range of proteins were observed for both infected and healthy IWF samples. Different bands were between the infected and healthy samples were clearly visible. The protein ladder ranged from 10kDa to 190 kDa.

The samples then went through a proteomic analysis using Matrix-assisted Laser Desorption/Ionisation coupled with Time of Flight (MALDI-TOF) (by Prof Cyril Zipfel and his team at The Sainsbury Laboratory, Norwich/UK). The results revealed 106 *Hpa*–sourced proteins and best BLAST hits corresponding to annotated *Hpa* gene numbers are given in Table 5.1 (for complete amino acid sequences see Appendix 2).

According to the best BLAST hits, the majority of the putative proteins were showing resemblance to proteins of *P. infestans*, including hydrolases, serine

proteases and glucanases, which are strong candidates of being apoplastic effectors. There were a significant amount of proteins showing similarities to *P. sojae* proteins, however they were annotated as hypothetical proteins. Out of 106, there were 5 proteins that BLASTed to *H. arabidopsidis* itself. From those proteins *Hpa* 813024, *Hpa* 806256, *Hpa* 806254 and *Hpa* 811332 were annotated as Cysteine-rich proteins. In addition there was *Hpa* 806975, which was predicted to be an elicitin-like protein. Both cysteine-rice and elicitin-like proteins are known to belong to apoplastic effector families. Therefore, out of these 5 *Hpa*-sourced proteins, two putative cysteine-rich, signal peptide-carrying *Hpa* 813024 (Fig.5.2a-b) and *Hpa* 806256 (Fig.5.3a-b) were selected as additional candidates and were ordered as ready-synthesized from GenScript Inc. (USA) to be used for *in vitro* expression assays (see section 4.2.5.).

MHVKTSAVLLLATVASGPATFSSAWDTPGVPPQYTLPKYCGTAVPCDGWDQ
SASSICTYNKDANVEFKDCCLKNCVHATIRPGDKCSDVGVHCDQLNEVASQT
CGEQYYRVGDFARCCTLKCQSQWQH

Figure 5.2a. Amino acid sequence of *Hpa* 813024. *Hpa* 813024 was a 128 amino acid-long *Hpa*-sourced Cysteine-rich protein according to the MALDI-TOF results.
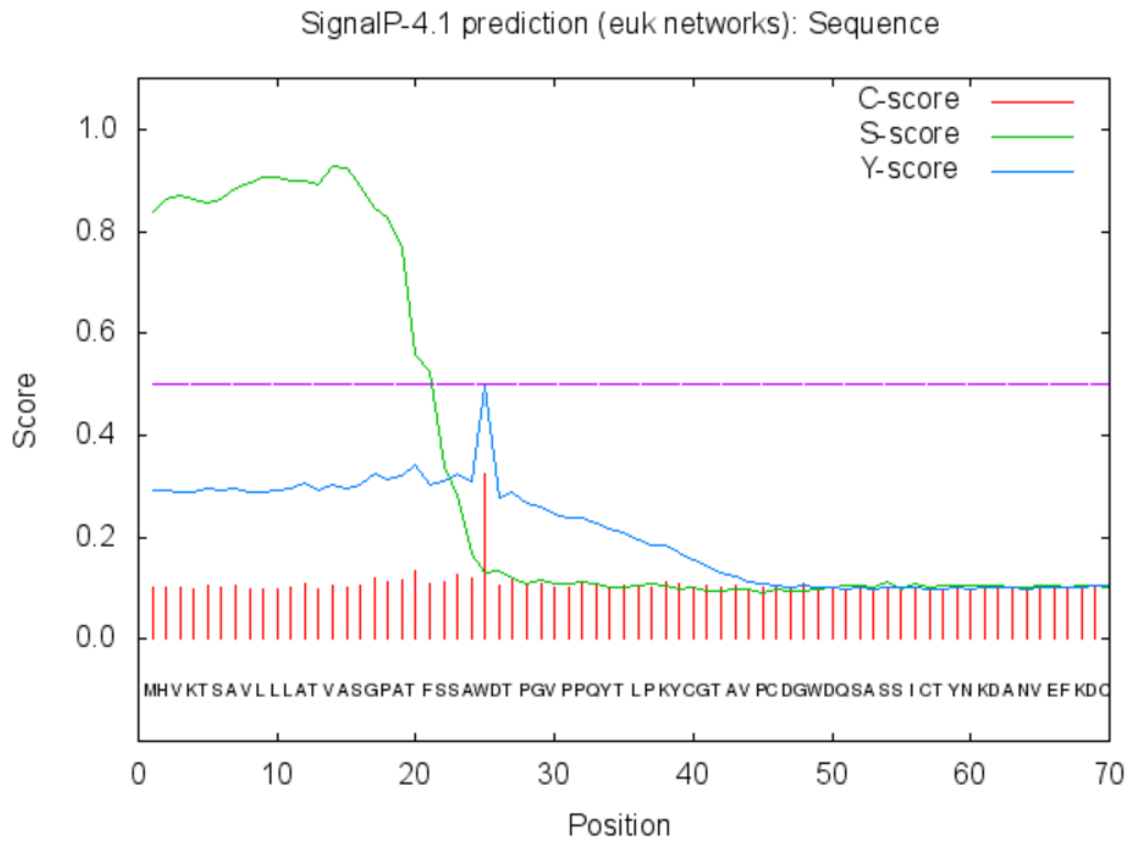
Figure 5.2b. The signal peptide cleavage site prediction for *Hpa* 813024. SignalP software predicted the cleavage site for *Hpa* 813024 as between amino acid 24 and 25.

MHVPSSLVLFMAAVASGSATEYAGGVSGSPSISPNTL**C**SAINVP**C**DGFDQDA SSF**C**SWQKDPNVEFKD**CC**LGR**C**SDARIPRGYK**C**SDSGVP**C**DQLNVDASQS**C** GVYETRVGQFRD**CC**IRE**C**RRQDQH

Figure 5.3a. Amino acid sequence of *Hpa* 806256. *Hpa* 806256 was a 127 amino acid-long *Hpa*-sourced Cysteine-rich protein according to the MALDI-TOF results.
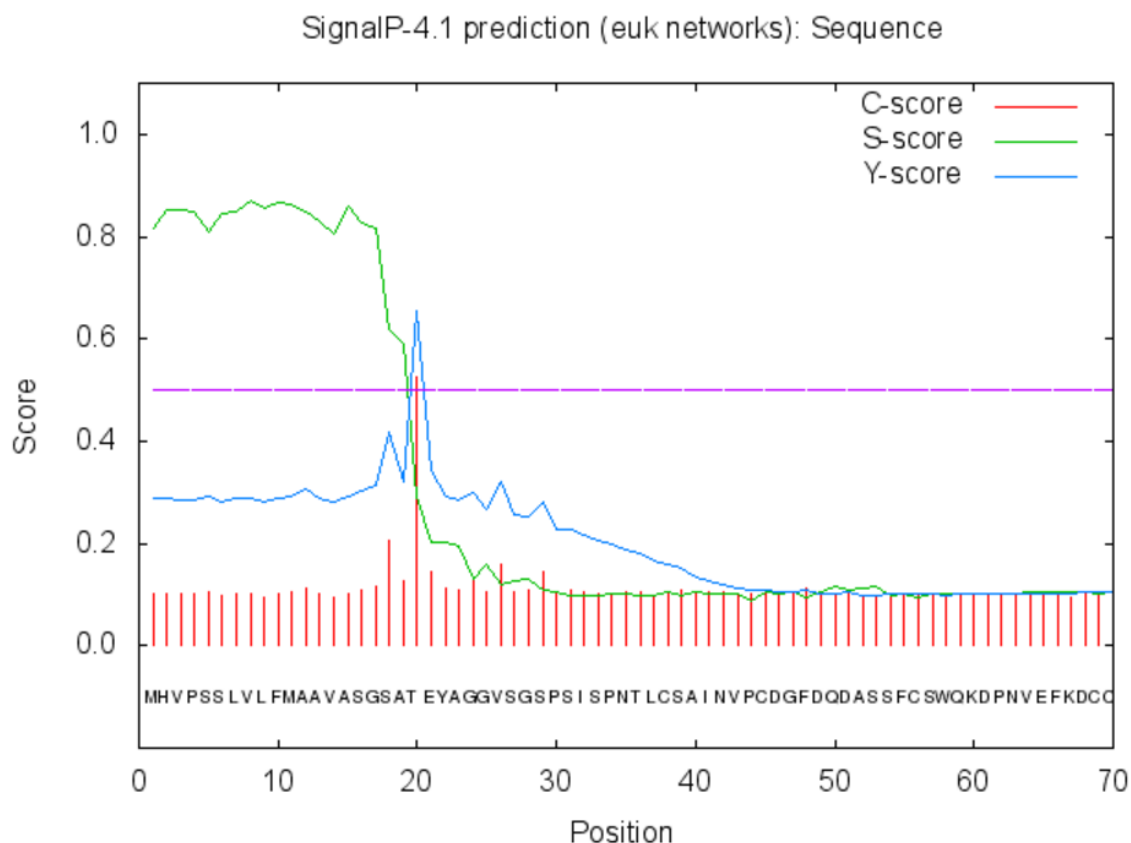
Figure 5.3b. The signal peptide cleavage site prediction for *Hpa* 806256. SignalP software predicted the cleavage site for *Hpa* 806249 as between amino acid 19 and 20.

Table 5.1. The *Hpa*-sourced results of MALDI-TOF with annotated Gene ID numbers and corresponding BLAST hits.

Hpa ID	BLAST best hits
800098	glycoside hydrolase, putative [*P. infestans*]
809344	hypothetical protein PHYSODRAFT_543825 [*P. sojae*]
814621	aldose 1-epimerase, putative [*P. infestans*]
809613	endo-1,3(4)-beta-glucanase, putative [*P. infestans*]
812220	serine protease family S33, putative [*P. infestans*]
808621	hypothetical protein PHYSODRAFT_518366 [*P. sojae*]
807735	aldose 1-epimerase, putative [*P. infestans*]
808748	metalloprotease family M17, putative [*P. infestans*]
803598	hypothetical protein PHYSODRAFT_362264 [*P. sojae*]
804884	serine protease family S33, putative [*P. infestans*]
806582	putative exo-1,3-beta-glucanase [*P. infestans*]

811928	chaperonin CPN60-1, mitochondrial precursor [*P. infestans*]
805873	hypothetical protein PHYSODRAFT_297590 [*P. sojae*]
808763	glutathione reductase [*P. infestans*]
807738	aldose 1-epimerase, putative [*P. infestans*]
803710	14-3-3 protein epsilon [*P. infestans*]
801564	hypothetical protein PHYSODRAFT_355326 [*P. sojae*]
811423	hypothetical protein PHYSODRAFT_520447 [*P. sojae*]
810555	hypothetical protein PHYSODRAFT_511802 [*P. sojae*]
806398	catalase [*P. infestans*]
808887	hypothetical protein PHYSODRAFT_485399 [*P. sojae*]
814664	glycoside hydrolase, putative [*P. infestans*]
806254	cysteine-rich protein [*H. arabidopsidis*]
806861	aconitate hydratase, putative [*P. infestans*]
803173	fructose-bisphosphate aldolase [*P. infestans*]
805299	hypothetical protein PHYSODRAFT_355994 [*P. sojae*]
805450	hypothetical protein PHYSODRAFT_478968 [*P. sojae*]
806256	cysteine-rich protein [*H. arabidopsidis*]
800014	glutathione S-transferase, putative [*P. infestans*]
800104	hypothetical protein PHYSODRAFT_287073 [*P. sojae*]
811239	nucleoside diphosphate kinase B [*P. infestans*]
802945	hypothetical protein PHYSODRAFT_284682 [*P. sojae*]
813525	hypothetical protein PHYSODRAFT_286509 [*P. sojae*]
802651	hypothetical protein PHYSODRAFT_561955 [*P. sojae*]
808043	hypothetical protein PHYSODRAFT_353705 [*P. sojae*]
803591	hypothetical protein PHYSODRAFT_355265 [*P. sojae*]
808107	elongation factor 2 [*P. infestans*]
810945	berberine-like protein [*P. infestans*]
801500	hypothetical protein PHYSODRAFT_481340 [*P. sojae*]
804167	inositol-3-phosphate synthase, putative [*P. infestans*]
801540	heat shock protein 70 [Phytophthora nicotianae]
808599	endoglucanase, putative [*P. infestans*]
801658	hypothetical protein PHYSODRAFT_347580 [*P. sojae*]
802132	hypothetical protein PHYSODRAFT_394585 [*P. sojae*]
807640	endoribonuclease L-PSP, putative [*P. infestans*]
809280	pectinesterase, putative [*P. infestans*]
813024	cysteine-rich protein [*H. arabidopsidis*]
811332	cysteine-rich protein [*H. arabidopsidis*]
806569	adenosylhomocysteinase [*P. infestans*]
806408	hypothetical protein PHYSODRAFT_498283 [*P. sojae*]
813371	hypothetical protein PHYSODRAFT_354299 [*P. sojae*]
810356	hypothetical protein PHYSODRAFT_467340 [*P. sojae*]

808503	hypothetical protein PHYSODRAFT_501475 [*P. sojae*]
814283	hypothetical protein PHYSODRAFT_295142 [*P. sojae*]
802113	hypothetical protein PHYSODRAFT_394585 [*P. sojae*]
813874	manganese superoxide dismutase [*P.nicotianae*]
803071	conserved hypothetical protein [*P. infestans*]
800270	family 31 glycoside hydrolase [*P. sojae*]
813080	hypothetical protein PHYSODRAFT_473521 [*P. sojae*]
804029	3-isopropylmalate dehydrogenase [*P. infestans*]
812913	peroxisomal acyl-coenzyme A oxidase, putative [*P. infestans*]
811591	hypothetical protein PHYSODRAFT_317226 [*P. sojae*]
800080	glycoside hydrolase, putative [*P. infestans*]
807220	Neprolysin CD1, peptidase family M13, neutral zinc metallopeptidase [*P. sojae*]
809066	protein disulfide-isomerase, putative [*P. infestans*]
802420	hypothetical protein PHYSODRAFT_289025 [*P. sojae*]
803975	hypothetical protein PHYSODRAFT_299659 [*P. sojae*]
814239	heat shock 70 kDa protein, mitochondrial precursor [*P. infestans*]
802267	elicitor-like mating protein M81 [*P. infestans*]
808833	serine protease family S10, putative [*P. infestans*]
812621	dihydrolipoyl dehydrogenase 1, mitochondrial precursor [*P. infestans*]
801302	hypothetical protein PHYSODRAFT_501339 [*P. sojae*]
809865	methylmalonate semialdehyde dehydrogenase [acylating], putative [*P. infestans*]
808413	conserved hypothetical protein [*P. infestans*]
806837	conserved hypothetical protein [*P. infestans*]
810322	glucose-6-phosphate isomerase [*P. infestans*]
807402	hypothetical protein PHYSODRAFT_542471 [*P. sojae*]
808045	calcineurin-like phosphoesterase, putative [*P. infestans*]
803469	conserved hypothetical protein [*P. infestans*]
804440	NAD-dependent malic enzyme, putative [*P. infestans*]
806975	elicitin-like protein [*H. arabidopsidis*]
809424	translation elongation factor 1-alpha, putative [*P. infestans*]
809723	6-phosphogluconate dehydrogenase [*P. infestans*]
812115	hypothetical protein PHYSODRAFT_555192 [*P. sojae*]
808637	hypothetical protein PHYSODRAFT_518702 [*P. sojae*]
803265	hypothetical protein PHYSODRAFT_534742 [*P. sojae*]
804695	family 30 glycoside hydrolase [*P. sojae*]
812337	putative GPI-anchored serine-threonine rich hypothetical protein [*P. infestans*]
807649	hypothetical protein PHYSODRAFT_355037 [*P. sojae*]
804180	hypothetical protein PHYSODRAFT_259407 [*P. sojae*]
807075	hypothetical protein PHYSODRAFT_554034 [*P. sojae*]
814377	cell 12A endoglucanase [Phytophthora ramorum]
804992	hypothetical protein PHYSODRAFT_354100 [*P. sojae*]

808954	hypothetical protein PHYSODRAFT_321069 [*P. sojae*]
804529	hypothetical protein PHYSODRAFT_285828 [*P. sojae*]
809219	cytochrome c [*P. sojae*]
802703	hypothetical protein PHYSODRAFT_537941 [*P. sojae*]
810520	short chain dehydrogenase, putative [*P. infestans*]
813189	hypothetical protein PHYSODRAFT_248120 [*P. sojae*]
806307	protease inhibitor EpiC2B [*P. infestans*]
811467	hypothetical protein PHYSODRAFT_539865 [*P. sojae*]
808734	proteasomal ubiquitin receptor ADRM1-like protein [*P. infestans*]
810603	conserved hypothetical protein [*P. infestans*]
809042	cyclophilin A [Phytophthora nicotianae]
810208	peptidyl-prolyl cis-trans isomerase CYP19-4 precursor [*P. infestans*]
811164	hypothetical protein PHYSODRAFT_286325 [*P. sojae*]

In addition to *Hpa*-sourced proteins the MALDI-TOF readings also revealed 243 *A. thaliana*-sourced proteins. Out of the 243 proteins, 159 were up-regulated (Table 5.2.) and 84 were down-regulated (Table 5.3) (See Appendix 3 for detailed predicted chromosomal locations of the proteins). In terms of regulation of gene expression, an up-regulation is a process by which a cell increases the quantity of one or more cellular proteins by increasing the expression of the genes encoding these proteins as response to an internal or external signal. Conversely, down-regulation is the process where the quantity of one or more cellular proteins is decreased resulting from decreased expression of the corresponding gene by the cell due to an external or internal variable. The MALDI-TOF analysis detected a wide range of both up and down-regulated proteins mostly acting in the intercellular area. However, out of these results, most attracting findings were the proteins predicted as LRR proteins that are known receptors for effectors (van der Hoorn and Kamoun, 2008; Bernoux et al., 2011). Since there is the possibility of those LRR proteins being cytoplasmic

effectors that contaminated the IWF sample due to human error, it was important to determine whether those proteins had signal peptides in their sequences to allow them to be secreted out of the cell and act freely in the apoplastic area or as cell surface receptors to interact with apoplastic effectors. To determine whether these proteins carried a signal peptide, the Gene ID numbers were carried over to The Arabidopsis Information Resource (TAIR) which is a database of genetic and molecular biology data for *A. thaliana* (http://www.arabidopsis.org/index.jsp). The protein sequences obtained from the database were examined with SignalP for signal peptides. After this screening, 3 LRR family proteins; 1 up-regulated AT1G33610.1 and 2 down-regulated AT3G20820.1 and AT1G49750.1 were selected for use in further assays. An up-regulation of an LRR protein with recognizing abilities would mean that in case of a pathogen attack the cell would be likely to produce more of this protein by increasing the expression levels of the encoding gene for an increased sensitivity and defence responses. On the other hand down-regulation of an LRR protein would suggest that their expression levels might depend on other players in plant defence responses (see 5.2.6 for assays carried out with the insertional mutants of the selected LRR proteins).

Overall, the results of the MALDI-TOF supported our theory of obtaining apoplastic pathogenesis and defence related proteins via collecting IWF.

Table 5.2. The *A. thaliana*-sourced results of MALDI-TOF listing up-regulated genes with annotated Gene ID numbers and corresponding protein BLAST hits.

AT1G26390.1	FAD-binding Berberine family protein
AT5G09590.1	mitochondrial HSO70 2
AT4G12290.1	Copper amine oxidase family protein
AT1G26380.1	FAD-binding Berberine family protein
AT4G26970.1	aconitase 2
AT4G01870.1	tolB protein-related
AT3G48000.1	aldehyde dehydrogenase
AT4G37520.1	Peroxidase superfamily protein
AT2G19500.1	cytokinin oxidase 2
AT1G66700.1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
AT1G30730.1	FAD-binding Berberine family protein
AT2G45220.1	Plant invertase/pectin methylesterase inhibitor superfamily
AT1G13750.1	Purple acid phosphatases superfamily protein
AT5G04360.1	limit dextrinase
AT3G23600.1	alpha/beta-Hydrolases superfamily protein
AT2G06050.1	oxophytodienoate-reductase 3
AT3G57260.1	beta-1,3-glucanase 2
AT5G20960.1	aldehyde oxidase 1
AT5G39580.1	Peroxidase superfamily protein
AT2G20420.1	ATP citrate lyase (ACL) family protein
AT5G38900.1	Thioredoxin superfamily protein
AT3G16400.1	nitrile specifier protein 1
AT5G11920.1	6-&1-fructan exohydrolase
AT4G01850.1	S-adenosylmethionine synthetase 2
AT3G22200.1	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein
AT5G07440.1	GDH2 glutamate dehydrogenase 2
AT4G14630.1	germin-like protein 9
AT5G18470.1	Curculin-like (mannose-binding) lectin family protein
AT2G37040.1	ammonia lyase 1
AT3G19010.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT1G06290.1	acyl-CoA oxidase 3
AT1G76680.1	12-oxophytodienoate reductase 1
AT2G30110.1	ubiquitin-activating enzyme 1
AT1G13060.1	PBE1 20S proteasome beta subunit E1
AT1G27020.1	unknown protein
AT4G19810.1	Glycosyl hydrolase family protein with chitinase insertion domain
AT1G78460.1	SOUL heme-binding family protein
AT2G27860.1	AXS1 UDP-D-apiose/UDP-D-xylose synthase 1
AT4G05180.1	PSII-Q photosystem II subunit Q-2
AT5G35100.1	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein
AT4G12490.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein

AT2G22970.1	SCPL11 serine carboxypeptidase-like 11
AT4G29260.1	HAD superfamily, subfamily IIIB acid phosphatase
AT4G21280.1	PSBQ-1 photosystem II subunit QA
AT2G26010.1	PDF1.3 plant defensin 1.3
AT4G16660.1	heat shock protein 70 (Hsp 70) family protein
AT1G50380.1	Prolyl oligopeptidase family protein
AT5G08300.1	Succinyl-CoA ligase, alpha subunit
AT1G26420.1	FAD-binding Berberine family protein
AT1G32960.1	ATSBT3.3, SBT3.3 Subtilase family protein
AT3G56650.1	Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein
AT5G12040.1	Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase family protein
AT2G21620.1	RD2 Adenine nucleotide alpha hydrolases-like superfamily protein
AT3G16460.1	Mannose-binding lectin superfamily protein
AT4G25100.1	FSD1, ATFSD1 Fe superoxide dismutase 1
AT5G41870.1	Pectin lyase-like superfamily protein
AT3G07390.1	AIR12 auxin-responsive family protein
AT1G30900.1	BP80-3;3 VACUOLAR SORTING RECEPTOR 6
AT1G53850.1	PAE1, ATPAE1 20S proteasome alpha subunit E1
AT1G15000.1	scpl50 serine carboxypeptidase-like 50
AT4G16155.1	dihydrolipoyl dehydrogenases
AT5G20830.1	ASUS1, atsus1 sucrose synthase 1
AT1G67550.1	URE urease
AT4G08770.1	Prx37 Peroxidase superfamily protein
AT1G08980.1	ATTOC64-I, TOC64-I amidase 1
AT4G34480.1	O-Glycosyl hydrolases family 17 protein
AT1G11580.1	PMEPCRA methylesterase PCR A
AT3G03640.1	GLUC, BGLU25 beta glucosidase 25
AT1G20510.1	OPCL1 OPC-8:0 CoA ligase1
AT5G54080.1	HGO homogentisate 1,2-dioxygenase
AT5G05600.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT3G06770.2	Pectin lyase-like superfamily protein
AT5G55480.1	SVL1 SHV3-like 1
AT3G61440.1	BSAS3;1, CYSC1 cysteine synthase C1
AT2G02390.1	ATGSTZ1, GST18, GSTZ1 glutathione S-transferase zeta 1
AT3G27890.1	NQR NADPH:quinone oxidoreductase
AT5G60640.1	ATPDI2, PDIL1-4 PDI-like 1-4
AT4G30910.1	Cytosol aminopeptidase family protein
AT3G10060.1	FKBP-like peptidyl-prolyl cis-trans isomerase family protein
AT3G11700.1	FLA18 FASCICLIN-like arabinogalactan protein 18 precursor
AT4G18250.1	receptor serine/threonine kinase, putative
AT1G02930.1	GSTF6, ATGST1 glutathione S-transferase 6
AT5G64120.1	Peroxidase superfamily protein
AT1G30700.1	FAD-binding Berberine family protein
AT1G02920.1	ATGST11 glutathione S-transferase 7

AT2G32240.1	FUNCTIONS IN: molecular_function unknown
AT4G16260.1	Glycosyl hydrolase superfamily protein
AT3G24503.1	REF1 aldehyde dehydrogenase 2C4
AT2G05710.1	ACO3 aconitase 3
AT1G21750.1	ATPDI5, PDI5, PDIL1-1 PDI-like 1-1
AT5G03630.1	Pyridine nucleotide-disulphide oxidoreductase family protein
AT3G08590.1	Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent
AT3G15730.1	PLD phospholipase D alpha 1
AT3G49120.1	PRX34 peroxidase CB
AT4G15530.1	pyruvate orthophosphate dikinase
AT5G54500.1	flavodoxin-like quinone reductase 1
AT4G34890.1	xanthine dehydrogenase 1
AT4G20850.1	TPP2 tripeptidyl peptidase ii
AT3G14310.1	pectin methylesterase 3
AT5G54960.1	pyruvate decarboxylase-2
AT3G16530.1	Legume lectin family protein
AT5G17530.1	phosphoglucosamine mutase family protein
AT1G22840.1	ATCYTC-A CYTOCHROME C-1
AT2G41220.1	GLU2 glutamate synthase 2
AT4G22470.1	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
AT2G43560.1	FKBP-like peptidyl-prolyl cis-trans isomerase family protein
AT4G20860.1	FAD-binding Berberine family protein
AT2G43590.1	Chitinase family protein
AT3G04720.1	PR-4 pathogenesis-related 4
AT4G16760.1	acyl-CoA oxidase 1
AT2G27190.1	purple acid phosphatase 12
AT5G13120.1	CYP20-2 cyclophilin 20-2
AT5G13420.1	Aldolase-type TIM barrel family protein
AT5G11670.1	NADP-malic enzyme 2
AT2G14610.1	pathogenesis-related gene 1
AT3G15356.1	Legume lectin family protein
AT5G62530.1	aldehyde dehydrogenase 12A1
AT5G50250.1	chloroplast RNA-binding protein 31B
AT3G28940.1	AIG2-like (avirulence induced gene) family protein
AT2G29350.1	senescence-associated gene 13
AT2G15220.1	Plant basic secretory protein (BSP) family protein
AT3G28930.1	AIG2-like (avirulence induced gene) family protein
AT3G62250.1	UBQ5 ubiquitin 5
AT1G33610.1	Leucine-rich repeat (LRR) family protein
AT4G26690.1	PLC-like phosphodiesterase family protein
AT2G43910.1	HOL1 HARMLESS TO OZONE LAYER 1
AT1G21680.1	DPP6 N-terminal domain-like protein
AT3G09440.1	Heat shock protein 70 (Hsp 70) family protein
AT4G37870.1	PEPCK phosphoenolpyruvate carboxykinase 1

AT1G16470.1	proteasome subunit PAB1
AT4G00570.1	NAD-dependent malic enzyme 2
AT1G13900.1	Purple acid phosphatases superfamily protein
AT3G54960.1	PDI-like 1-3
AT4G37530.1	Peroxidase superfamily protein
AT1G75040.1	PR-5 pathogenesis-related gene 5
AT2G44920.2	Tetratricopeptide repeat (TPR)-like superfamily protein
AT4G20830.2	FAD-binding Berberine family protein
AT5G56590.1	O-Glycosyl hydrolases family 17 protein
AT1G53310.1	phosphoenolpyruvate carboxylase 1
AT2G27150.1	abscisic aldehyde oxidase 3
AT5G17710.1	Co-chaperone GrpE family protein
AT4G03200.1	catalytics
AT3G51260.1	20S proteasome alpha subunit PAD1
AT4G35830.1	aconitase 1
AT5G37600.1	glutamine synthase clone R1
AT1G17290.1	alanine aminotransferas
AT5G06860.1	polygalacturonase inhibiting protein 1
AT1G07890.1	ascorbate peroxidase 1
AT5G04590.1	SIR sulfite reductase
AT4G12730.1	FASCICLIN-like arabinogalactan 2
AT2G13560.1	NAD-dependent malic enzyme 1
AT2G33120.1	synaptobrevin-related protein 1
AT5G07470.1	peptidemethionine sulfoxide reductase 3
AT5G45680.1	FK506-binding protein 13
AT5G60360.1	aleurain-like protease
AT1G27130.1	glutathione S-transferase tau 13
AT5G40370.1	Glutaredoxin family protein
AT1G79720.1	Eukaryotic aspartyl protease family protein
AT5G55450.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein

Table 5.3. The *A. thaliana*-sourced results of MALDI-TOF listing down-regulated genes with annotated Gene ID numbers and corresponding protein BLAST hits.

AT1G70730.1	Phosphoglucomutase/phosphomannomutase family protein
AT1G72150.1	PATELLIN 1
AT5G56010.1	heat shock protein 81-3
AT5G60600.1	4-hydroxy-3-methylbut-2-enyl diphosphate synthase
AT1G23190.1	Phosphoglucomutase/phosphomannomutase family protein
ATCG00120.1	ATP synthase subunit alpha
AT1G64190.1	6-phosphogluconate dehydrogenase family protein
AT3G20820.1	Leucine-rich repeat (LRR) family protein

AT5G65010.1	asparagine synthetase 2
AT2G40840.1	disproportionating enzyme 2
AT1G10760.1	Pyruvate phosphate dikinase, PEP/pyruvate binding domain
AT4G22670.1	HSP70-interacting protein 1
AT5G19770.1	tubulin alpha-3
AT2G36390.1	starch branching enzyme 2.1
AT3G22960.1	Pyruvate kinase family protein
AT1G28600.1	GDSL-like Lipase/Acylhydrolase superfamily protein
AT5G40450.1	unknown protein
AT5G19820.1	ARM repeat superfamily protein
AT4G39330.1	cinnamyl alcohol dehydrogenase 9
AT4G29840.1	Pyridoxal-5'-phosphate-dependent enzyme family protein
AT5G35360.1	acetyl Co-enzyme a carboxylase biotin carboxylase subunit
AT5G45390.1	CLP protease P4
AT5G57560.1	Xyloglucan endotransglucosylase/hydrolase family protein
AT3G25230.1	rotamase FKBP 1
AT2G46520.1	cellular apoptosis susceptibility protein/ importin-alpha re-exporter
AT3G09820.1	adenosine kinase 1
AT5G16390.1	chloroplastic acetylcoenzyme A carboxylase 1
AT3G05350.1	Metallopeptidase M24 family protein
AT3G18060.1	transducin family protein / WD-40 repeat family protein
AT3G26380.1	Melibiase family protein
AT1G20440.1	cold-regulated 47
AT1G49240.1	actin 8
AT3G48420.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
AT1G22530.1	PATELLIN 2
AT1G29900.1	carbamoyl phosphate synthetase B
AT5G26830.1	Threonyl-tRNA synthetase
AT5G01410.1	Aldolase-type TIM barrel family protein
AT2G41530.1	S-formylglutathione hydrolase
AT5G03650.1	starch branching enzyme 2.2
AT2G30200.1	catalytics;transferases;[acyl-carrier-protein] S-malonyltransferases
AT5G64050.1	glutamate tRNA synthetase
AT4G25370.1	Double Clp-N motif protein
AT1G20950.1	Phosphofructokinase family protein
AT2G35840.1	Sucrose-6F-phosphate phosphohydrolase family protein
AT5G65730.1	xyloglucan endotransglucosylase/hydrolase 6
AT2G21660.1	cold, circadian rhythm, and rna binding 2
AT5G23120.1	photosystem II stability/assembly factor, chloroplast (HCF136)
AT1G15140.1	FAD/NAD(P)-binding oxidoreductase
AT4G18810.1	NAD(P)-binding Rossmann-fold superfamily protein

AT5G57550.1	xyloglucan endotransglucosylase/hydrolase 25
AT5G53480.1	ARM repeat superfamily protein
AT5G41670.1	6-phosphogluconate dehydrogenase family protein
AT1G49750.1	Leucine-rich repeat (LRR) family protein
AT2G34810.1	FAD-binding Berberine family protein
AT4G09670.1	Oxidoreductase family protein
AT2G27920.1	serine carboxypeptidase-like 51
AT4G27450.1	Aluminium induced protein with YGL and LRDR motifs
AT1G27400.1	Ribosomal protein L22p/L17e family protein
AT1G18270.1	ketose-bisphosphate aldolase class-II family protein
AT1G11750.1	CLP protease proteolytic subunit 6
AT1G12410.1	CLP protease proteolytic subunit 2
AT5G55730.1	FASCICLIN-like arabinogalactan 1
AT4G23500.1	Pectin lyase-like superfamily protein
AT2G25080.1	glutathione peroxidase 1
AT1G06570.1	phytoene desaturation 1
AT2G20890.1	photosystem II reaction center PSB29 protein
AT1G13930.1	Involved in response to salt stress
AT2G06850.1	xyloglucan endotransglucosylase/hydrolase 4
AT2G44160.1	methylenetetrahydrofolate reductase 2
AT5G64260.1	EXORDIUM like 2
AT1G05560.1	UDP-glucosyltransferase 75B1
AT1G03230.1	Eukaryotic aspartyl protease family protein
AT1G12230.1	Aldolase superfamily protein
AT5G58250.1	unknown protein
AT4G01130.1	GDSL-like Lipase/Acylhydrolase superfamily protein
AT1G09210.1	calreticulin 1b
AT1G13280.1	allene oxide cyclase 4
AT2G02560.1	cullin-associated and neddylation dissociated
AT2G34480.1	Ribosomal protein L18ae/LX family protein
AT1G69830.1	alpha-amylase-like 3
AT3G12145.1	Leucine-rich repeat (LRR) family protein
AT3G12390.1	Nascent polypeptide-associated complex (NAC)
AT3G14210.1	epithiospecifier modifier 1
AT5G59320.1	lipid transfer protein 3

5.2.2. Infected IWF sample triggers immune responses

Jefferson et al. (1987) introduced a method in which GUS gene is fused to the gene of interest and when the gene of interest is activated, GUS gene is activated as well; hence the GUS enzyme is produced. Then, this production is detected by observing the colour change to a clear blue when a glucoronide substrate (e.g. X-gluc) is processed. Here, we took advantage of this reporter system in order to see whether the IWF samples were able to trigger a reaction in transformed plants carrying PR-GUS or RLK-GUS fusions in Col-0 background. Besides the infected and healthy IWFs, boiled and proteinase-K treated versions thereof were used as well, in order to make sure any reaction to be observed will be resulting from a structurally stable protein source, and no other agents. The treated samples were compared to the originals on an SDS-PAGE gel (Fig. 5.4), loaded at adjusted concentrations. On the gel, it was observed that there was a slight decrease in the amount of the samples for boiled ones and proteinase-K was successful in degrading all the proteins in the IWF samples. The gel also showed that after degradation, the proteinase-K was staying intact.
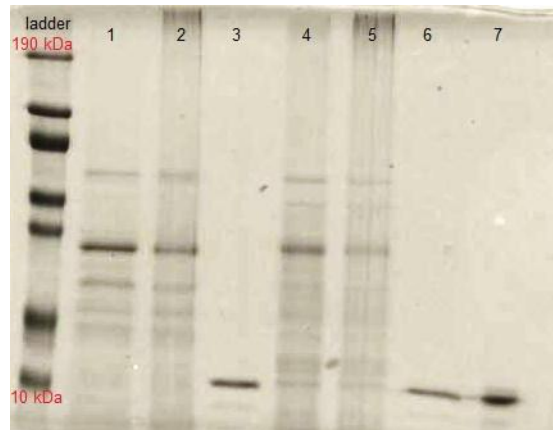
Figure 5.4. SDS-PAGE of untreated and treated IWF samples. (Lane 1: infected IWF, 2: boiled infected IWF, 3: proteinase K-treated infected IWF, 4: healthy IWF, 5: boiled healthy IWF, 6: proteinase-K treated healthy IWF, 7: proteinase K). There was a decrease in the amount of the protein samples in boiled ones. Proteinase-K was successful in degrading all the proteins in the IWF samples. The proteinase-K itself was staying intact after serving its purpose.

The original and treated samples were pressure-injected to the leaves of 4-5 week-old Col-*PR-GUS* and Col-*RLK-GUS* plants, along with controls. *E. amylovora* (O.D.$_{600}$= 0.25) was used as positive control since it is known to trigger defence responses in *A. thaliana* (Degrave et al., 2008). Sterile dH$_2$O was used as negative control as all the samples were prepared in sdH$_2$O. In addition, proteinase- K was also injected as control to see whether it had an effect on its own, since it was not eliminated from the samples after treatment. After the injections the plants were incubated (Col-*RLK-GUS* for 24 h, Col-*PR-GUS* for 48 h) to allow reactions to develop. The incubated leaves were then subjected to GUS staining solution and were destained in methanol. The results showed that the infected IWF caused a significant reaction whereas; boiled-infected, healthy, boiled-healthy were showing almost negligible results as negative control on both transgenic lines (Fig 5.5a-b). However, proteinase-K

itself also caused a strong reaction; therefore it was not possible to evaluate the reaction caused by proteinase-K-treated samples.
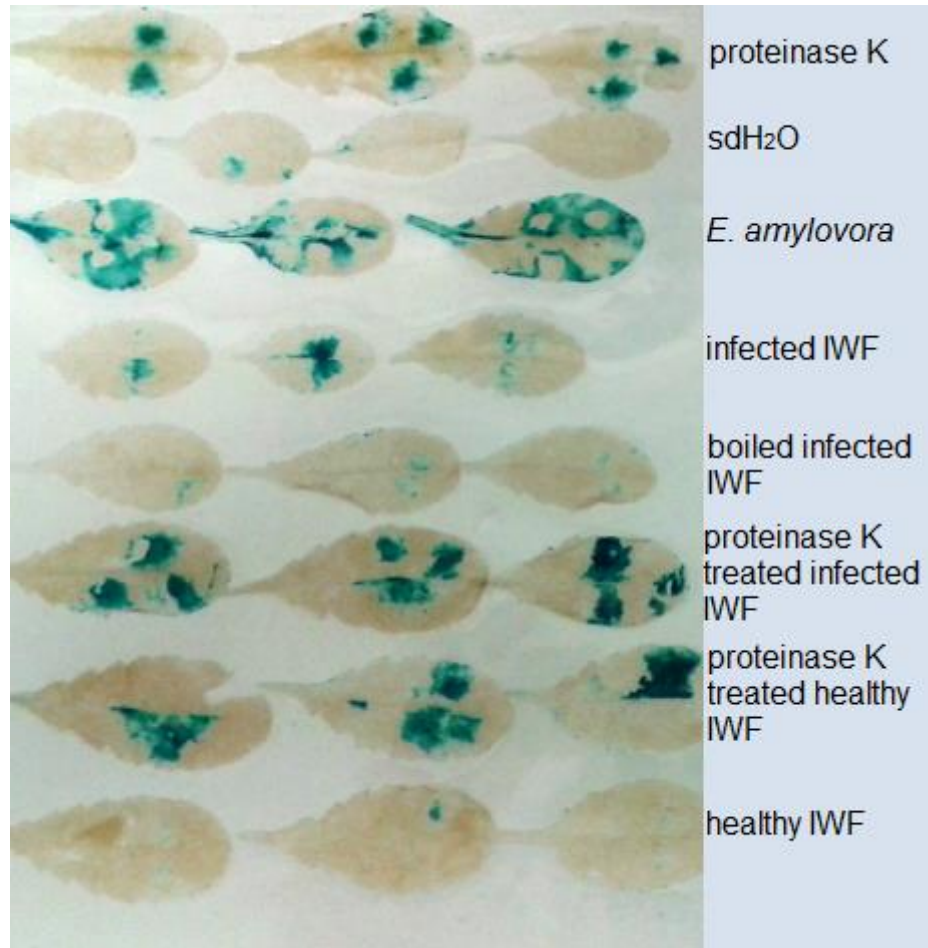
Figure 5.5a.GUS assay on Col-*RLK-GUS* plants treated with various IWF samples. Reactions observed as blue colouring. The infected IWF caused a significant reaction. Boiled-infected, healthy, boiled-healthy were showing almost negligible results as negative control. Proteinase-K caused a strong reaction; thus it was not possible to evaluate the reaction caused by proteinase-K-treated samples.
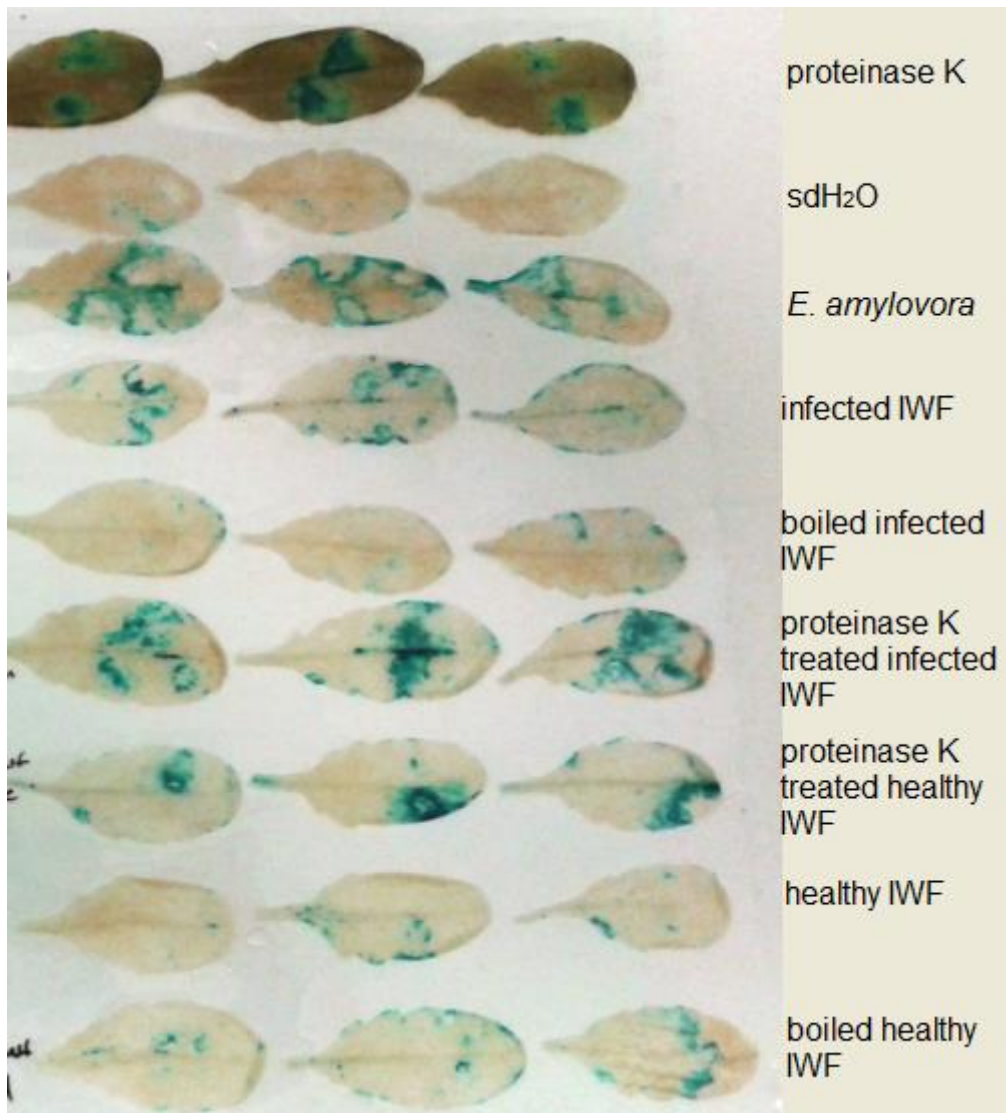
Figure 5.5b.GUS assay on Col-*PR-GUS* plants treated with various IWF samples. Reactions observed as blue colouring. The infected IWF caused a significant reaction. Boiled-infected, healthy, boiled-healthy were showing almost negligible results as negative control. Proteinase-K caused a strong reaction; thus it was not possible to evaluate the reaction caused by proteinase-K-treated samples.

5.2.3. IWF samples trigger ROS production

After observing that IWF samples can induce PR gene expression, determined by GUS assay, ROS production triggered by IWF was investigated in detail, an indicative of activated defence responses. For measuring the ROS production, Felix et al. (1999) proposed a method where plant leaves are treated with samples and then presented with a substrate, 5-aminosalicylic acid (5-AS) to measure the activity of extracellular peroxidases that are involved in ROS production. If peroxidases are produced, they will convert the colourless substrate to an orange-brown product, and this change is measured by a spectrophotometer (at 450 nm). Following this method, healthy, infected, boiled healthy and boiled infected IWF samples were screened for ability to trigger ROS on Col-0, Col-*rpp4* and Col-*bak1-5/bkk1-1* and *N. benthamiana* plants. These plants were chosen, since we wanted to compare the reactions given by the resistant wild type and non-host plants, as well as the plants impaired in a known receptor for the effector ATR4 and a co-receptor BAK1; in order to see whether these receptors were involved in recognition of the molecules in the IWF samples. In the assays, pure horseradish peroxidase (0.2M prepared in sodium phosphate buffer, added just before quantification) and flg22 (100 nM final concentration) were used as positive control, and the negative controls was sdH$_2$O. Measurements were taken immediately after substrate addition, and after 2 min. According to the numbers obtained from treatment of Col-0, the infected IWF sample was causing a significantly high reaction immediately, even greater than the positive controls; whereas the healthy and boiled samples were in accordance with the negative control. After 2 min of incubation with the

substrate, reaction caused by the healthy IWF and boiled samples were not showing a noticeable change, however the healthy IWF sample was showing a considerable reaction, almost four times greater than the initial reaction (Fig. 5.6a).
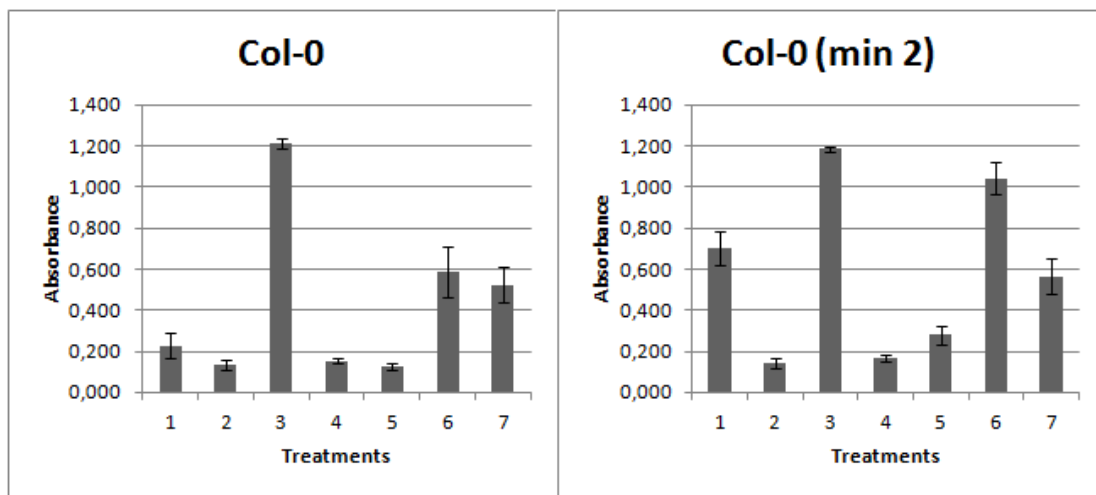
Figure 5.6a. Quantitative ROS assays on Col-0 plants treated with various IWF and control samples. Measurements taken immediately and 2 min after treatments. (Treatments: 1: healthy IWF, 2: boiled healthy IWF, 3: infected IWF, 4: boiled infected IWF, 5: water 6: flg22 7: horseradish peroxidase). Infected IWF sample caused a significant reaction immediately, greater than the positive controls. Healthy and boiled samples' results were in accordance with the negative control. After 2 min of incubation with the substrate, reaction caused by the healthy IWF and boiled samples did not show a noticeable change, however the healthy IWF sample was showing a considerable reaction, almost four times greater than the immediate reaction.

In case of Col-*rpp4* and Col-*bak1-5/bkk1-1* plants, the results were in consistence with the wild type plants (Fig. 5.6b-c).
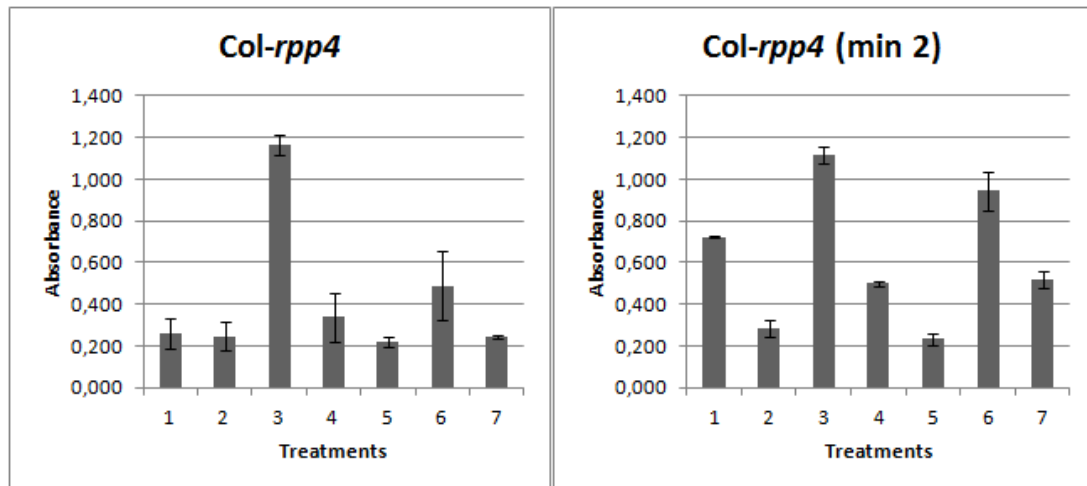
Figure 5.6b. Quantitative ROS assays on Col-*rpp4* plants treated with various IWF and control samples. Measurements taken immediately and 2 min after treatments. (Treatments: 1: healthy IWF, 2: boiled healthy IWF, 3: infected IWF, 4: boiled infected IWF, 5: water 6: flg22 7: horseradish peroxidase). The ROS results for Col-*rpp4* were similar to wild type Col-0.
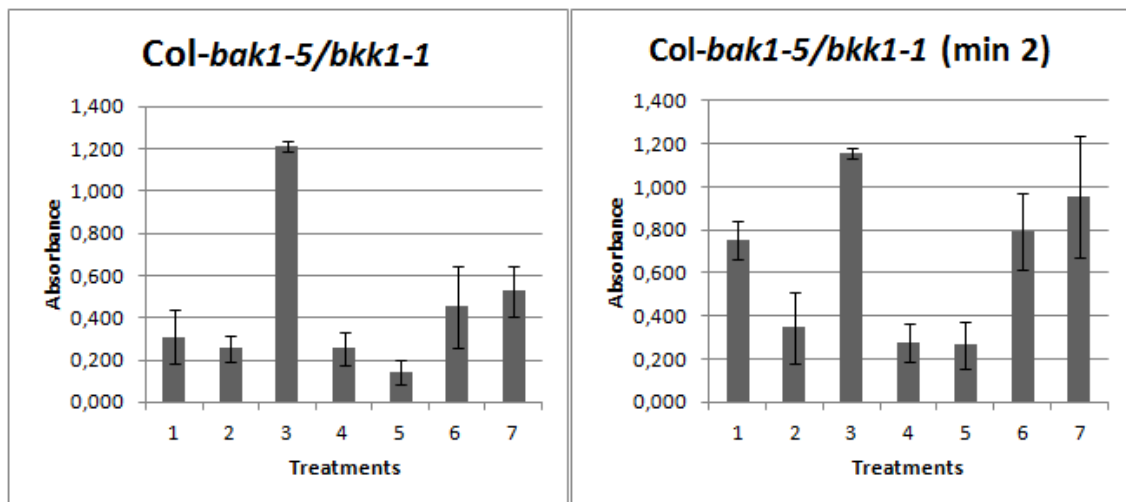
Figure 5.6c. Quantitative ROS assays on Col-*bak1-5/bkk1-1* plants treated with various IWF and control samples. Measurements taken immediately and 2 min after treatments. (Treatments: 1: healthy IWF, 2: boiled healthy IWF, 3: infected IWF, 4: boiled infected IWF, 5: water 6: flg22 7: horseradish peroxidase). The ROS results for Col-*bak1-5/bkk1-1* were similar to wild type Col-0.

For *N. benthamiana* plants, healthy IWF sample was also causing a noteworthy reaction at immediate measurements, and after 2 min its reaction was almost as high as the infected IWF sample. The reactions caused by the boiled IWF samples were in agreement with the reactions they caused in the other plant materials (Fig. 5.6d)
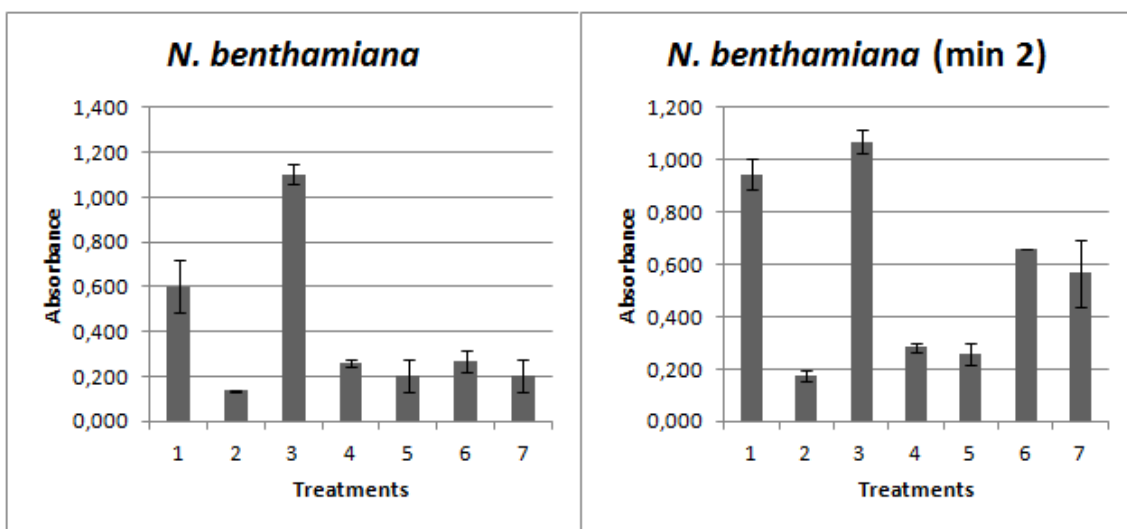
Figure 5.6d. Quantitative ROS assays on *N. benthamiana* plants treated with various IWF and control samples. Measurements taken immediately and 2 min after treatments. (Treatments: 1: healthy IWF, 2: boiled healthy IWF, 3: infected IWF, 4: boiled infected IWF, 5: water 6: flg22 7: horseradish peroxidase). Healthy IWF sample caused a significant reaction at immediate measurements. At 2 min measurements its reaction was almost as high as the infected IWF sample. The reactions caused by the boiled IWF samples were similar to the reactions they caused in the other plant materials.

5.2.4. Ion exchange chromatography is a reliable method for fractionating IWF samples

With the aim of narrowing down the focus to identify the active portion of the compound IWF samples, ion exchange chromatography method was utilized to separate the contents of the samples. For obtaining fractions of the both infected and healthy IWF samples, we worked with manually conducted strong anion exchange columns. Initially, the samples were passed through a desalting column and their composition was adjusted to the start buffer. Desalted samples were re-filter-sterilized and then fractions were collected with sterile increasing concentrations of NaCl solutions (0.1 M, 0.15 M, 0.2 M, 0.25 M, 0.3 M, 0.35 M, 0.4 M, 0.45 M, 0.5 M and 0.55 M). Eleven fractions were collected including the flow through (FT). The fractions were quantified using Bradford assay and visualized via 12.5% SDS-PAGE (Fig.5.7). The SDS-PAGE results showed successful separations of both infected and healthy IWF samples. All fractions were loaded at same concentrations and it was observed that towards higher salt concentrations (0.45 M, 0.5 M and 0.55 M), the number of protein bands was decreasing. On the other hand, the FT fractions also contained a considerable number of proteins showing that a portion of the original sample was not binding to the column.
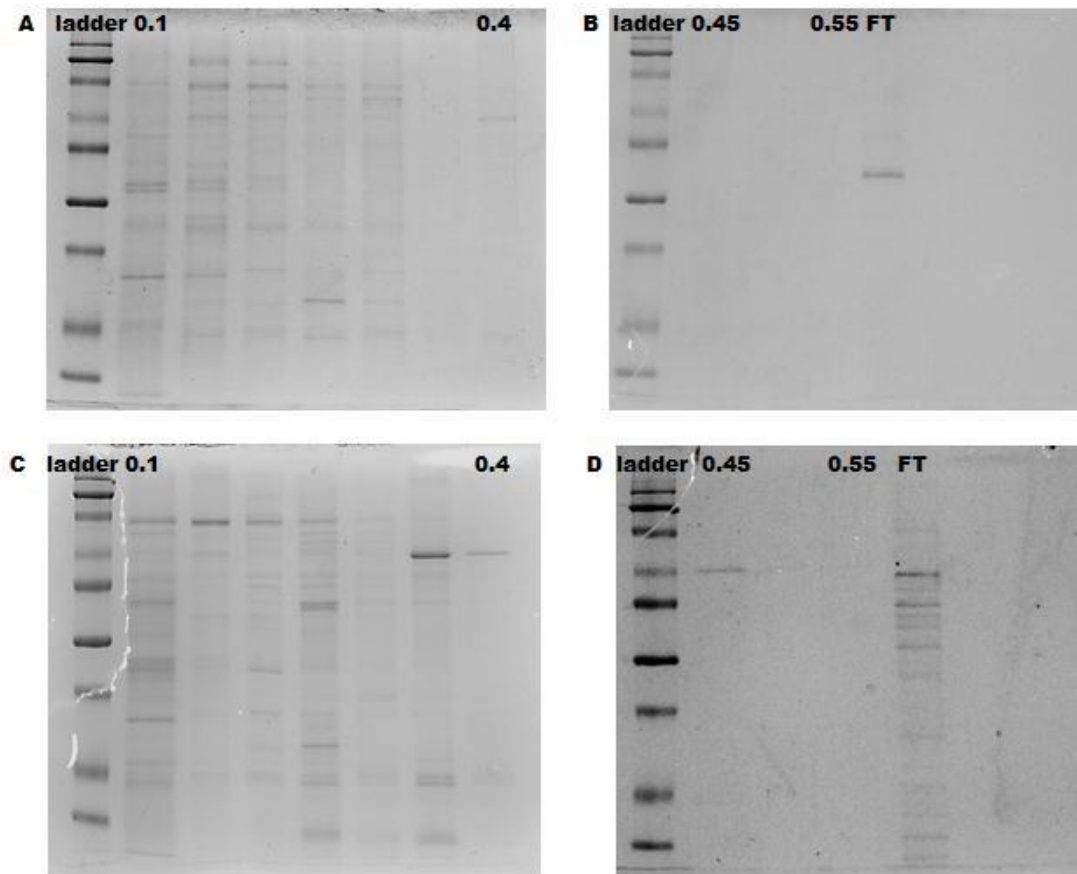
Figure 5.7. SDS-PAGE of the fractionated protein samples. Ladder 10 kDa-190kDa. (A and B: infected IWF fractions, C and D: healthy IWF fractions, collected eluting with 0.1 M, 0.15 M, 0.2 M, 0.25 M, 0.3 M, 0.35 M, 0.4 M, 0.45 M, 0.5 M and 0.55 M NaCl). All fractions were loaded at same concentrations. Both infected and healthy IWF samples were separated successfully. Towards higher salt concentrations (0.45 M, 0.5 M and 0.55 M), the number of protein bands was decreasing. The Flow through (FT) fractions also contained a considerable number of proteins showing that a portion of the original sample was not binding to the column.

5.2.5. ROS assays with fractions suggest the activity lies within the flow-through

After successful separation of the IWF samples, the fractions were subjected to ROS assays to find out which ones were triggering reactions observed, hence containing the molecules causing defence activation. All fractions including the FT of both healthy and infected IWF were used as treatments on Col-0 plants. The resuspension buffer and sdH_2O were used as negative controls; horseradish peroxidase and flg22 were used as positive controls. The results showed that only FT fraction of infected IWF sample was causing a strong reaction, and all the reactions caused by the rest of the samples were almost negligible (Fig. 5.8). The assays were repeated several times and the results were reproducible and consistent. Therefore this active fraction was used in further experiments.
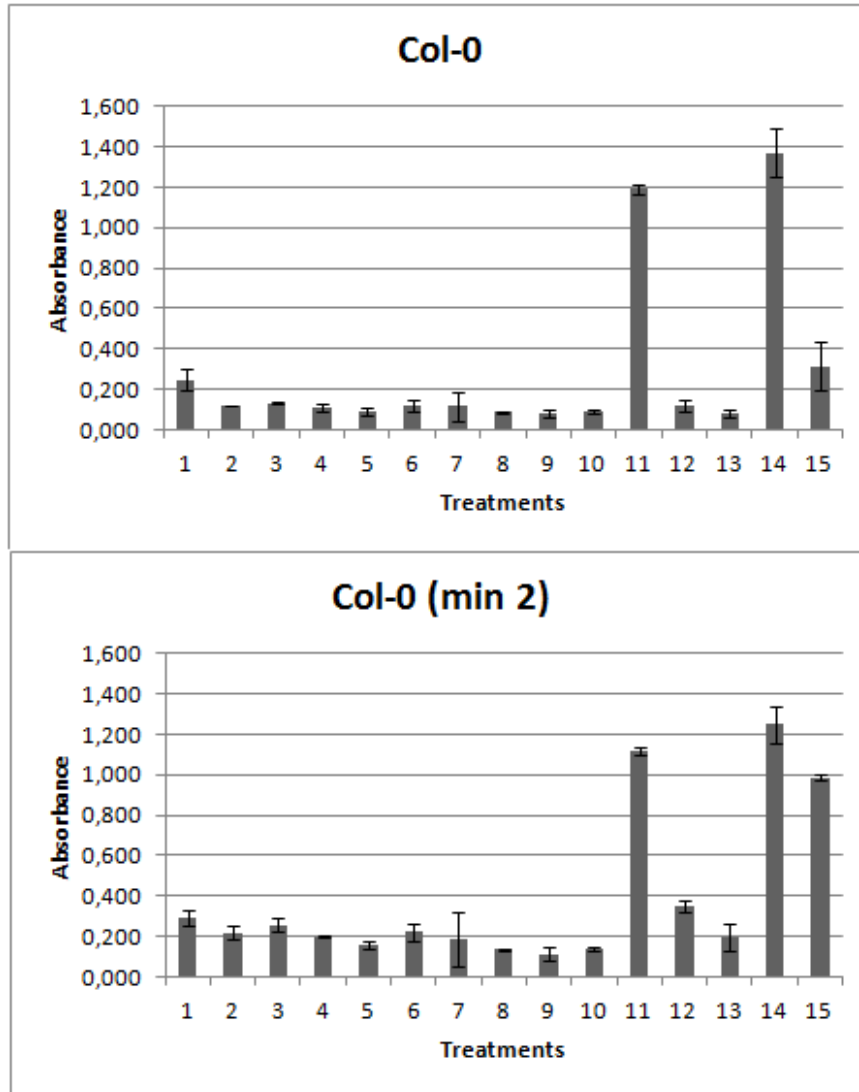
Figure 5.8. ROS assay on Col-0 with fractions of infected IWF sample. (1-10: Fractions collected with increasing NaCl concentrations 0.1 M - 0.55 M, 11: FT-flow through, 12: water, 13: resuspension buffer, 14: horseradish peroxidase, 15: flg22). FT fraction of infected IWF sample caused a strong reaction, and all the reactions caused remaining samples were very low.

5.2.6. Sporulation and ROS assays with T-DNA insertional mutant lines of selected LRR proteins, suggest there might be other players involved in perception of the pathogen

MALDI-TOF analysis of the IWF samples revealed proteins annotated as LRR protein family, indicating that these putative receptors may be involved in the *Hpa-Arabidopsis* interactions. Therefore, the LRR proteins were checked for existence of a signal peptide using SignalP, thus they would be secreted out of the plant cell and act as apoplastic or cell-surface receptors. After this screening, 3 LRR family proteins; 1 up-regulated AT1G33610.1 and 2 down-regulated AT3G20820.1 and AT1G49750.1 were selected. T-DNA the mutant lines were obtained from Nottingham Arabidopsis Stock Centre (NASC). Two different mutant lines for each gene (Table 2.4) and given left and right border primers for each mutant were ordered, as well as the common 3rd primer to prove the insertion and select homozygous mutant lines (Table 2.5).

All seeds were grown for 4-5 weeks and DNA was isolated and amplified using Extract-N-Amp™ Plant PCR Kit (Sigma-Aldrich). In order to select the homozygous lines, 3 primers were used in the PCRs: gene-specific primers (Left border primer, LP and right border primer, RP) and a forward primer (Border primer, BP) designed to match a sequence within the T-DNA. After PCRs, it was expected to observe a single band for homozygous mutants due to amplification of the segment between the border primer and right border; a single band for wild type due to amplification of the fragment between left

border (LB) and right border (RB) and both bands for the heterozygous mutant lines (Fig 5.9).
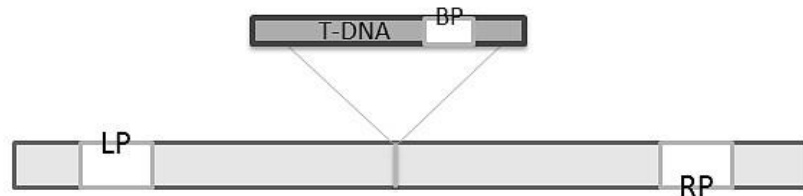
Figure 5.9. The corresponding positions of the left and right primers on the genome and the forward border primer in the T-DNA insert. The homozygous lines were selected using 3 primers in the PCRs: gene-specific primers (Left border primer, LP and right border primer, RP) and a forward primer (Border primer, BP) designed to match a sequence within the T-DNA. The expected outcome of the PCRs were to observe a single band for homozygous mutants due to amplification of the segment between the border primer and right border. For heterozygous mutant lines the expectation was to see both bands, and a single band for wild type due to amplification of the fragment between left border (LB) and right border (RB).

According to the agarose gel results, homozygous lines were selected. However, it was not possible to obtain homozygous lines for N547337 and N579946. All the seeds turned out to be heterozygous; thereby those plants were grown to obtain seeds for further segregation in T_2 lines. Unfortunately, these lines were lost to secondary infections, and were not possible to carry forward. Therefore, sporulation assays were performed with the remaining lines.

In order to see whether the mutations were affecting the resistance or susceptibility against different *Hpa* isolates; seeds (3 pots/each mutant) were sown on 40-pot modules, excluding the edge pots to prevent edge effect, each tray including Ws-*eds1* and Col-0. The seeds were grown for 7 days and were inoculated separately with *Hpa*-Emoy2, *Hpa*-Cala2 and *Hpa*-Noks1 isolates at 5×10^4 spores/ml concentrations. A week later, infections were observed. For

Hpa-Emoy2 and *Hpa*-Cala2 isolates, the number of sporangiophores was counted from 15 seedlings from each pot. On the other hand, for *Hpa*-Noks1, the infection levels were measured by suspending 15 spore bearing seedlings in 250 µl chilled sdH$_2$O and counting the number of spores using a haemocytometer.

The inoculations of *Hpa*-Emoy2 and *Hpa*-Cala2 were controlled by the counts of Ws-*eds1*, which turned out to be as expected, highly-sporulated, and Col-0, which showed very low or no sporulation at all. All the mutants inoculated with these isolates also showed resistance marks and no sporulation, suggesting that mutating these LRR genes did not result in enhanced susceptibility when compared to the controls. On the other hand, Col-0 and Ws-*eds1* are both known to be susceptible to *Hpa*-Noks1 isolate, and spore counts of the mutant lines revealed similar results to controls again supporting the outcome of the isolates *Hpa*-Emoy2 and *Hpa*-Cala2 (Fig. 5.10). Overall, it can be said that absence or reduced expression of these selected receptor proteins do not seem to be affecting the resistance or susceptibility of the plant on their own.

In addition to sporulation assays, a ROS assay was also performed with these mutant lines using flow-through fraction of infected IWF as treatments. Mutants lines (different homozygous plants) were compared to Col-0 and the results did not show any significant difference, suggesting that the mutations are not directly affecting the perception of the proteins within the FT (Fig. 5.11)
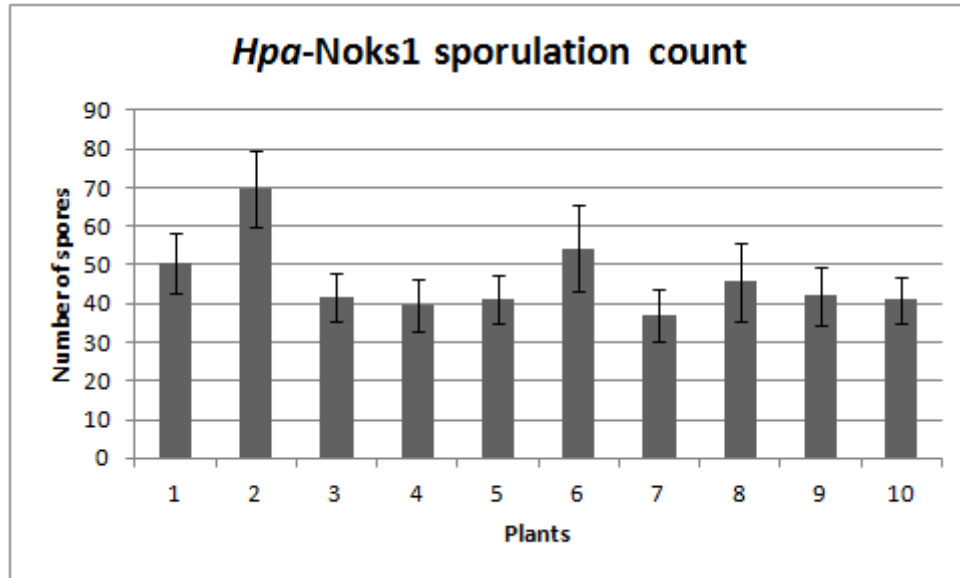
Figure 5.10. Sporulation count of *Hpa*-Noks1 on T-DNA mutant lines, Col-0 and Ws-*eds1*. (1: Col-0, 2: Ws-*eds1*, 3 and 4: SALK_151177C, 5 and 6: SALK_055913C, 7 and 8: SALK_119747C, 9 and 10: SALK_047896). Col-0 and Ws-*eds1* are susceptible to *Hpa*-Noks1 isolate. Spore counts of the mutant lines showed similar results to those controls. The results were in consistence with *Hpa*-Emoy2 and *Hpa*-Cala2.
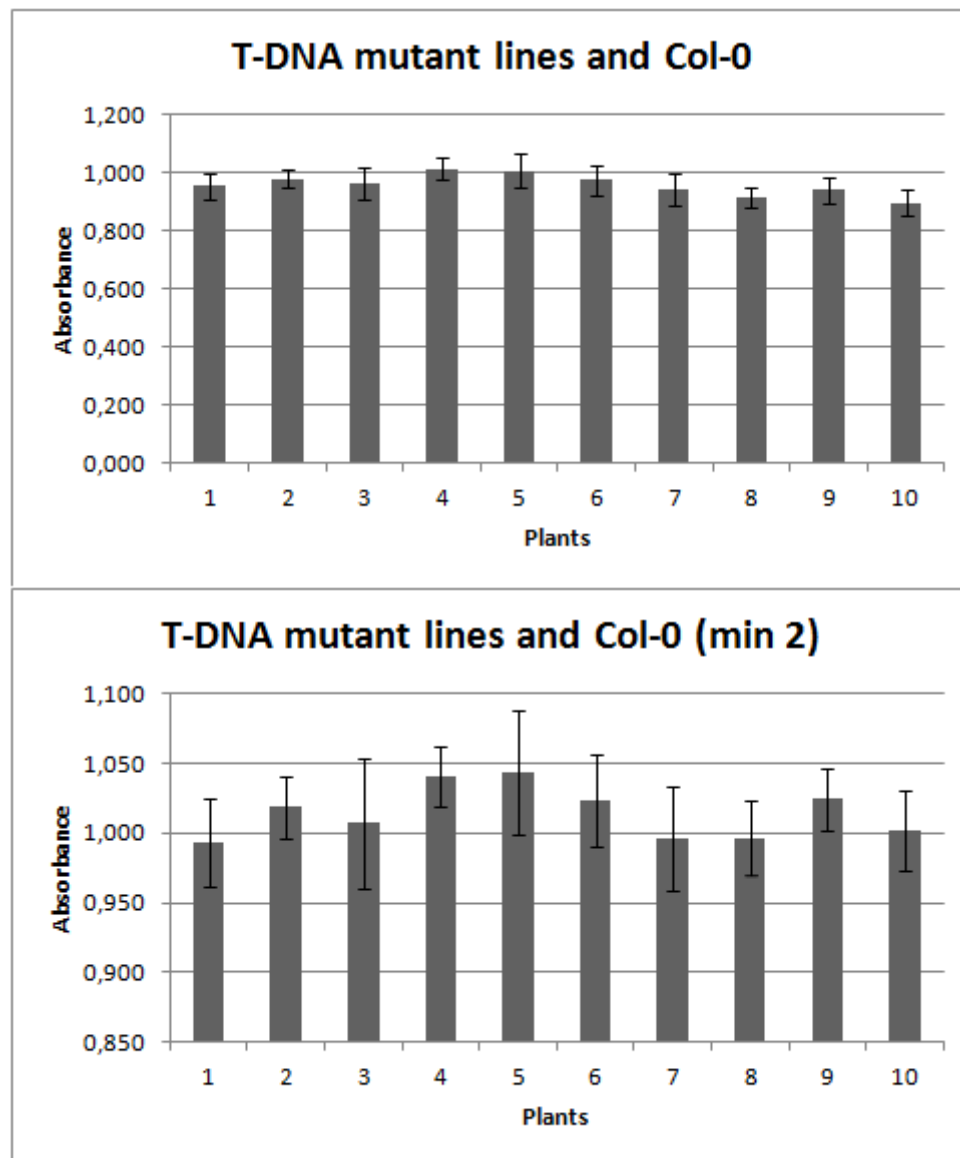
Figure 5.11. ROS assay with homozygous T-DNA mutant lines and Col-0 using infected FT as treatment. (1 and 8: SALK_119747C; 2 and 3: SALK_151177C, 4 and 9: SALK_047896, 5-7: SALK_055913C, 10: Col-0). Homozygous mutants were compared to Col-0 and the results did not show any significant difference.

5.3. Summary

In this chapter, a proteomic approach was taken with the intention of identifying novel plastic effectors and their corresponding receptors. With this purpose Intercellular washing fluid, presumably containing apoplastic molecules from both plant and the pathogen, was collected from *Hpa*-Emoy2 infected Ws-*eds*1 plants as well as healthy plants as neative control. Both the infected and healthy IWF samples were used in subsequent assays. With the GUS reporter system, was shown that the infected IWF sample was carrying such molecules that are playing a role in defence activation in transgenic Col-*PR-GUS* and Col-*RLK-GUS* plants. Additionally, a ROS assay, in which the production of extracellular peroxidases is measured, was also performed with wild-type Col-0 and mutants impaired in pathogen recognition as well as non-host plant materials. Both GUS reporter and ROS assays demonstrated consistent results suggesting that the infected IWF had defence triggering molecules when compared to healthy samples. Additionally, the infected IWF sample was analysed by MALDI-TOF mass spectrometry, which detected a significant number of protein from both *Hpa* and *A. thaliana*. The Hpa-sourced proteins also included cysteine-rich proteins which carry the potential of being apoplastic effectors. Therefore 2 of those Cys-rich proteins were selected to use in *in vitro* espression assays along with the bioinformatically selected candidates. On the other hand, *A.thaliana*-sourced protein contained LRR proteins which are known effector receptors. Out of those proteins signal-peptide containing 1 up-regulated and 2 down-regulated proteins were selected. In order to investigate the roles of those proteins in pathogen recognition, T-DNA insertional

homozygous mutant lines were obtained which were impaired in the genes encoding these proteins. The mutant lines were then subjected to interaction phenotype assays to observe changes in resistance or susceptibility by challenging them with different *Hpa* isolates and comparing to wild-type and known susceptible accessions. The results showed no significant differences for the mutant lines, suggesting that there might be other players involved in their activities.

Moreover, the IWF was fractionated via ion exchange chromatography technique. The fractionation and different bands were clearly visible on SDS-PAGE. Therefore the fractions were used in ROS assays to narrow down the focus to the most-reaction-causing fraction to take the research further. The fractions however, did not result in noteworthy reactions on their own, except for the flow-through. It was concluded that in order to be able to carry the fractionation further and test them in functionality assays, a greater amount of starting material was necessary.

6. DISCUSSIONS

A. thaliana-Hpa pathosystem has been an extensively used model system for investigating the molecular interactions and understanding the underlying mechanisms of the downy mildew disease (Slusarenko and Schlaich, 2003). Despite the increasing attention, there is still a huge gap especially in the repertoire of apoplastic effectors. Therefore, in this study, the objective was to identify novel apoplastic effectors of *Hpa* and investigate their roles in immune responses of *A. thaliana*. With this purpose we selected five candidate genes taking advantage of the constructed EST library of *Hpa*-Emoy2. The candidates were scrutinized to reveal their expression patterns, capabilities of triggering defence response *in planta* and variations across different isolates. *In vitro* expression of the candidate genes was also tried. We observed that the candidate genes were not triggering an HR and were mostly conserved among isolates. In addition, the intercellular washing fluid was obtained from both healthy and infected plant materials in order to observe whether the molecules in the apoplastic matrix were able to activate plant's immune system via various assays. GUS reporter assays suggested that the infected IWF contains molecules that are able to trigger defence responses in transgenic Col-*PR-GUS* and Col-*RLK-GUS* plants, when compared to the healthy IWF. A ROS assay, in which the production of extracellular peroxidases

was measured, revealed similar results. The complex IWF was then fractionated, and in ROS assays, the flow-through fraction was found to be the active fraction, containing molecules responsible for the reactions. This active fraction was then tested on mutant plants that were impaired in known effector receptors, and we observed that RPP4 and BAK1 co-receptor are not involved in the recognition of the molecules in the flow-through fraction. Furthermore, IWF samples were analysed by MALDI-TOF and the results showed that the mixture contains *Hpa*-specific proteins such as cysteine-rich and elicitin-like proteins with a potential of being apoplastic effectors. Two cysteine-rich proteins from the *Hpa*-specific ones were selected and were used in *in vitro* expression trials. In addition, according to the MALDI-TOF results, there were apoplastic LRR family proteins, which might have a role in perception of effectors. Therefore, T-DNA insertional mutant lines for three selected LRR proteins from the MALDI-TOF results were also tested for interaction phenotypes for enhanced resistance/susceptibility. A significant change in those attributes was not observed. Finally, the flow-through fraction was tested on LRR mutants in a ROS assay, and we observed no significant change in the magnitude of the reactions. In this chapter, these finding will be discussed in detail.

6.1. A bioinformatic approach towards identifying novel apoplastic effectors

For many researchers, sequencing cDNA has become an attractive approach for building databases for various eukaryotic plant pathogens (Kamoun et al., 1999b; Qutob et al., 2000; Skinner et al., 2001; Dautova et al.,

2001). Using the data generated by these genomic studies and advancing bioinformatic technologies, sequence-based screening has become a fast, efficient and inexpensive method to annotate and discover the functions of the genes. The model system for downy mildew '*A. thaliana–Hpa*' used in this study also had the advantage of data-availability (accession Col-0 and isolate *Hpa*-Emoy2, respectively) (AGI, 2000; Baxter et al., 2010); and therefore determining candidate genes and investigating their roles in host-pathogen interactions by screening EST databases was opted.

6.1.1. Choosing secreted apoplastic effector candidates from EST databases using bioinformatic tools

In the recent years oomycete effectors have become an attractive research topic. Significant amount of data has been collected on the effector trafficking, evolution and function (Birch et al., 2006). Especially, the discovery of the translocation motifs for cytoplasmic effectors was a breakthrough. This enabled developing data-mining strategies in order to identify novel effectors (Win et al., 2008; Haas et al., 2009). On the other hand, proteins such as enzyme inhibitors, small cysteine rich proteins and Nep1-like proteins were defined as effectors that function in the apoplast (Kamoun, 2006). Since the apoplastic effectors lack a conserved motif, and they have the possibility of being recognized by apoplastic receptors or cell surface receptors, as well as inducing intracellular pathways made their identification a challenging topic. There is a limited number of defined apoplastic effectors, therefore they require a special attention in order to compile a larger repertoire. Thus, with the

purpose of identifying novel apoplastic effectors, we assessed short sub-sequences from cDNA (ESTs) database of *Hpa*-Emoy2 using the existing knowledge from the previous studies in order to define the criteria and select putative secreted proteins. More than 3000 ESTs were screened using several different online bioinformatics tools combined with the Geneious software. Five candidates without a function assigned were selected, all of which contained a signal peptide, lacked the signatures of cytoplasmic effectors such as RxLR, dEER, LxLFLAK or CHxC. It was also important to choose relatively shorter sequences for easier manipulation. While determining candidate genes, all the tools were found to be handy and easy to use; however the whole process was time consuming as it required using several different platforms to check and confirm different criteria. Developing a new, user-friendly and inexpensive software, dedicated to *Hpa* isolates that would offer a combination and extension of the features would be really helpful for future studies.

6.2. Molecular characterization and evaluation of the candidate genes

6.2.1. Amplification of the candidate genes

There are several studies, in which, apoplastic effectors were discovered via bioinformatic prediction. For instance, the Serine protease-inhibitors EPI1 and EPI10 were first selected through data-mining and were then used in functionality assays (Tian et al., 2004) Similarly, investigation on EST libraries led to discovery of two cysteine protease inhibitors, EPIC1 and EPIC2 (Tian et

al., 2007). In addition, small cysteine-rich proteins, PcF-like SCR74 and SCR91, were also identified after bioinformatic prediction based on previous findings on PcF proteins (Orsomando et al., 2003; Orsomando et al., 2011). In all, the candidates were determined with a sequence-based analysis, and were validated for their expressions before moving on to *in vitro* and *in planta* functionality assays. Therefore, here we followed the same route and first observed expression of our apoplastic effector candidates. DNA was isolated from infected Ws-*eds1* seedlings and the candidate genes were amplified via PCR, in order to show that the candidates actually existed in the *Hpa*-Emoy2 isolate, and to use for cloning purposes. After observing amplification of the candidate genes, they were cloned for subsequent *in vitro* and *in planta* expression and functionality assays in order to assess whether they were able to trigger defence responses.

6.2.2. Expression pattern of the candidate genes

In order to check at which point and what level the candidate genes were expressed by the pathogen after inoculation, samples were collected from infected material with 24 h intervals. Healthy samples were also processed as negative controls. To evaluate the expression of the candidate genes, RT-PCR (touchdown) was performed as it allows qualitative observation of gene expression via generating cDNA transcripts from RNA and amplifying them. To do that, initially RNA was isolated from plant samples using Qiagen's RNeasy Plant Mini Kit. The kit was easy to use and results were achieved fast; however, DNA contamination was considerably high in the products. Therefore, it was

crucial to eliminate DNA from these samples with an optimized protocol. Following the elimination of DNA, the samples were brought to same concentration prior to RT-PCR. In the results of RT-PCR, it was possible to observe many non-specific bands on the gels, which could have resulted from operating conditions and non-specific binding of the primers to the cDNAs. This occurrence of non-specific bands was observed in the repetitions of the experiments; therefore we carried on focusing on the bands-of-interest and evaluating the results accordingly. The intensity and the timing of the expression showed variation between the candidate genes. It could be suggested that some encoded proteins by the candidate genes can be put forward by the pathogen right away or as a second layer of attack to suppress the PTI later during the infection. On the other hand, the proteins can be produced in an increasing manner to establish pathogenicity by evading the defence responses or they can be recognized during the infection which may result in a decrease in the expression level. In addition to the candidate genes, *actin* gene was also tested for the expression level after the infection as a reference. Actin has been a useful tool as a control, especially since it was shown that during a biotrophic infection, due to rearrangements of the host cytoskeleton, *actin* is expressed in an increasing manner as a response to pathogen invasion (Jin et al., 1999; Kamoun et al., 1998). In case of *Hpa-Emoy2* infection, the actin gene was starting to be expressed on the 3rd dai and the expression level was following an increasing trend at first, then staying constant. In order to have a better understanding of the expression levels a

quantitative PCR can be performed and the level of expressions can be observed for an extended period of time.

6.2.3. Looking for signs of HR by *Agrobacterium-mediated* transient expression of the candidate genes

Agrobacterium-mediated transient expression is a widely-used system, where the *A. tumefaciens* culture carrying the gene of interest is infiltrated into plant leaves, especially tobacco, tomatoes and potatoes (Zupan and Zambryski, 1995; Li et al., 2009). *A. tumefaciens* is naturally able to transfer genes when it is used to infect a plant. Simply put, the bacteria transfer T-DNA into the plants cells, a segment that is a part of the bacterial tumour-inducing plasmid; and when the bacteria is engineered to carry the gene of interest, it is delivered alongside. The gene of interest is then expressed for only a short amount of time (usually several days) since it is not integrated into the plant genome, thus is not replicated. Previously, Bos et al. (2006) showed that the RxLR effector AVR3a was able to trigger HR response and suppress INF1-induced cell death in *N. benthamiana* plants via *Agrobacterium*-mediated transformation. In addition, Liu et al. (2011) used their bioinformatically-selected, expression – validated putative cytoplasmic effectors and observed their role in HR induction via the same method. Here, we also used this type of transformation in order to see whether our candidate genes were causing any HR on the non-host plant: *N.* benthamiana and *N. tabacum*. *E. amylovora* was used as positive control as it is known to trigger HR (Kim and Beer, 1998). Since HR is a rapid response and transient transformation is effective for a limited time only, any reaction that

would appear were expected to be observed within 2-3 days after injection. The expected HR reaction was only observed from the positive control, whereas the candidates were not showing signs of HR even after 5-6 days. Therefore, the results suggested that the candidates were not likely to cause an HR, since chlorosis was not observed soon after infiltration. This may indicate that a) the candidate genes may not be expressed within tobacco. However, our RT-PCR results indicated that they do express (data not shown), b) tobacco may not be the right host for these genes to trigger defence; c) these apoplastic genes may not trigger HR but other responses such as callose deposition or ion flux.

6.2.4. Floral dipping as a straight-forward method for stable transformation of *Arabidopsis*

Creating transgenic lines is a key technique for introducing new traits to a plant and studying gene functions, using recombinant DNA technologies. There are several methods to achieve plant transformation, such as gene bombardment, electroporation and microinjection. One other method proposed by Clough and Bent (1998), uses *Agrobacterium*-mediated transformation based on simple dipping of developing floral tips (female gametes) of the plant into a solution containing the *Agrobacterium* cells carrying the gene of interest. The cell suspension is added sucrose and a surfactant (Silwet L-77) to increase the transformation efficiency. For a better yield, it is also important to perform the dipping when there is maximum amount of closed buds. Based on this procedure, it was decided to transform resistant Col-0 plants with the candidate genes, and check whether the insertions of the genes of interest into host

genome would cause any difference in resistance/susceptibility attributes. It was possible to transform the plants finely; however the seeds were catching fungal infections, a common problem encountered with this method (Das and Joshi, 2011). Therefore, despite repeated attempts it was not possible to carry the functional assays to further steps.

6.2.5. *In vitro* gene expression for producing the candidate putative effector proteins

It was essential to express the candidate genes in order to use the proteins they encode in activity assays for understanding whether they carried effector features. For example, to study the recognition of the elicitor INF1 protein, Kamoun et al. (1998) inserted the encoding gene into FLAG-tagged expression vector. On the other hand, Sohn et al. (2007) showed that cytoplasmic effectors ATR1 and ATR13, were promoting susceptibility in *A. thaliana* plants, and their first step was to clone the genes of interest via restriction enzyme digestion and ligation method to an expression vector. Additionally, another RxLR effector Avr1b was fused GFP and inserted into expression vector, prior to translocation assays (Dou et al., 2008). Moreover, Bos et al. (2010) used the Gateway® cloning method to generate constructs to express the Avr3a to carry out the assays for observing its role in manipulation of plant immunity. With the same intentions, several different methodologies were followed in order to express our candidate genes *in vitro*. First, pET28a expression system, which adds His-tags to both N- and C-terminals for further purifications, was chosen. The candidates were inserted in the vector via

restriction enzyme digestion and ligation method. For the ligation step, it was important to check whether the ligase enzyme buffer contained polyethylene glycol (PEG). PEG is important because it is a hydrophobic molecule that piles the aqueous solution of the ligation mix, hence increasing the DNA and ligase concentration, and boosting reaction efficiency. However, PEG could also decrease the transformation efficiency if the enzyme is heat inactivated or ligation time is extended. Paying attention to these points, the candidate genes were tried to be inserted in expression vector. However, it was not possible to observe successful ligations, despite the variations in the parameters and colonies observed after transformation, which could have resulted by the self-repairing of the vector after digestion, and random DNA fragments might have been generated.

Another attempt was to make use of the Gateway® recombination cloning method offered by Invitrogen that is based on site-specific recombination properties of bacteriophage λ (Landy, 1989). With this method, the risks of restriction enzyme digestion are avoided and fewer steps are followed, thereby is time-saving and cost-effective and highly efficient (Hartley et al., 2000). Following this technique, the candidate and ready-synthesized genes were all inserted into Gateway® destination vectors, verifying the sequences of the inserts and confirming the transformations at each required step. The destination vectors chosen were pDEST15™, pDEST17™, pEXP1-DEST, and GW-adapted pET29 and pET32, all operating under T7 promoter. The vectors containing the target genes were transformed into BL21 (DE3) pLysE cells for induction, which are chemically competent cells designed to be

used with T7 promoter-based expression vectors. These cells contain DE3 lysogen that carries the gene encoding the T7 RNA polymerase under the control of the *lac*UV5 promoter; therefore the cells are IPTG inducible. They also carry the pLysE plasmid that reduces the basal level expression of the gene of interest due to production of higher amounts of T7 lyzozyme. For induction of the transformed cells, various parameters had to be concerned such as the induction temperature and duration, the concentration of the culture and the concentration of the IPTG. In addition, as selective antibiotic, ampicillin was alternated with carbenicillin to eliminate the stability concerns. The induced cells were then processed both under native and denaturing conditions using Imidazole and urea-guanidine, respectively, as well as with BugBuster extraction reagent. Alternative methods were followed, since there was the possibility of the proteins retaining in the inclusion bodies. After these treatments the lysed cells were passed through Ni-NTA spin columns designed for purification of His-tagged proteins, as all the expression vectors used for this purpose was granting His-tags to the products. Non-induced cells were used as negative control. Samples were also collected after loading the treated cells onto the spin columns and after the wash steps in order to check whether the target proteins was in them, in case they were not binding to the Anti-His column. The eluted samples, along with the others were loaded onto the gel with a colourless loading buffer not to misinterpret the dye as a smaller-sized protein. Unfortunately, none of the trials revealed production of the proteins of interest. In order to make sure that the whole method was working and the results were not due to errors in operation, a positive control was also tested. In

that experiment, the DspF protein was effectively produced. The reason that the candidate genes and the ready-synthesized gene were not induced could be due to the nature of the proteins.

One last attempt was to try inducing the genes of interest with another vector, pFLAG-ATS that allows induction of the inserted gene also by IPTG. Here, the product protein is secreted out by the bacteria as a fusion protein that contains the SP of the *ompA* gene followed by epitope tag FLAG at the N-terminus, to allow purification by ANTI-FLAG magnetic beads or ANTI-FLAG monoclonal antibodies. The method was easy and straight forward, however it required using large batches of cultures which caused a handling problem as well as time management, unless an automated system was available. In addition, it was not possible to express any of the candidate and ready-synthesized genes, with this vector either. In order to check and validate the efficiency of the method, pFLAG-ATS carrying EPIC gene was transformed into BL21 (DE3) pLysE cells and the induction was tried the same way. The results showed that the protein production was induced, suggesting that human-error and consumable problems are unlikely.

Despite using different methodologies, induction of the expression of the candidate genes were not achieved, therefore the proteins they encoded were not tested in subsequent assays such as ROS and GUS assays or seedling growth assays for investigating their roles in suppression and/or activation of the plant defence responses. In order to do that, a complete different expression system, can be tried from scratch in future studies.

6.2.6. Allelic variation of the candidate genes among other *Hpa* isolates

It has been shown that avirulent effectors, especially the members of RxLR family, are under constant pressure for evolution in an arms race for survival (Win et al., 2008; Allen et al., 2004; Rose et al., 2004; Haas et al., 2009). On the other hand, it was observed that, LysM effectors and apoplastic Nep1-like proteins are widely conserved across pathogens (Gizjen and Nürnberger, 2006; Kamoun, 2006). Since PAMPs are also defined as conserved molecules, it was suggested that conserved effectors may not be significantly distinct from PAMPs and can be recognized by cell surface PRRs as well (Thomma et al., 2011). A conserved nature for a pathogen protein increases the potential of getting recognized by the plant receptors, thus puts pressure on the protein to evolve in order to avoid recognition. However, if the protein stays conserved despite the evolutionary demands, they can become useful tools for engineering resistant plants (Bart et al., 2012). Based on this knowledge, the candidate genes' variations were analysed across isolates, to have a better understanding on their recognition, the trend they follow in terms of evolution as well as to evaluate their potential for further use in generating resistant plants. With this purpose, candidates were sequence-analyzed across available *Hpa* isolates, and were compared to the reference sequence of *Hpa*-Emoy2. In general, the results showed most of the candidates were conserved across isolates, and few SNPs resulted in non-synonymous mutations. A closer look to these mutations revealed that the amino acids were mostly changing into a closely-related amino acids, thus the mutation may not result in a

structural or functional alteration for the protein. There were also mutations observed within the signal peptide, raising a concern that the secretion of the protein could be affected. For instance, for *Hpa* 814014, *Hpa*-Cala2 and *Hpa*-Emwa1 showed a change from Valine to Leucine within the SP sequence. This mutation is unlikely to affect the function of the protein or alter the SP since Valine and Leucine are quite similar amino acids with hydrophobic side chains. Additionally, *Hpa*-Emco5, *Hpa*-Edco1 and *Hpa*-Goco1 possessed a change in the coding sequence from Glutamate to Aspartate, again similar acidic amino acids with negatively charged side chains.

For *Hpa* 804480, *Hpa*-Cala2, *Hpa*-Noks1, *Hpa*-Edco1, *Hpa*-Goco1, *Hpa*-Hiks1 and *Hpa*-Maks9 were containing a Serine instead of a Threonine at the same position within the SP. Serine and Threonine are similar amino acids both carrying polar uncharged side chains. In addition, *Hpa*-Maks9 had another switch within the SP from Leucine to Phenylalanine; amino acids both containing hydrophobic side chains. These differences within the SP would not likely affect the secretion of the protein. In addition, *Hpa*-Cala2, *Hpa*-Noks1, *Hpa*-Edco1, *Hpa*-Goco1, *Hpa*-Hiks1 and *Hpa*-Maks9 also demonstrated two other mutations at the same positions where a Proline was switched to Threonine, and Lysine became Arginine. Lysine and Arginine are also quite close in structure as they both contain positively charged side chains. However, a change from Proline to Threonine was significant. Moreover, *Hpa*-Edco1, *Hpa*-Goco1, *Hpa*-Hiks1 and *Hpa*-Maks9 shared a mutation at the same position, from Glycine to Asparagine, which was creating a polar side chain in the protein structure. *Hpa*-Noks1 had a few differences with respect to the

reference sequence, which were changes from Glutamate to Glycine, Isoleucine to Valine, and Glycine to Serine, suggesting a possible functional variation.

On the other hand, few non synonymous mutations were also observed for *Hpa* 814231. *Hpa*-Noks1 possessed a Serine instead of Proline within the SP, adding a polar uncharged side chain to the structure of the signal peptide. *Hpa*-Noks1, *Hpa*-Emco5, *Hpa*-Edco1, *Hpa*-Goco1 and *Hpa*-Hiks1 had a Proline instead of a Threonine in the coding sequence resulting in addition of a cyclic side chain to the protein structure. Additionally, one other mutation was observed on *Hpa*-Goco1, a change of Proline to Serine.

For another candidate, *Hpa* 813915, *Hpa*-Noks1 had noteworthy differences: a Glutamate instead of Lysine, and Arginine instead of a Threonine within the coding sequence. In this case, a negatively charged acidic amino acid was converted into a positively charged basic amino acid and a positively charged amino acid was switched to a polar uncharged amino acid, respectively. *Hpa*-Cala2 and *Hpa*-Maks9 shared a mutation in the coding sequence at the same position from Serine to Arginine. *Hpa*-Cala2 also had a mutation resulting in a change from Serine to Proline, and *Hpa*-Maks9 had Serine instead of a Cysteine, an amino acid with a tendency of building disulphide bridges with other Cysteines, towards the C-terminus. Last but not least, a nonsense mutation observed was in *Hpa*-Edco1 and *Hpa*-Goco1 isolates, in which a Serine was switched to a premature stop codon towards C-terminal, at the same position, resulting in an incomplete protein. A nonsense mutation would result in a serious alteration in the structure, since the incomplete protein probably wouldn't function.

Despite repetitions of PCRs and sequencing, it was not possible to perform a conclusive analysis with the remaining candidate. Nevertheless, it is possible to comment that; overall, most of the mutations seemed unlikely to cause huge changes in the protein folding and function due to the biochemical features of the amino acids they were switched into with few exceptions. The number of variations was still significantly less when compared to the previous findings on other effector proteins, especially the cytoplasmic ones. The findings also raised the question of whether the conserved candidates were in fact perceived as PAMPs and the reason why they did not cause an HR reaction could be that they essentially trigger PTI instead of ETI. In order to have a more comprehensive idea on the extent of the pressure of evolution on the candidate genes, a more detailed and expanded analysis should be performed regarding the Ka/Ks ratio, which is the ratio of the number of non-synonymous substitutions per non-synonymous site (Ka) to the number of synonymous substitutions per synonymous site (Ks), with additional *Hpa* isolates. Based on that, a conclusive note can be put on whether the candidates are widely conserved or in tendency for neutral selection, purifying selection or positive selection. Lastly, determining the receptors corresponding to candidate genes would clear the blur on whether they are functioning as PAMPs or effectors.

6.3. A proteomic approach towards identifying novel apoplastic effectors

Oomycete plant pathogens occupy the plant's intercellular regions, where they initiate the pathogenesis (Kamoun, 2006). This apoplastic area is also rich

in the proteins associated with plant's immune responses (Hammerschmidt, 2010). Therefore, the apoplast can be considered as a reasonable start point for the search of apoplastic effectors and their corresponding receptors. Several studies (Klement, 1965; Rathmell and Sequiera, 1974; Rathmell and Sequiera 1975; De Wit and Spikman, 1982; Boller and Métraux, 1988) showed obtaining the intercellular fluid from infected plant materials via vacuum infiltration-centrifugation as a valid method. Based on that, we aimed to find novel apoplastic effectors, using the *Hpa*-Emoy2-infected and healthy Ws-*eds1* plants as the IWF source, and to investigate their roles in pathogenicity and defence through assays i.e. β -glucuronidase (GUS) reporter system and cumulative quantification of ROS. The total protein material in the IWF was also separated with ion exchange chromatography to use in deductive screenings. In addition, the IWF samples were analyzed via Matrix-assisted laser desorption/ionization coupled to Time-of-flight (MALDI-TOF) in order to identify the proteins looking at their masses and sequences. This MALDI-TOF analysis also allowed selection of T-DNA insertional mutant lines in order to observe the interaction phenotypes and see whether there was a change in resistance/susceptibility attributes.

6.3.1. Confirming intercellular washing fluid as a source of apoplastic molecules via MALDI-TOF analysis

With the purpose of obtaining the apoplastic fluid, a vacuum-infiltration/centrifugation method was followed. IWF was collected not only from the infected but also non-infected plant materials, in order to have control samples for activity assays. Despite being a straight-forward method, there

were many important points to pay attention. Since bigger leaves (and heavily sporulated ones from infected plants) were preferred to end up with a large volume of sample with high concentration of proteins, it was crucial to keep the plants intact during the 5-6 week period. Therefore, extra care was taken in order to keep the non-infected plants healthy and the infected plants protected from secondary infections. Plant leaves were harvested in a Class II Biosafety Cabinet to prevent cross-contamination and thoroughly washed with chilled sdH$_2$O several times to eliminate the spores and other contaminants. The leaves were applied vacuum under a fume hood and dried off just enough to get rid of the remaining excess water on the leaves not to have a dilute sample. It was important to handle the vacuum-infiltrated leaves delicately, as they became brittle, to prevent breaking the cells and contaminating the sample with intracellular fluids, as much as possible. Collected samples were filter-sterilized to remove any remaining contaminants. The samples were not kept in a specific buffer and not added protease inhibitors or any other preservatives, not to have any effects caused by such agents on the results of subsequent assays. However, since the samples were vulnerable for disruption they were tried to be used immediately or within a short amount of time after production. The IWF samples were visualized on SDS-PAGE after every collection, and the followed method was proven to be easy, efficient, reproducible and precise. Both infected and control materials revealed numerous bands on the gel, thereby looking merely at the gel was not enough to have a conclusive idea about the content before using in the assays. Therefore, the samples were processed via MALDI-TOF, which is a handy, advantageous, and sensitive mass spectrometry

that makes use of a soft ionization technique and time of flight for identifying biomolecules and organic molecules including proteins and peptides. There are many studies that have used MALDI-TOF mass spectrometry to investigate the dynamic apoplast (Kwon et al., 2005; Soares et al., 2007, 2009; Szuba et al., 2015). For instance, Haslam et al. (2003) revealed that the extracellular matrix of *Oryza sativa* (Asian rice) contains proteins such as glucanases and germin-like proteins; *A. thaliana* contains substilisin-like proteinases, ascorbate oxidases and glucan endo-1,3-β-D-glucosidases and compared these results to the contents of the whole leaf. Furthermore, Diaz-Vivancos et al. (2006) were able to suggest that the apoplast of the plum pox virus infected *Prunus* spp was rich in ROS (e.g. superoxide dismutases and peroxidases). Moreover, Iglesias et al. (2006) found that the apoplastic fluid of *A. thaliana* was rich in glycanases necessary for the degradation of xyloglucan oligosaccharides. In addition, Delaunois et al. (2013) recovered 177 proteins from grapevine apoplast including PR proteins such as chitinases and proteases.

On the other hand, the MALDI-TOF analysis of our IWF samples revealed 105 proteins from *Hpa*; of those 100 proteins showed similarities to proteins of *Phytophthora* spp and included glucanases, serine proteases and hydrolases. Remaining 5 proteins were linked to *Hpa*, carried potentials for being apoplastic effectors and all contained signal peptides: *Hpa* 806254, *Hpa* 806256, *Hpa* 813024 and *Hpa* 811332 were named as cysteine-rich proteins and *Hpa* 806775 was an elicitin-like protein. Additionally, from *A. thaliana*, 159 up-regulated proteins were found, of which 61 had putative signal peptides. Among them, there were germin-like, chitinase family and peroxidase

superfamily proteins and an LRR family protein. Moreover, 84 down-regulated proteins were annotated, 20 of them had predicted signal peptides. These also included LRR family proteins. The results of MALDI-TOF were quite satisfying and motivating, especially since the BLAST searches indicated towards putative apoplastic effector proteins as well as plant receptors with a high potential of functioning in the apoplastic area.

6.3.2. β-glucuronidase (GUS) reporter activity assays to look for indications of defence activation

Benefiting from the fact that plants lack a detectable β-glucuronidase (GUS) activity, an enzyme that catalyses the breakdown of glucuronides, a technique was developed where GUS gene is used as a gene fusion marker for observing the expression of the desired gene in transformed plants (Jefferson et al., 1987). The expression of the gene of interest can be spotted by observing the GUS enzyme's activity by introducing its colourless or non-fluorescent glucoronide substrates. The reaction will yield coloured or fluorescent products. With the intention of observing whether the IWF samples were triggering any reactions, Col-*RLK-GUS* and Col-*PR-GUS* transgenic lines were used. The samples were boiled to disrupt the 3D structure of the proteins, and were also treated with Proteinase -K enzyme to digest and eliminate the proteins from the samples. The purpose of boiling and degrading the samples was to alter the conformation and destroy the proteins respectively, since it was important to show that any reaction was coming from the proteins in the sample and not from other molecules or contaminants and also to show the nature of

the molecules i.e. the activator is a protein rather than a carbohydrate. These treatments were observed on a gel, and there was a decrease in the amount of the proteins when boiled, and proteinase K was successful in digesting all the proteins but it was staying in the sample.

For injecting the samples, healthy leaves (3 plants, 3 leaves each) were carefully selected and pressure-injected. After the incubation, the treated leaves were gently picked and immersed in GUS staining solution containing the antibiotic chloramphenicol, in order to prevent any contamination and bacterial growth at 37°C. The GUS reporter assay carried out with these samples showed that infected sample was resulting in a considerable reaction. *E. amylovora* was used as positive control in this experiment, at O.D.$_{600}$= 0.25 concentration adjusted in sdH$_2$O, and caused a noteworthy reaction as expected. Despite having merely proteins in its content, the reaction caused by the infected IWF was comparable to *E. amylovora*'s. On the other side, the boiled-infected, healthy and healthy-boiled IWF samples showed very little to no reaction, blue-colour only appearing at the sites where the edges of the syringe tip was pressing and not at the infused sites, similar to the negative control, indicating structurally disrupted proteins are not capable of activating defence responses, and the proteins in the healthy IWF sample are not recognized as non-self to trigger immune activation. Proteinase-K-treated samples were not constructive for the evaluation of the results since the proteinase-K itself was already causing a great reaction. According to these preliminary outcomes, it was clear that the contents of the infected IWF sample were triggering a reaction that was causing the expression of the PR and RLK encoding

genes. The significant decrease in the reaction when the sample was boiled, can point out that an intact protein structure is essential, and boiling the samples is sufficient to alter this structure. At this point, it can also be said that the infected IWF sample in fact contains the pathogen derived molecules that are directly or indirectly involved in the mechanism of activation of the defence response.

6.3.3. Reactive oxygen species as a strong proof of activation of immune responses

MALDI-TOF analysis and GUS assays results; showed that the IWF samples contained molecules that are able to trigger defence responses. It order to support this idea, another assay was performed with these samples: quantitative investigation of ROS accumulation. It is known that, immediately after pathogen entry, plants produce ROS molecules in order to create a hostile environment for the pathogen and stop the further invasion as a part of the first layer of defence, the PTI. While, pathogens try to overcome PTI with their effectors, plants respond with a greater and stronger production of as a part of ETI. For that reason, measuring the production of ROS can indicate a host defence activation. Felix et al. (1999) suggested observing the ROS production with an assay, in which, plant materials are subjected to treatments and subsequently a substrate (5-AS) is introduced that would be processed by the extracellular peroxidases. As a result of that, the colourless substrate will be converted to orange-brown coloured product that can be measured by spectrophotometry at 450 nm. Overall, in this assay, the activity of extracellular

peroxidases which are involved in ROS production and are a part of a system parallel to the NADPH-dependent oxidase pathway (Torres et al., 2005), will be quantified as a sign of triggered immune responses. Therefore, in this assay, plant materials were subjected to IWF samples and were allowed enough time for the second wave of ROS production. The experiments were carried on a range of plant materials: Col-0, Col-*rpp4* and Col-*bak1-5/bkk1-1* and *N. benthamiana* plants. These plants were carefully selected, since it was desired to observe the reactions given by resistant and non-host plants in addition the ones with impaired receptors for pathogen perception. Col-0 plants are resistant to *Hpa*-Emoy2 relying on ATR4 recognition by RPP4 (Holub et al., 1994); hence the mutant line Col-*rpp4* is susceptible to infection. On the other hand, Col-*bak1-5/bkk1-1* were also screened since, the mutated genes have a role as an adaptor kinase and a co-receptor in perception of the pathogen and involve in disease resistance against *Hpa* (Roux et al., 2011).

During the experiments, it was crucial to keep the environment free from contaminants the whole time. In addition, the plant disks required delicate care. After cutting off the disks, they were given a wash with the buffer that will be used in the assay, in order to remove any contaminants and molecules released by the damage. Moreover, it was important to keep the damage caused on the disks to a minimum during transferring to the microtiter plates, to prevent subsequent release of the DAMPs, as they could alter the results. On the plate, the edge wells were solely filled with sdH$_2$O to minimize evaporation and prevent the edge effect, and the plates were sealed and incubated under light with gentle agitation, to keep the plants active for the maximum development of

the ROS production. All assays were repeated multiple times, since it was important to show reproducibility and consistency of the assay. Additionally, assays were executed by obtaining leaf disks from the same leaf per treatment, different leaves on the same plant, leaves on the same batch and different batches, as well as using treatment samples obtained at different times, to show that results were independent of these materials.

Basic statistical analysis of the results showed that the infected IWF sample was causing an immediate reaction with a high magnitude, compared to all other treatments in all of the assays performed with different plants. On the other hand, healthy IWF samples and positive control flg22 were developing their reactions with time and reaching saturation. In comparison, the healthy samples were causing considerably less reaction than infected ones even after 2 min of incubation. These measurements also presented a routine for these samples among different plant materials. Moreover, both boiled healthy and infected IWF samples demonstrated results in accordance to negative control, supporting the findings of the GUS assay and proposing that an intact structure of the proteins is essential to trigger any reaction. In a general look, it can be said that the cumulative quantification of ROS production can be reckoned as a result of activation of ETI by the samples. The figures revealed by the mutants Col-*rpp4* and Col-*bak1-5/bkk1-1* compared to Col-0, were suggesting that the detection of the molecules in the IWF was not solely or at all, depending on these receptors; therefore other players might have a role in their recognition as a result of an arms-race evolution. However, the positive control flg22 was expected to cause a rather less reaction in the assay on Col-*bak1-5/bkk1-1*

plants, since this double mutant is impaired in the perception of flg22. One other issue was that the healthy IWF sample was causing an evident ROS production in *N. benthamiana* plants, which could be due to recognition of the molecules coming from *Arabidopsis* as non-self. In addition, the reactions observed from the negative controls can be a result of any damage on the leaves caused during the conduction of the experiment, as normally they should cause negligible reactions. A conclusive word can be that the difference between the reactions caused by the infected and healthy IWF materials are due to the molecules coming from the pathogen, and this assay is a suitable way to assess defence activations.

6.3.4. Simplification of the complex IWF samples via ion exchange chromatography

Since both GUS and ROS assays were revealing consistent and supporting positive results, a simplification of the complex nature of the IWF samples was planned in order to define the actual actors that were triggering reactions within IWF samples. There are many available techniques for separating compound samples or purifying a target protein. Among them, ion exchange chromatography has been a popular method, which is exploits a separation depending on the charges of the proteins (Cummins et al., 2011). The technique is grounded on the reversible bond built between a charged protein and an oppositely charged chromatography medium. Simply put, if surrounding pH is higher than the protein's isoelectric point (pI), the protein will bind to a positively charged chromatography medium (anion exchanger) and

when the pH is lower than its pI, the protein will bind to a negatively charged solid support (cation exchanger) (Amersham Biosciences, 2002). Where the case is pH is equal to protein's pI, then the protein will not be attached to a charged medium. The technique can be practiced using columns filled with the preferred medium, by loading the column with the sample to be separated and eluting with an increasing concentration of salt. Strong anion exchange chromatography column was used for the IWF samples. The columns were found to be easy to use, cost-efficient and reliable, however slightly more time consuming with respect to the automated systems. Prior to the separation, it was crucial to desalt the samples, since linear salt gradients were going to be used to elute fractions. During desalting, the buffer was switched to the prepared start buffer as well, as was suggested in manufacturer's instructions. While the samples were being loaded into the column, the flow-through was also collected in order to catch the proteins that were not binding to the column. This FT was also considered as a fraction and was visualized on the SDS-PAGE along with the other fractions. According to the gel, it was possible to say that the separation was successful and, the fractions as well as the FT contained a broad range of assorted proteins. It was natural to expect that the first fraction obtained with the lowest salt gradient (0.1 M) would show similar bands to that of FT, since that elution would contain the proteins making the weakest bond with the medium; however on the gel different bands were clearly observed. On the other hand, there were also similar bands for the fractions eluted with close salt concentrations. It is hard to say that the similar bands represent the same proteins, solely by looking at the gel that separated the

protein based on their sizes. In addition, the number of proteins was smaller for the fractions eluted with higher salt concentrations, which suggested that there not many proteins retaining in the column, thus not many proteins with a pI value too low, existed in the sample. Overall, the method was found to be solid and easy to conduct, yet the elutions were very dilute and required condensation. Therefore, the elutions were precipitated in cold acetone and resuspended in a buffer to achieve effective solubilisation (see section 2.2.33).

6.3.5. Identification of the fraction containing the molecules responsible for activating immune responses

The fractions obtained by chromatography from the infected IWF sample were assessed for their abilities to trigger ROS production. All concentration-adjusted fractions were used to treat Col-0 plants in a ROS assay, using the buffer they were resuspended in as negative control in addition to sdH_2O, and horseradish peroxidase and flg22 as positive controls. Multiple repetitions of the assays consistently revealed that the most activity was achieved by the FT fraction, and the reactions caused by the other fractions were not significant. It was also observed that the resuspension buffer was not causing any reaction which relieved the concern of the results being affected by the precipitation and resuspension steps. The fractions obtained from the healthy IWF samples were also used in ROS assays; and no significant reaction was observed for any of the fractions. It could be supposed that the previous reactions caused by the non-separated healthy IWF samples might be due to protein-protein interactions and they were inactive after separation. On the other hand, the positive results

observed from the FT fraction of the infected IWF sample, suggested that this fraction should be the one to go further with subsequent assays. The only concern with using fractions in activity assays was that the volume of sample obtained in every batch was quite small and assays were requiring considerable amount of proteins. Due to the limited capacity of the ion exchange columns it was not possible to load more samples to start with. Instead, to end up with larger volume of fractions with higher concentrations, we had to repeat the chromatography multiple times and combine the fractions. This also required collection of IWF samples in an extremely controlled, dedicated and continuous manner, a process that can be easily altered by secondary reasons.

6.3.6. Insertional mutagenesis as a handy tool for protein function discovery

The gene knock-outs and knock-downs are important tools for finding out the function of that gene's product using reverse genetics approach. One widely-used method to perform these alterations is insertion of foreign DNA in the target gene, a process called insertional mutagenesis, which can be achieved by *Agrobacterium*-mediated transformation of T-DNA. This technique allows investigation the function of a protein by interfering with the structure of the encoding gene and observing the phenotypic changes. When T-DNA is inserted into the genome, it not only interrupt's the target gene's function but also acts as a marker (Krysan et al., 1999; Wang, 2008). This method was also applied to *Arabidopsis* Col-0 accession in order to create a genome-wide mutagenesis and a library of mutant lines (Alonso et al., 2003). Taking

advantage of this library, we chose three proteins annotated as LRR protein family from MALDI-TOF results, It has been known that LRR protein family members are known to carry effector recognition properties (van der Hoorn and Kamoun, 2008), all containing a signal peptide; hence most probably function in the apoplast. The aim was to use the mutant lines to determine their involvement in recognition of the proteins in the infected IWF and in resistance/susceptibility for different *Hpa* isolates. With this purpose, the mutant lines of the selected proteins were ordered from the NASC library. It was necessary to confirm the insertions and select homozygous lines for further assays. After analysing all mutant lines, homozygous plant were determined for, N660880, N665050, N859846 and N661217; however, it was not possible to obtain homozygous lines for N547337 and N579946. All the seeds in the ordered batch turned out to be heterozygous. It should have been possible to further grow the plants to obtain seeds and select the homozygous lines out of segregation, however, these lines were lost to secondary infections, and were not possible to carry forward. Therefore, sporulation and ROS assays were performed with the remaining lines. Two different homozygous plants were used in the assays for each mutant line.

First assay was for observing the interaction phenotypes of these mutants with *Hpa*-Emoy2, *Hpa*-Cala2 and *Hpa*-Noks1 isolates, and comparing to Col-0 and Ws-*eds1* plants. Col-0 is naturally resistant to *Hpa*-Emoy2 and *Hpa*-Cala2, whereas, Ws-*eds1* is highly susceptible. The mutants inoculated with these isolates also showed resistance marks and no sporulation similar to Col-0, suggesting that selected LRR genes' mutations did not result in enhanced

susceptibility when compared to the controls. On the other hand, Col-0 is susceptible to *Hpa*-Noks1 isolate and Ws-eds1 is highly susceptible. The mutant lines were also showing similar results to Col-0, proposing no significant enhancement in disease susceptibility. These results were demonstrating that the mutation of these LRR on their own was not effective on changing to perception the effectors of these isolates, leading to the idea that the receptor might be collaborating with co-receptors, or the site of mutation is not causing a noteworthy change in the function of the gene.

The second assay was a ROS assay, assessed by using the FT fraction of the infected IWF. The attributes of the mutant lines were compared to Col-0 and previous findings, and the reactions caused by the mutant lines were in accordance with the Col-0 plants, supporting the sporulation assay results.

In order to have a better understanding of the roles of these LRR proteins, a wider range of mutant lines, as well as double mutants, can be obtained or generated to be used in functionality assays.

7. CONCLUSIONS

Healthy and sustainable crop production has been a challenge for centuries, due to many threats, one of which being the plant diseases caused by a broad range of plant pathogens (Boyd et al., 2012). Oomycetes, a distinct group of eukaryotic phytopathogens, are also responsible for a significant number of catastrophic and commonly-encountered plant diseases including late blight, sudden oak death, root rot, white rust and downy mildew (Kamoun, 2003). With the motivation of contributing to improvement of control measures against downy mildew by highlighting the underlying molecular mechanisms, in this study we focused on an established model pathosystem *of A. thaliana and H. arabidopsidis* (Coates and Beynon, 2010). Oomycetes, including the *Hpa*, are shown to be remarkably rich in the number of effectors, molecules which interfere with the plant defence by acting in the apoplast or plant cell's cytoplasm to create a favourable environment to establish pathogenicity. With improving technologies and genomic discoveries, especially in the last decade, substantial progress has been made in the study of effectors; yet, there is still so much to uncover. Due to the fact that the apoplastic effectors lack motifs for translocation, can be recognized by apoplastic plant receptors as well as cell surface receptors, and may trigger intracellular defence mechanisms; identifying apoplastic effectors have been more changeling with respect to the RxLR effectors (Tör, 2008). Here, we were inspired by the bioinformatic tools and

available genomic data libraries, and decided to investigate novel apoplastic effectors from *Hpa* and scrutinize their roles in activation/suppression of the plant immunity. We took two approaches with this purpose: First was, to dig through the EST data of the *Hpa*-Emoy2 isolate to determine candidate putative apoplastic effectors, regarding the criteria defined by previous studies. Therefore, five small genes with signal peptides, without RxLR (and dEER) or LxLFLAK motifs, without potential cytoplasmic functions were chosen. Initially, the expression of the genes was validated. Following that, the timing and the expression levels of the candidate genes were analysed. According to the expression patterns, the genes of interest were expressed at different times and with different intensities, suggesting that, they may differ in pathways of secretion from the pathogen and perception by the plant, as well as their exact roles on triggering plant immunity. For a more detailed analysis on the expression patterns, it might be a good idea to extend the time period and continue observing the daily or hourly changes in the expression levels for the duration of the infection. On the other hand, the candidate genes were transiently expressed *in planta* in order to observe whether they were able to trigger hypersensitive response, a local cell death at the infection site that occurs as a result of effector recognition. When compared to positive control, the candidate genes were not triggering an HR. This may indicate that the *Nicotiana* spp may not be suitable plants for examining this response or the candidate genes may not be recognized as effectors. Therefore, an extended assay for transient expression can be tried on different plants; and additional assays can be performed to measure other immune responses such as callose

deposition and changes of ion flux as a part of future work. In addition, same system can be used for localization assays by fusing the genes with fluorescent proteins. Another attempt for investigating the candidate genes' functions was stable transformation of wild type resistant Col-0 plants to create transgenic lines and observe the changes in interaction phenotypes. However, due to secondary fungal infections, it was not possible to carry this effort further. In future studies, alternative methods for generating transgenic lines can be practiced such as gene bombardment. Separately, *in vitro* expression of the candidate genes by inducing bacterial cultures was tried. Several different techniques were practiced to achieve this however; we were not able to obtain the proteins they encoded. It would be ideal to have the encoded proteins to investigate their roles in pathogenicity, ability to trigger defence responses and corresponding receptors. Therefore, in future studies, alternative transformation methods and expression systems such as yeast or virus-based can be utilized. Last but not least, allelic variations of the candidate genes were investigated. The genes of interest were amplified in different *Hpa* isolates in order to check whether they were conserved or showing polymorphisms. This analysis revealed that, mostly the candidates were conserved; if not, few SNPs were detected, majority of them resulting in synonymous mutations. This suggested that, with respect to cytoplasmic effectors, the candidates are probably under less pressure to evolve to avoid recognition, or they are acting like PAMPs and adopting a conserved nature. A more comprehensive future study for analysing the evolutionary stress on these candidates can involve other *Hpa* isolates and

go into statistical details to determine whether these genes are subjected to positive or purifying selection.

The second approach was collection of the intercellular washing fluid, which is the fluid in the dynamic apoplast, to investigate the content in deductive screenings. To observe whether the IWF of the infected plant material was able to trigger defence responses, GUS and ROS assays were performed. The former was performed on transgenic Col plants carrying PR-GUS and RLK-GUS fusions, the proteins involved in pathogen perception. In this assay, the infected IWF was causing significant reactions, indicating defence activation by the molecules in the infected IWF. The latter was quantification of the extracellular peroxidases as an indication of activated host immunity. For this assay, we used not only wild type plant but also mutants impaired in RPP4 receptor of the effector ATR4 (Holub et al., 1994), and BAK1, co-receptor that helps with signal transduction (Liebrand et al., 2014). The infected IWF was again causing a considerably high reaction in both wild type and mutant plants, supporting previous findings. In addition, the results showed that these receptors are highly unlikely to be involved in perception of the molecules within the infected IWF. According to these results, we decided to simplify the complex IWF sample via strong anion exchange chromatography, in order to define an active fraction for subsequent assays. One fraction, the flow-through, consisting of the proteins that did not bind to the chromatography column was found to be the active one in ROS assays, hinting that the proteins triggering the observed reactions are within this fraction. On the other hand, the complex IWF samples were analysed via MALDI-TOF mass spectrometry, which revealed a large

number of proteins, both coming from the pathogen and the plant. Majority of these proteins also contained a signal peptide, meaning they will be secreted hence will take position in the apoplast. Most of the proteins coming from the pathogen were showing similarities to proteins of *Phytophthora* spp, the remaining was assigned to *Hpa*. These included cysteine-rich proteins and elicitin-like proteins that carried potential of being apoplastic effectors. Two of these genes were used in *in vitro* expression experiments. An additional study can focus on all of these *Hpa*-sourced proteins to scrutinize their behaviours and assign functions. On the other hand, among the proteins coming from the *A. thaliana,* there were proteins annotated as LRR family proteins which are known receptors for PAMPs and effectors. Mutant lines for three of the LRR family proteins were obtained from the library and were used for interaction phenotype assays and ROS assays with the flow-through fraction. Our findings did not suggest that these receptors were involved in recognizing the molecules in the flow-through fraction, and impairment of these receptors did not have an effect on enhancing susceptibility or resistance of the plants to *Hpa* isolates on their own. Based on these findings, the active fraction can be further separated to narrow down the complexity and used in pull-down experiments for hunting their receptors.

Overall, it can be said that the candidate genes still carry the potential of being apoplastic effectors, but also can be perceived as PAMPs; therefore they have to be thoroughly investigated to annotate a function. In addition, the active fraction of IWF should be further simplified and tested on functionality assays.

Moreover, both the proteins encoded by candidate genes and the ones within the active fraction should be confirmed for their apoplastic localisations. Following that, a broader perspective would be making use of these proteins to generate resistant transgenic lines and apply to closely-related higher plants such as Brassicae, in terms of control measures against the downy mildew disease.

Additional work:

Involvement of the Electrophilic Isothiocyanate Sulforaphane in Arabidopsis Local Defense Responses

During my PhD study I had the chance to contribute to the research on the involvement of the electrophilic isothiocyanate sulforaphane in Arabidopsis local defence responses by carrying out the trypan blue staining stages of the experiments. The paper was published in *Plant Physiology* journal (Vol. 167 pp. 251-261) in January 2015 (Andersson et al., 2015, see Paper). Below I provide a summary of the study including a brief background information along with the part I was involved in.

Summary

Plants are continuously under attack by microbial pathogens (Faulkner and Robatzek, 2012), and they defend themselves by means of thorough and complex physical and molecular mechanisms. Plants make use of their firm cell walls made of complex polysaccharides, the waxy layer and the cuticles as a physical barrier against pathogens. In order to penetrate the plant, the pathogens force mechanical pressure and secrete enzymes to overcome these barriers. The degradation of plant cell walls however, produce phenolic and toxic substances that creates an unfavourable environment for the pathogens (Hématy et al., 2009). Despite that, the pathogens generally survive and proceed forward by penetrating the host plant cell and releasing molecules to ascertain the infection. As a response, plants recognize the molecules produced by the damage given by the pathogen in addition to the molecules the pathogens release when they are inside the plant (Staal and Dixelius, 2009). In some cases, induced defence responses such as thickening of the cell wall or production of

phytoalexins may stop the advancing pathogens (Monaghan and Zipfel, 2012). However, an adapted and evolved pathogen can surmount these defence responses by secreting molecules called effectors (Bozkurt et al., 2012). The effector molecules can manipulate the host plant's cellular processes to their own advantages by avoiding recognition and delaying or suppressing defence responses (Birch et al., 2008; Bent and Mackey, 2007). In return, plants have evolved to recognize these effectors by host R proteins (Coates and Beynon, 2010). This recognition often results in hypersensitive responses (HR) that is a defence response involving programmed cell death at the infection site, while activating other local and systemic defence responses (Mur et al., 2008). Generally, the cell death around the infection site can be clearly observed (Coates and Beynon, 2010). In addition, other defence pathways are induced away from the infection site and systemic tissue is alerted for defence. Such molecules and compounds that are systemically acting have been investigated and candidates have been identified (Dempsey and Klessig, 2012).

However, despite the HR phenomenon has been known for a long time (Stakman, 1915), the role players in the HR at the local infection sites are mostly under shadow. Therefore, in this study our main objective was to try to identify molecules and substances that are produced and released during HR that are inducing cell death in naïve tissue.

Following thorough investigations we reported that, after effector recognition, the model plant *Arabidopsis thaliana*'s leaf tissue undergoing the HR, produces and releases a reactive electrophilic compound named sulforaphane (4-methylsulfinylbutyl isothiocyanate), which is a plant defensive compound derived from glucosinolate precursors (Fahey et al., 2001), and this compound induces cell death as well as alerting the naïve tissue. A reduced programmed cell death following bacterial and oomycete effector recognition was observed in addition to decreased resistance to

several isolates of the pathogen *Hyaloperonospora arabidopsidis* in two different mutants impaired in pathogen-induced accumulation of sulforaphane. Moreover, sulforaphane-pre-treated plants displayed enhanced resistance against a virulent isolate of *H. arabidopsidis*. Furthermore, mutants affected in sulforaphane production as well as other glucosinolate breakdown products, revealed reduced or delayed cell death after effector recognition and decreased resistance to an oomycete pathogen. It was deduced that sulforaphane and similar compounds are highly likely to have antimicrobial properties and they also trigger local defence responses in Arabidopsis tissue.

Contribution (Fig 6B-C in paper - pp. 256)

In order to further investigate the function of sulforaphane in plant's local defence responses, the oomycete *H. arabidopsidis* was used as pathogen. Arabidopsis wild type Col-0 carries the Recognition Peronospora Parasitica2 (RPP2) gene that enable recognition of isolates *Hpa* Cala2 and trigger HR (van der Biezen et al., 2002; Sinapidou et al., 2004). Wild type Col-0 and double mutant plants *myb28 myb29* and *tgg1 tgg2* were inoculated with the isolate *Hpa* Cala2, and the cell death at the infection sites was evaluated via trypan blue staining. The Arabidopsis seedlings were sprayed with 400 mM sulforaphane suspended in deionized water 24 h before or after inoculation with *Hpa* spores as indicated and the cell death was observed by trypan blue staining method described by Koch and Slusarenko (1990). A rapid localized cell death was observed on 90% of the interaction sites and there was no visible pathogen growth in wild type Col-0 seedlings (Fig. 6B). Approximately 10% of the interaction sites showed a certain level of trailing necrosis and growth of hyphae within the given time period (Fig. 6C).

Figure 6. Reduced resistance to a biotrophic oomycete in sulforaphane-deficient mutants. The indicated lines were inoculated at the cotyledon stage with *Hpa* conidia of the isolates Cala2 (A–D) or Emwa1 (E). The cotyledons were stained with Trypan Blue 7 (A–C) or 2 (D) dpi, and the extent of cell death was determined or sporophores were counted at 7 dpi (E). Rapid cell death in the wild type is shown in B, and trailing necrosis is shown in C. Average and range values for three replicate experiments, each including 200 interaction sites, are shown in A and D. Average and sp for 15 cotyledons are shown in E. Letters a to c indicate statistically significant groups (one-way ANOVA with Tukey's posthoc test, $P < 0.05$). The experiments were performed three times with similar results. Bars in B and C = 0.5 mm.

References

Abramovitch, R.B. & Martin, G.B. (2004) Strategies used by bacterial pathogens to suppress plant defenses. *Current Opinions in Plant Biology,* 7, 356-364.

Agrios, G. N. 1997. Plant Pathology. Fourth ed. New York: Academic Press. 635 pp.

Allen, R.L., Bittner-Eddy, P.D., Grenville-Briggs, L.J., Meitz, J.C., Rehmany, A.P., Rose, L.E. & Beynon, J.L. (2004) Host–parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science*, 306, 1957–1960.

Alonso, J.M., Stepanova, A.N., Leisse T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Horn, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. & Ecker, J.R. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana. Science*, 301, 653-657.

Amersham Biosciences. (2002) *Ion exchange chromatography, principles and methods*. Björkgatan, Uppsala, Sweden, Amersham Pharmacia Biotech.

Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408, 796-815.

Armstrong, M.R., Whisson, S.C., Pritchard, L., Bos, J.I., Venter, E., Avrova, A. O., Rehmany, A.P., Böhme, U., Brooks, K., Cherevach, I., Hamlin, N., White, B., Fraser, A., Lord, A., Quail, M.A., Churcher, C., Hall, N., Berriman, M., Huang, S., Kamoun, S., Beynon, J.L., & Birch, P.R. (2005) An ancestral oomycete locus contains late   blight avirulence geneAvr3a, encoding a protein that is recognized in the host cytoplasm. *PNAS USA*, 102, 7766–7771.

Bailey, K., Çevik, V., Holton, N., Byrne-Richardson, J., Sohn, K.H., Coates, M., Woods-Tör, A., Aksoy, H.M., Hughes, L., Baxter, L., Jones, J.D., Beynon, J., Holub, E.B. & Tör, M. (2011) Molecular cloning of ATR5 (Emoy2) from *Hyaloperonospora arabidopsidis*, an avirulence determinant that triggers RPP5-mediated defense in *Arabidopsis*. *Molecular Plant Microbe Interactions*, 24 (7), 827-838.

Baldauf, S.L., Roger, A.J., Wenk-Siefert, I. & Doolittle, W.F. (2000) A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science,* 290*,* 972–977.

Bart, R., Cohn, M., Kassen, A., McCallum, E.J., Shybut, M., Petriello, A., Krasileva, K., Dahlbeck, D., Medina, C., Alicai, T., Kumar, L., Moreira, L.M., Rodrigues Neto, J., Verdier, V., Santana, M.A., Kositcharoenkul, N., Vanderschuren, H., Gruissem, W., Bernal, A. & Staskawicz, B.J. (2012) High-throughput genomic sequencing of cassava bacterial blight strains

identifies conserved effectors to target for durable resistance. *PNAS USA*, 109 (28), E1972-1979.

Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral A., Kemen E., Thines, M., Ah-Fong, A., Anderson, R., Badejoko, W., Bittner-Eddy, P., Boore, J.L., Chibucos, M.C., Coates, M., Dehal, P., Delehaunty, K., Dong, S., Downton, P., Dumas, B., Fabro, G., Fronick, C., Fuerstenberg, S.I., Fulton, L., Gaulin, E., Govers, F., Hughes, L., Humphray, S., Jiang, R.H.Y., Judelson, H., Kamoun, S., Kyung, K., Meijer, H., Minx, P., Morris, P., Nelson, J., Phuntumart, V., Qutob, D., Rehmany, A., Rougon-Cardoso, A., Ryden, P., Torto-Alalibo, T., Studholme, D., Wang, Y., Win, J., Wood, J., Clifton, S.W., Rogers, J., Van den Ackerveken, G., Jones, J.D.G., McDowell, J.M., Beynon, J. & Tyler, B.M. (2010) Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science*, 330, 1549-1551.

Bendtsen, J.D., Nielsen, H., von Heijne, G. & Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology*, 340, 783-795.

Bernoux, M., Ellis, J.G. & Dodds, P.N. (2011) New insights in plant immunity signalling activation. *Current Opinion in Plant Biology*, 14, 1-7.

Birch, P.R.J., Armstrong, M., Bos, J., Boevink, P.C., Gilroy, E.M., Taylor, R.M., Wawra, S., Pritchard, L., Conti, L., Ewan, R., Whisson, S.C., Van West, P., Sadanandom, A. & Kamoun, S. (2009) Towards understanding the virulence

functions of RXLR effectors of the oomycete plant pathogen *Phytophthora infestans*. *Journal of Experimental Botany*, 60 (4), 1133-1140.

Birch, P.R.J., Boevink, P.C., Gilroy, E.M., Hein, I., Pritchard, L. & Whisson S.C. (2008) Oomycete RXLR effectors: delivery, functional redundancy and durable disease resistance. *Current Opinion in Plant Biology*, 11 (4), 373-379.

Birch, P.R., Rehmany, A.P., Pritchard, L., Kamoun, S. & Beynon, J.L. (2006) Trafficking arms: oomycete effectors enter host plant cells. *Trends in Microbiology*, 14 (1), 8-11.

Bishop, J.G., Ripoll, D.R., Bashir, S., Damasceno, C.M., Seeds, J.D. & Rose, J.K. (2004) Selection on *Glycine* beta-1,3-endoglucanase genes differentially inhibited by a *Phytophthora* glucanase inhibitor protein. *Genetics,* 169, 1009–1019.

Bittner-Eddy, P.D., Allen, R.L., Rehmany, A.P., Birch, P. & Beynon, J.L. (2003) Use of suppression subtractive hybridization to identify downy mildew genes expressed during infection of *Arabidopsis thaliana*. *Molecular Plant Pathology*, 4, 501–507.

Bittner-Eddy, P., Can, C., Gunn, N., Pinel, M., Tör, M., Crute, I., Holub, E.B. & Beynon, J. (1999) Genetic and physical mapping of the RPP13 locus, in *Arabidopsis*, responsible for specific recognition of several *Peronospora parasitica* (downy mildew) isolates. *Molecular Plant Microbe Interactions*, 12, 792–802.

Bittner-Eddy, P.D., Crute, I.R., Holub, E.B. & Beynon, J.L. (2000) *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. *Plant Journal*, 21, 177–188.

Boissy, G., O'Donohue, M., Gaudemer, O., Perez, V., Pernollet, J.C., Brunie, S. (1999) The 2.1 A structure of an elicitin-ergosterol complex: a recent addition to the Sterol Carrier Protein family. *Protein Science*, 8, 1191–1199.

Boller, T. & Métraux, J.P. (1988) Extracellular localization of chitinases in cucumber. *Physiological and Molecular Plant Pathology*, 33, 11-16.

Bos, J.I., Armstrong, M.R., Gilroy, E.M., Boevink, P.C., Hein, I., Taylor, R.M., Zhendong, T., Engelhardt, S., Vetukuri, R.R., Harrower, B., Dixelius, C., Bryan, G., Sadanandom, A., Whisson, S.C., Kamoun, S. & Birch, P.R. (2010) *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. PNAS USA, 107 (21), 9909-9014.

Bos, J.I., Kanneganti, T.D., Young, C., Çakır, C., Huitema, E., Win, J., Armstrong, M.R., Birch, P.R. & Kamoun, S. (2006) The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*. *Plant Journal*, 48 (2), 165-176.

Botella, M.A., Parker, J.E., Frost, L.N., Bittner-Eddy, P.D., Beynon, J.L., Daniels, M.J., Holub, E.B. & Jones, J.D. (1998) Three genes of the

Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants. *Plant Cell*, 10,1847–1860.

Bouwmeester, K., Meijer, H.J. & Govers, F. (2011) At the frontier; RXLR effectors crossing the *Phytophthora*-host interface. *Frontiers in Plant Science*, 2, 75.

Boyd, L.A., Ridout, C., O'Sullivan, D.M., Leach, J.E. & Leung, H. (2012) Plant-pathogen interactions: disease resistance in modern agriculture. *Trends in Genetics*, 29 (4), 233-240.

Bozkurt, T.O., Schornack, S., Banfield, M.J. & Kamoun, S. (2012) Oomycetes, effectors, and all that jazz. *Current Opinion in Plant Biology*, 15, 483-492.

Bradford, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248-254.

Brunner, F., Rosahl, S., Lee, J., Rudd, J.J., Geiler, C., Kauppinen, S., Rasmussen G., Scheel, D. & Nürnberger, T. (2002) Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora* transglutaminases. *EMBO Journal*, 21 (24), 6681-6688.

Bruno, S., Duschak, V.G., Ledesma, B., Ferella, M. & Andersson, B. (2004) Identification and characterization of serine proteinase inhibitors from *Neospora caninum. Molecular Biochemical Parasitology,* 136, 101–107.

Cabral, A., Oome, S., Sander, N., Küfner, I., Nürnberger, T. & Van den Ackerveken, G. (2012) Nontoxic Nep1-like proteins of the downy mildew pathogen *Hyaloperonospora arabidopsidis*: repression of necrosis-inducing activity by a surface-exposed region. *Molecular Plant Microbe Interactions*, 25 (5), 697-708.

Cabral, A., Stassen, J.H.M., Seidl, M.F., Bautor, J., Parker, J.E. & Van den Ackerveken, G. (2011) Identification of *Hyaloperonospora arabidopsidis* transcript sequences expressed during infection reveals isolate-specific effectors. *PLoS One,* 6 (5), 1-13.

Canut, H., Carrasco, A., Galaud, J.P., Cassan, C., Bouyssou, H., Vita, N., Ferrara, P. & Pont-Lezica, R. (1998) High affinity RGD-binding sites at the plasma membrane of *Arabidopsis thaliana* links the cell wall. *Plant Journal*, 16 (1), 63-71.

Catanzariti, A.M., Dodds, P.N, & Ellis, J.G. (2007) Avirulence proteins from haustoria-forming pathogens. *FEMS Microbiology Letters,* 269, 181–188.

Chinchilla, D., Shan, L., He, P., De Vries, S. & Kemmerling, B. (2009) One for all: the receptor-associated kinase BAK1. *Trends in Plant Science*, 14 (10), 535-541.

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J.D., Felix, G. & Boller, T. (2007) A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defense. *Nature*, 448 (7152), 497-500.

Chisholm, S.T., Coaker, G., Day, B. & Staskawicz, B.J. (2006) Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell,* 124, 803-814.

Chung, Y., Choe, V., Fujioka, S., Takatsuto, S., Han, M., Jeon, J.S., Park, Y.I., Lee, K.O. & Choe, S. (2012) Constitutive activation of brassinosteroid signaling in the *Arabidopsis* elongated-D/bak1 mutant. *Plant Molecular Biology*, 80 (4-5), 489-501.

Clough, S.J. & Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, 16, 735-743.

Coates, M.E. & Beynon, J.L. (2010) *Hyaloperonospora arabidopsidis* as a pathogen model. *Annual Review of Pythopathology*, 48, 329-345.

Collmer, A., Badel, J.L. , Charkowski, A.O., Deng, W., Fouts, D.E., Ramos, A.R., Rehm, A.H., Anderson, D.M., Schneewind, O., Van Dijk, K. & Alfano, J.R. (2000) *Pseudomonas syringae* Hrp type III secretion system and effector proteins. *PNAS*, 97 (16), 8770-8777.

Constantinescu, O. & Fatehi, J. (2002) *Peronospora*-like fungi (Chromista, Peronosporales) parasitic on *Brassicaceae* and related hosts. *Nova Hedwigia*, 74, 291–338.

Cummins. P.M., Downling, O. & O'Connor. B.F. (2011) Ion-exchange chromatography: basic principles and application to the partial purification of

soluble mammalian prolyl oligopeptidase. *Methods in Molecular Biology*, 681, 215-228.

Dardick, C. & Ronald, P. (2006) Plant and animal pathogen recognition receptors signal through non-RD kinases. *PLOS Pathogens*, 2 (1), e2.

Dardick, C., Schwessinger, B. & Ronald, P. (2012) Non-arginine-aspartate (non-RD) kinases are associated with innate immune receptors that recognize conserved microbial signatures. *Current Opinion in Plant Biology*, 15, 358-366.

Das, P. & Joshi, N.C. (2011) Minor modifications in obtainable *Arabidopsis* floral dip method enhances transformation efficiency and production of homozygous transgenic lines harboring a single copy of transgene. *Advances in Bioscience and Biotechnology*, 2, 59-67.

Dautova, M., Rosso, M., Abad, P., Gommers, F.J., Bakker, J. & Smant, G. (2001) Single pass cDNA sequencing—A powerful tool to analyze gene expression in preparasitic juveniles of the southern root-knot nematode *Meloidogyne incognita. Nematology*, 3,129-139.

Daxberger, A., Nemak, A., Mithofer, A., Fliegmann, J., Ligterink, W., Hirt, H. & Ebel, J. (2007). Activation of members of a MAPK module in beta-glucan elicitor mediated non-host resistance of soybean. *Planta,* 225, 1559–1571.

De Jonge, R., Bolton, M.D. & Thomma, B.P.H.J. (2011) How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. *Current Opinion in Plant Biology*, 14, 1-7.

De Wit, P.J.G.M. & Spikman, G. (1982) Evidence of the occurrence of race and cultivar-specific elicitors of necrosis in intercellular fluids of compatible interactions of *Cladosporium fulvum* and tomato. *Physiological Plant Pathology*, 21, 1-8.

Degrave, A., Fagard, M., Perino, C., Brisset, M.N., Gaubert, S., Laroche, S., Patrit, O. & Barny, M.A. (2008) *Erwinia amylovora* type three-secreted proteins trigger cell death and defense responses in *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions*, 21 (8), 1076-1086.

Delaunois, B., Colby, T., Belloy, N., Conreux, A., Harzen, A., Baillieul, F., Clément, C., Schmidt, J., Jeandet, P. & Cordelier, S. (2013) Large-scale proteomic analysis of the grapevine leaf apoplastic fluid reveals mainly stress-related proteins and cell wall modifying enzymes. *BMC Plant Biology*, 13, 24.

Diaz-Vivancos, P., Rubio, M., Mesonero, V., Periago, P.M., Barceló, A.R., Martínez-Gómez, P. & Hernández, J.A. (2006) The apoplastic antioxidant system in Prunus: response to long-term plum pox virus infection. *Journal of Experimental Botany*, 57 (14), 3812-3824.

Dodds, P.N., Rafiqi, M., Gan, P.H.P., Hardham, A.R., Jones, D.A. & Ellis, J.G. (2009) Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytologist,* 183, 993-1000.

Dodds, P.N. & Rathjen, J.P. (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics*, 11, 539-548.

Dong, S., Yin, W., Kong, G., Yang, X., Qutob, D., Chen, Q., Kale, S.D., Sui, Y., Zhang, Z., Dou, D., Zheng, X., Gijzen, M., Tyler, B.M. & Wang, Y. (2011) *Phytophthora sojae* avirulence effector Avr3b is a secreted NADH and ADP-ribose pyrophosphorylase that modulates plant immunity. *PloS Pathogens*, 7 (11): e1002353.

Dou, D., Kale, S.D., Wang, X., Jiang, R.H.Y., Bruce, N.A., Arredondo, F.D., Zhang, X. & Tyler, B.M. (2008) RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. *Plant Cell*, 20 (7), 1930-1947.

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K. & Pikaard, C.S. (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal*, 45, 616-629.

Ellis, J.G. & Dodds, P.N. (2011) Showdown at the RXLR motif: serious differences of opinion in how effector proteins from filamentous eukaryotic pathogens enter plant cells. *PNAS*, 108(35), 14381-14382.

Ensemble Protists. http://www.ensembl.org/index.html & http://protists.ensembl.org/Hyaloperonospora_arabidopsidis/Info/Index.

Erwin, D.C., Bartnicki-Garcia, S. & Tsao, P.H. (1983) *Phytophthora: its biology, taxonomy, ecology, and pathology*. St Paul, Minnestota, APS Press.

Faulkner, C. & Robatzek, S. (2012) Plants and pathogens: putting infection strategies and defence mechanisms on the map. *Current Opinion in Plant Biology*, 15, 699-707.

Felix, G., Duran, J.D., Volko, S. & Boller, T. (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal,* 18 (3), 265-276.

Gassman, W. & Bhattacharjee, S. (2012) Effector-triggered immunity signalling: from gene-for-gene pathways to protein-protein interaction networks. *Molecular Plant-Microbe Interactions*, 25 (7), 862-868.

Gäumann, E. (1918) Über die formen der *Peronospora parasitica* (Pers.) fries: ein beitrag zur spezies frage bei den parasitischenpilzen. *Beihefte zum Botanischen Zentralblatt,* 35 (Abt 1), 395–533.

Gijzen, M. & Nürnberger T. (2006) Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. *Phytochemistry*, 67, 1800–1807.

Gómez-Gómez, L. & Boller, T. (2000) FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell*, 5 (6), 1003-1011.

Gouget, A., Senchou, V., Govers, F., Sanson, A., Barre, A., Rougé, P., Pont-Lezica, R. & Canut, H. (2006) Lectin receptor kinases participate in protein-protein interactions to mediate plasma membrane-cell wall adhesions in *Arabidopsis. Plant Physiology*, 140 (1), 81-90.

Göker, M., Riethmüller, A., Voglmayr, H., Weiss, M. & Oberwinkler, F. (2004) Phylogeny of *Hyaloperonospora* based on nuclear ribosomal internal transcribed spacer sequences.*Mycoogical.Progress,* 3, 83–94.

Greeff, C., Roux, M., Mundy, J. & Petersen, M. (2012) Receptor-like kinase complexes in plant innate immunity. *Frontiers in Plant Science,* 3, 1-7.

Grouffaud, S., Whisson, S. C., Birch, P.R.J. & Van West, P. (2010) Towards an understanding on how RxLR-effector proteins are translocated from oomycetes into host cells. *Fungal Biology Reviews*, 24, 27-36.

Gust, A.A. & Felix, G. (2014) Receptor like proteins associate with SOBIR1 - type of adaptors to form bimolecular receptor kinases. *Current Opinion in Plant Biology*, 21, 104-111.

Haas, B.J., Kamoun, S., Zody, M.C., Jiang, R.H., Handsaker, R.E., Cano, L.M., Grabherr, M., Kodira, C.D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A.M., Alvarado, L., Anderson, V.L., Armstrong, M.R., Avrova, A. Baxter, L., Beynon, J., Boevink, P.C., Bollmann, S.R., Bos, J.I., Bulone, V., Cai, G., Çakır, C., Carrington, J.C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M.A., Fugelstad, J., Gilroy, E.M., Gnerre, S., Green, P.J., Grenville-Briggs, L.J., Griffith, J., Grünwald, N.J., Horn, K., Horner, N.R., Hu, C.H., Huitema, E., Jeong, D.H., Jones, A.M., Jones, J.D., Jones, R.W., Karlsson, E.K., Kunjeti, S.G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M.C., McDonald, H., McWalters, J., Meijer, H.J., Morgan, W., Morris, P.F., Munro, C.A., O'Neill, K., Ospina-Giraldo, M., Pinzón, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D.C., Schumann, U.D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D.J., Sykes, S., Thines, M., Van de Vondervoort, P.J.,

Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B.C., Van West, P., Ristaino, J., Govers, F., Birch, P.R., Whisson, S.C., Judelson, H.S. & Nusbaum, C. (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*, 461 (7262), 393-398.

Haldar, K., Kamoun, S., Hiller, N.L., Bhattacharje, S. & Van Ooij, C. (2006) Common infection strategies of pathogenic eukaryotes. *Nature Reviews Microbiology*, 4 (12), 922-931.

Hammerschmidt, R. (2010) The dynamic apoplast. *Physiological and Molecular Plant Pathology*, 74, 199-200.

Hardham, A.R., Cahill, D.M., Cope, M., Gabor, B.K., Gubler, F. & Hyde, G.J. (1994) Cell surface antigens of *Phytophthora* spores: biological and taxonomic characterization. *Protoplasma,* 181, 213–232.

Hartley, J.L., Temple, G.F. & Brasch, M.A. (2000) DNA cloning using *in vitro* site-specific recombination. *Genome Research*, 10, 1788-1795.

Haslam, R.P., Downie A.L., Raveton, M.K., Gallardo, K., Job, D., Pallett, K.E., John, P., Parry, M.A.J. & Coleman, J.O.D. (2003) The assessment of enriched apoplastic extracts using proteomic approaches. *Annals of Applied Biology*, 143 (1), 81-91.

Hématy, K., Cherk, C. & Somerville, S. (2009) Host-pathogen warfare at the plant cell wall. *Current Opinion in Plant Biology*, 12, 406–413.

Hiller, N.L., Bhattacharjee, S., Van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C., Haldar K. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science*, 306, 1934–1937.

Hogenhout, S.A., Van der Hoorn, R.A.L., Terauchi, R. & Kamoun, S. (2009) Emerging concepts in effector biology of plant-associated organisms. *Molecular Plant-Microbe Interactions*, 22 (2), 115-122.

Holub, E.B. (2006) Evolution of symbioses between plants and filamentous parasites. *Current Opinion of Plant Biology*, 9, 397–405.

Holub, E.B. (2008) Natural history of *Arabidopsis thaliana* and oomycete symbioses. *European Journal of Plant Pathology,* 122, 91-109.

Holub, E.B., Beynon, J.L. & Crute, I.R. (1994) Phenotypic and genotypic characterization of interactions between isolates of *Peronospora parasitica* and accessions of *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions,* 7, 223-239.

Iglesias, N., Abelenda, J.A., Rodiño, M., Sampedro, J., Revilla, G. & Zarra, I. Apoplastic glycosidases active against xyloglucan oligosaccharides of *Arabidopsis thaliana*. *Plant Cell Physiology*, 47 (1), 55-63.

Jefferson, R.A., Kavanagh, T.A. & Bevan, M.W. (1987) GUS-fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal*, 6 (13), 3901-3907.

Jiang, R.H. & Tyler, B.M. (2012) Mechanisms and evolution of virulence in oomycetes. *Annual Review of Phytopathology*, 50, 295-318.

Jin, S., Xu, R., Wei, Y. & Goodwin, P.H. (1999) Increased expression of a plant actin gene during a biotrophic interaction between round-leaved mallow, *Malva pusilla*, and *Colletotrichum gloeosporioides f. sp. malvae*. *Planta*, 1999, 209 (4), 487-494.

Johnson, L.N., Noble, M.E. & Owen, D.J. (1996) Active and inactive protein kinases: structural basis for regulation. Cell, 85 (2), 149-158.

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

Judelson, H.S. & Blanco, F.A. (2005) The spores of *Phytophthora*: weapons of the plant destroyer. *Nature Reviews Microbiology*, 3, 47-58.

Kale, S.D. (2012) Oomycete and fungal effector entry, a microbial Trojan horse. *New Phytologist*, 193 (4), 874-881.

Kale, S.D. & Tyler, B.M. (2011) Entry of oomycete and fungal effectors into plant animal host cells. *Cellular Microbiology*, 13 (12), 1839-1848.

Kamoun, S., (2003) Molecular genetics of pathogenic oomycetes. *Eukaryotic Cell*, 10, 191-199.

Kamoun, S. (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Annual Review of Phytopathology*, 44, 41-60.

Kamoun, S., Honée, G., Weide, R., Laugé, R., Kooman-Gersmann, M., De Groot, K. Govers, F. & De Wit P.J.G.M. (1999a) The fungal gene Avr9 and the oomycete gene inf1 confer avirulence to potato virus X on tobacco. *Molecular Plant-Microbe Interactions*, 12, 459–462.

Kamoun, S., Hraber, P., Sobral, B., Nuss, D. & Govers, F. (1999b) Initial assessment of gene diversity for the oomycete pathogen *Phytophthora infestans* based on expressed sequences. *Fungal Genetics and Biology*, 28, 94-106.

Kamoun, S., Lindqvist, H. & Govers, F. (1997b) A novel class of elicitin-like genes from *Phytophthora infestans. Molecular Plant-Microbe Interactions.* 10, 1028–1030.

Kamoun, S., Van West, P., De Jong, A.J., Vleeshouwers, V.G.A.A. & Govers, F. (1997a) A gene encoding a protein elicitor of *Phytophthora infestans* is down-regulated during infection of potato. *Molecular Plant-Microbe Interactions*, 10, 13-20.

Kamoun, S., van West, P., Vleeshouwers, V.G., de Groot, K.E. & Govers, F. (1998) Resistance of nicotiana benthamiana to phytophthora infestans is mediated by the recognition of the elicitor protein INF1. *Plant Cell*, 10 (9), 1413-1426.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P. & Drummond, A. (2012). Geneious Basic: an integrated and

extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28* (12), 1647-1649.

Keeling, P.J., Burger, G., Durnford, D.G., Lang, B.F., Lee, R.W., Pearlman, R.E., Roger, A.J. & Gray, M.W. (2005) The tree of eukaryotes. *Trends in Ecology & Evolution (Personal Edition)*, 20, 670-676.

Kemen, E., Gardiner, A., Schultz-Larsen, T., Kemen, A.C., Balmuth, A.L., Robert-Seilaniantz, A., Bailey, K., Holub, E., Studholme, D.J. & Maclean, D. (2011) Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana. PLoS Biology*, 9, e1001094.

Kemen, E. & Jones, J.D.G. (2012) Obligate biotroph parasitism: can we link genomes to lifestyles? *Trends in Plant Science*, 17 (8) 448-457.

Kim, J.F. & Beer, S.V. (1998) HrpW of *Erwinia amylovora*, a new harpin that contains a domain homologous to pectate lyases of a distinct class. *Journal of Bacteriology*, 180 (19), 5203-5210.

Klement, Z. (1965) Method of obtaining fluid from the intercellular spaces of foliage and the fluid's merit as substrate for phytobacterial pathogens. *Phytopathology*, 53, 1033-1034.

Knoth, C. & Eulgem, T. (2008) The oomycete response gene *LURP1* is required for defense against *Hyaloperonospora parasitica* in *Arabidopsis thaliana. The Plant Journal*, 55, 53-64.

Koch, E. & Slusarenko, A. (1990) *Arabidopsis* is susceptible to infection by a downy mildew fungus. *The Plant Cell*, 2, 437-445.

Krajaejun, T., Khositnithikul, R., Lerksuthirat, T., Lowhnoo, T., Rujirawat, T. Petchthong, T, Yingyong, W., Suriyaphol, P., Smittipat, N., Juthayothin, T., Phuntumart, V. & Sullivan, T.D. (2011) Expressed sequence tags reveal genetic diversity and putative virulence factors of the pathogenic oomycete *Pythium insidiosum*. *Fungal Biology*, 115, 683-696.

Krupa, A., Preethi, G. & Srinivasan, N. (2004) Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases. *Journal of Molecular Biology*, 339, 1025-1039.

Krysan, P.J., Young, J.C. & Sussman, M.R. (1999) T-DNA as an insertional mutagen in *Arabidopsis*. *The Plant Cell*, 11, 2283-2290.

Kunjeti, S.G., Evans, T.A., Marsh, A.G., Gregory, N.F., Kunjeti, S., Meyers, B.C., Kalavacharla, V.S. & Donofrio, N.M. (2012) RNA-Seq reveals infection-related global gene changes in *Phytophthora phaseoli*, the causal agent of lima bean downy mildew. *Molecular Plant Pathology*, 13, 454-466.

Kwon, H.K., Yokoyama, R. & Nishitani, K. (2005) A proteomic approach to apoplastic proteins involved in cell wall regeneration in protoplasts of *Arabidopsis* suspension-cultured cells. *Plant Cell Physiology*, 46 (6), 843-857.

Kwon-Chung, K.J. (1994) Phylogenetic spectrum of fungi that are pathogenic to humans. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 19 (Suppl. 1), S1-S7.

Landy, A.(1989) Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annual Review of Biochemistry*, 58, 913-949.

Lapin, D. & Van den Ackerveken, G. (2013) Susceptibility to plant disease: more than a failure of host immunity. *Trends in Plant Science*, 18 (10), 546-554.

Larroque, M., Ramirez, D., Lafitte, C., Borderies, G., Dumas, B. & Gaulin, E. (2011) Expression and purification of a biologically active *Phytophthora parasitica* cellulose binding elicitor lectin in *Pichia pastoris. Protein Expression and Purification*, 80 (2), 217-223.

Laugé, R. & De Wit, P.J.M. (1998) Fungal avirulence genes: structure and possible functions. *Fungal Genetics and Biology*, 24, 285–297.

Lee, S.W., Han, S.W., Bartley, L.E. & Ronald, P.C. (2006) From the Academy: Colloquium review. Unique characteristics of *Xanthomonas oryzae pv. oryzae* AvrXa21 and implications for plant innate immunity. *PNAS USA*, 103 (49), 18395-18400.

Letunic, I., Doerks, T. & Bork, P. (2014) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research*, 43 (D1), D257-D260.

Lévesque, C.A., Brouwer, H., Cano, L., Hamilton, J.P., Holt, C., Huitema, E., Raffaele, S., Robideau, G.P., Thines, M., Win, J., Zerillo, M.M., Beakes, G.W., Boore, J.L., Busam, D., Dumas, B., Ferriera, S., Fuerstenberg, S.I., Gachon, C.M., Gaulin, E., Govers, F., Grenville-Briggs, L., Horner, N., Hostetler, J., Jiang, R.H., Johnson, J., Krajaejun, T., Lin, H., Meijer, H.J., Moore, B., Morris, P., Phuntmart, V., Puiu, D., Shetty, J., Stajich, J.E., Tripathy, S., Wawra, S., Van West, P., Whitty, B.R., Coutinho, P.M., Henrissat, B., Martin, F., Thomas, P.D., Tyler, B.M., De Vries, R.P., Kamoun, S., Yandell, M., Tisserat, N. & Buell, C.R. (2010) Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biology,* 11, R73.

Li, J., Park, E., Von Arnim, A.G. & Nebenführ, A. (2009) The FAST technique: a simplified *Agrobacterium*-based transformation method for transient gene expression analysis in seedlings of *Arabidopsis* and other plant species. *Plant Methods*, 5 (6), 1-15.

Li, J., Wen, J., Lease, K.A., Doke, J.T., Tax, F.E. & Walker, J.C. (2002) BAK1, an *Arabidopsis* LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell*, 110 (2), 213-222.

Liebrand, T.W.H., Van de Burg, H.A. & Joosten, M.H. (2014) Two for all: receptor associated kinases SOBIR1 and BAK1. *Trends in Plant Science*, 19 (2), 123-132.

Lindh, J.G., Botero-Kleiven, S., Arboleda, J.I. & Wahlgren, M. (2001) A protease inhibitor associated with the surface of *Toxoplasma gondii*. *Molecular Biochemical Parasitology*, 116, 137–145.

Links, M.G., Holub, E., Jiang, R.H., Sharpe, A.G., Hegedus, D., Beynon, E., Sillito, D., Clarke, W.E., Uzuhashi, S. & Borhan, M.H. (2011) De novo sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. *BMC Genomics*, 12, 503.

Liu, Z., Bos, J.I.B., Armstrong, M., Whisson, S.C., Da Cunha, L., Torto-Alalibo, T., Win, J., Avrova, A.O., Wright, F., Birch P.R.J. & Kamoun, S. (2005) Patterns of diversifying selection in the phytotoxin-like scr74gene family of *Phytophthora infestans. Molecular Biology and Evolution*, 22, 659–672.

Luderer, R., De Kock, M.J.D., Dees, R.H.L., De Wit, P.J.M. & Joosten, M.H. (2002) Functional analysis of cysteine residues of ECP elicitor proteins of the fungal tomato pathogen *Cladosporium fulvum. Molecular Plant Pathology*, 2, 91-95.

Luna, E., Pastor, V., Robert, J., Flors, V., Mauch-Mani, B. & Ton, J. (2011) Callose deposition: a multifaceted plant defense response. *Molecular Plant-Microbe Interactions*, 24 (2), 183-193.

Margulis, L., & Schwartz, K.V. (2000) Five kingdoms: an illustrated guide to the phyla of life on earth. W. H. Freeman and Co., New York, N.Y.

Mauch, F. & Staehelin, L.A. (1989) Functional implications of the subcellular localization of ethylene-induced chitinase and β-1,3-glucanase in bean leaves. *The Plant Cell*, 1, 447-457.

McDowell, J.M. (2011) Genomes of obligate plant pathogens reveal adaptations for obligate parasitism. *PNAS USA*, 108 (22), 8921-8922.

Medzhitov, R. (2007) Recognition of microorganisms and activation of the immune response. *Nature*, 449, 819–826.

Meinke, D.W., Cherry, J.M., Dean, C., Rounsley, S.D. & Koornneef, M. (1998) *Arabidopsis thaliana*: a model plant for genome analysis. *Science*, 282, 662-682.

Mikes, V., Milat, M.L., Ponchet, M., Panabieres, F., Ricci, P. & Blein, J.P. (1998) Elicitins, proteinaceous elicitors of plant defense, are a new class of sterol carrier proteins. *Biochemical and Biophysical Research Communications*, 245, 133–139.

Mikes, V., Milat, M.L., Ponchet, M., Ricci, P. & Blein, J.P. (1997) The fungal elicitor cryptogein is a sterol carrier protein. *FEBS Letters*, 416, 190–192.

Mims, C.W., Richardson, E.A, Holt, B.F.III & Dangl, J. L. (2004) Ultrastructure of the host-pathogen interface in *Arabidopsis thaliana* leaves infected by the downy mildew *Hyaloperonospora parasitica*. *Canadian Journal of Botany*, 82, 1001–1008.

Misas-Villamil, J.C. & Van der Hoorn, R.A.L. (2008) Enzyme-inhibitor interactions at the plant-pathogen interface. *Current Opinion in Plant Pathology*, 11, 380-388.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H. & Shibuya, N. (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis. PNAS USA*, 104 (49), 19613-19618.

Monaghan, J. & Zipfel, C. (2012) Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology*, 15, 349-357.

Muthamilarasan, M. & Prasad, M. (2013) Plant innate immunity: An updated insight into defense mechanism. *Journal of Biosciences*, 38 (2), 1-17.

Nagaraj, S.H., Gasser, R.B. & Ranganathan, S. (2006) A hitchhiker's guide to expressed sequence tag (EST) analysis. *Briefings in Bioinformatics*, 8 (1), 6-21.

Nespoulous, C., Gaudemer, O., Huet, J.C. & Pernollet, J.C. (1999) Characterization of elicitin like phospholipases isolated from *Phytophthora capsici* culture filtrate. *FEBS Letters*, 452, 400–406.

Nicastro, G., Orsomando, G., Ferrari, E., Manconi, L., Desario, F., Amici, A., Naso, A., Carpaneto, A., Pertinhez, T.A., Ruggieri, S. & Spisni, A. (2009) Solution structure of the phytotoxic protein PcF: the first characterized member of the *Phytophthora* PcF toxin family. *Protein Science*, 18 (8), 1786-1791.

Orsomando, G., Brunetti, L., Pucci, K., Ruggeri, B. & Ruggieri, S. (2011) Comparative structural and functional characterization of putative protein effectors belonging to the PcF toxin family from *Phytophthora* spp. *Protein Science*, 20 (12), 2047-2059.

Orsomando, G., Lorenzi, M., Ferrari, E., de Chiara, C., Spisni, A. & Ruggieri, S. (2003) PcF protein from *Phytophthora cactorum* and its recombinant homologue elicit phenylalanine ammonia lyase activation in tomato. *Cellular and Molecular Life Sciences*, 60, 1470–76.

Osman, H., Mikes, V., Milat, M.L., Ponchet, M., Marion, D., Prangé, T., Maume, B.F., Vauthrin, S. & Blein, J.P. (2001a). Fatty acids bind to the fungal elicitor cryptogein and compete with sterols. *FEBS Letters*, 489, 55–58.

Osman, H., Vauthrin, S., Mikes, V., Milat, M. L., Panabières, F., Marais, A., Brunie, S., Maume, B., Ponchet, M. & Blein, J.P. (2001b) Mediation of elicitin activity on tobacco is assumed by elicitin-sterol complexes. *Molecular Biology of the Cell*, 12, 2825–2834.

Ottmann, C., Luberacki, B., Küfner, I., Koch, W., Brunner, F., Weyand, M., Mattinen, L., Pirhonen, M., Anderluh, G., Seitz, H.U., Nürnberger, T. & Oecking, C. (2009) A common toxin fold mediates microbial attack and plant defense. *PNAS USA*, 106 (25), 10359-10364.

Pais, M., Win, J., Yoshida, K., Etherington, G.J., Cano, L.M., Raffaele, S., Banfield, M.J., Jones, A., Kamoun, S. & Saunders, D.G. (2013) From

pathogen genomes to host plant processes: the power of plant parasitic oomycetes. *Genome Biology*, 14 (6), 211.

Parker, J.E., Holub, E.B., Frost, L.N., Falk, A., Gunn, N.D. & Daniels, M.J. (1996) Characterization of *eds1*, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different *RPP* genes. *The Plant Cell*, 8, 2033-2046.

Peresen, C., Van Themaat, E.V.L., McGuffin, L.J., Abbott, J.C., Burgis, T.A., Barton, G., Bindschedler, L.V., Lu, X., Maekawa, T., Wessling, R., Cramer, R., Thordal-Christensen, H., Panstruga, R. & Spanu, P.D. (2012) Structure and evolution of barley powdery mildew effector candidates. *BMC Genomics*, 13, 694-713.

Pemberton, C. L. & Salmond, G.P.C. (2004) The Nep1-like proteins a growing family of microbial elicitors of plant necrosis. *Molecular Plant Pathology*, 5, 353-359.

Ponchet, M., Panabières, F., Milat, M. L., Mikes, V., Montillet, J.L., Suty, L., Triantaphylides, C., Tirilly, Y. & Blein, J.P. (1999) Are elicitins cryptograms in plant-Oomycete communications? *Cellular and Molecular Life Sciences*, 56 (11-12), 1020-1047.

Punja, Z.K. (2004) *Fungal Disease Resistance in Plants*. Binghamton, NY: Food Products Press. 266 pp.

Qutob, D., Hraber, P.T., Sobral, B.W. & Gijzen, M. (2000) Comparative analysis of expressed sequences in *Phytophthora sojae*. *Plant Physiology*, 123, 243-254.

Qutob, D., Kemmerling, B., Brunner, F., Küfner, I., Engelhardt, S., Gust, A.A., Luberacki, B., Seit, H.U., Stahl, D., Rauhut, T., Glawischnig, E., Schween, G., Lacombe, B., Watanabe, N., Lam, E., Schlichting, R., Scheel, D., Nau, K., Dodt, G., Hubert, D., Gijzen, M., & Nürnberger, T. (2006) Phytotoxicity and innate immune responses induced by Nep1-like proteins. *Plant Cell*, 18 (12), 3721–3744.

Raffaele, S., Farrer, R.A., Cano, L.M., Studholme, D.J., MacLean, D., Thines, M., Jiang, R.H., Zody, M.C., Kunjeti, S.G. & Donofrio, N.M. (2010) Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science*, 330, 1540-1543.

Randall, T.A., Dwyer, R.A., Huitema, E., Beyer, K., Cvitanich, C., Kelkar, H., Fong, A.M., Gates, K., Roberts, S., Yatzkan, E., Gaffney, T., Law, M., Testa, A., Torto-Alalibo, T., Zhang, M., Zheng, L., Mueller, E., Windass, J., Binder, A., Birch, P.R., Gisi, U., Govers, F., Gow, N.A., Mauch, F., van West, P., Waugh, M.E., Yu, J., Boller, T., Kamoun, S., Lam, S.T. & Judelson, H.S. (2005) Large-scale gene discovery in the oomycete Phytophthora infestans reveals likely components of phytopathogenicity shared with true fungi. *Molecular Plant Microbe Interactions,* 18 (3), 229-243.

Rathmell, W.G. & Sequiera, L. (1974) Soluble peroxidase in fluid from intercellular spaces of tobacco leaves. *Plant Physiology*, 53, 317-318.

Rathmell, W.G. & Sequiera, L. (1975) Induced resistance in tobacco leaves – role of inhibitors of bacterial growth in intercellular fluid. *Physiological Plant Pathology*, 5, 65-73.

Rehmany, A.P., Gordon, A., Rose, L.E., Allen, R.L., Armstrong, M.R., Whisson, S.C., Kamoun, S., Tyler, B.M., Birch, P.R.J. & Beynon, J.L. (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two *Arabidopsis* lines. *The Plant Cell*, 17 (6), 1839-1850.

Rehmany, A.P., Grenville, L.J., Gunn, N.D., Allen, R.L., Paniwnyk, Z., Byrne, J., Whisson, S.C., Birch, P.R.J. & Beynon, J.L. (2003) A genetic interval and physical contig spanning the *Peronospora parasitica* (*At*) avirulence gene locus ATR1*Nd. Fungal Genetics and Biology*, 38, 33–42.

Ricci, P., Bonnet, P., Huet, J.C., Sallantin, M., Beauvais-Cante, F., Bruneteau, M., Billard, V., Michel, G. & Pernollet, J.C. (1989) Structure and activity of proteins from pathogenic fungi *Phytophthora* eliciting necrosis and acquired resistance in tobacco. *European Journal of Biochemistry*, 183 (3), 555-563.

Robatzek, S., Chinchilla, D. & Boller, T. (2006) Ligand-induced endocytosis of the pattern recognition receptor FLS2 in *Arabidopsis. Genes and Development*, 20, 537–542.

Ronald, P. (2011) Plant genetics, sustainable agriculture and global food security. *Genetics*, 188, 11-20.

Ronald, P.C. & Shirasu, K. (2012) Front-runners in plant-microbe interactions. *Current Opinion in Plant Biology*, 15, 345-348.

Rose, L.E., Bittner-Eddy, P.D., Langley, C.H., Holub, E.B., Michelmore, R.W. & Beynon, J.L. (2004) The maintenance of extreme amino acid diversity at the disease resistance gene, RPP13, in *Arabidopsis thaliana*. *Genetics*, 166 (3), 1517-1527.

Rose, J.K., Ham, K.S., Darvill, A.G. & Albersheim, P. (2002) Molecular cloning and characterization of glucanase inhibitor proteins: coevolution of a counterdefense mechanism by plant pathogens. *Plant Cell*, 14, 1329–1345.

Roux, M., Schwessinger, B., Albrecht, C., Chinchilla, D., Jones, A., Holton, N., Malinovsky, F.G., Tör, M., de Vries, S. & Zipfel, C. (2011) The *Arabidopsis* leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. *Plant Cell*, 23, 2440-2455.

Sasabe, M., Takeuchi, K., Kamoun, S., Ichinose, Y., Govers, F., Toyoda, K., Shiraishi, T. & Yamada, T. (2000) Independent pathways leading to apoptotic cell death, oxidative burst and defense gene expression in response to elicitin in tobacco cell suspension culture. *European Journal of Biochemistry*, 267, 5005–5013.

Schindler, M., Meiners, S., Cheresh, D.A. (1989) RGD-dependent linkage between plant cell wall and plasma membrane: consequences for growth. *Journal of Cell Biology*, 108, 1955-1965.

Schmidt, S.M. & Panstruga, R. (2011) Pathogenomics of fungal plant parasites: what we have learnt about pathogenesis? *Current Opinion in Plant Biology*, 14, 1-8.

Schornack, S., Van Damme, M., Bozkurt, T.O., Cano, L.M., Smoker, M., Thines, M., Gaulin, E., Kamoun, S. & Huitema, E. (2010) Ancient class of translocated oomycete effectors targets the host nucleus. *PNAS USA*, 107, 17421-17426.

Schultz, J,. Milpetz, F., Bork, P. & Ponting, C.P. (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *PNAS USA*, 95 (11), 5857-5864.

Schwessinger, B. & Zipfel, C. (2008) News from the frontline: recent insights into PAMP-triggered immunity in plants. *Current Opinion in Plant Biology*, 11, 389-395.

Senchou, V., Weide, R., Carrasco, A., Bouyssou, H., Pont-Lezica, R., Govers, F. & Canut, H. (2004) High affinity recognition of a *Phytophthora* protein by *Arabidopsis* via an RGD motif. *Cellular and Molecular Life Sciences*, 61 (4), 502-509.

Shan, W., Cao, M., Leung, D. & Tyler, B.M. (2004) The *Avr1b* locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene *Rps1b*. *Molecular Plant Microbe Interactions*, 17, 394–403.

Sierra, R., Rodriguez, R.L., Chaves, D., Pinzon, A., Grajales, A., Rojas, A., Mutis, G., Cárdenas, M., Burbano, D., Jiménez, P., Bernal, A. & Restrepo, S. (2010) Discovery of *Phytophthora infestans* genes expressed *in planta* through mining of cDNA libraries. *PLoS One*, 5(3), e9847.

Skinner, W., Keon, J. & Hargreaves, J. (2001) Gene information for fungal plant pathogens from expressed sequences. Current Opinion in *Microbiology*, 4, 381-386.

Slusarenko, A.J. & Schlaich, N.L. (2003) Downy mildew of *Arabidopsis thaliana* caused by *Hyaloperonospora parasitica* (formerly *Peronospora parasitica*). *Molecular Plant Pathology*, 4, 159–170.

Soares, N.C., Francisco, R., Ricardo, C.P. & Jackson, P.A. (2007) Proteomics of ionically bound and soluble extracellular proteins in *Medicago truncatula* leaves. *Proteomics*, 7 (12), 2070-2082.

Soares, N.C., Francisco, R., Vielba, J.M., Ricardo, C.P. & Jackson, P.A. (2009) Associating wound-related changes in the apoplast proteome of *Medicago* with early steps in the ROS signal-transduction pathway. *Journal of Proteome Research*, 8 (5), 2298-2309.

Sohn, K.H., Lei, R., Nemri, A. & Jones, J.D. (2007) The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in *Arabidopsis thaliana. Plant Cell*, 19 (12), 4077-4090.

Song, W.Y., Wang, G.L., Chen, L.L., Kim, H.S., Pi, L.Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H., Fauquet, C. & Ronald, P. A receptor

kinase-like protein encoded by the rice disease resistance gene, Xa21. Science, 270 (5243), 1804-1806.

Song, J., Win, J., Tian, M., Schornack, S., Kaschani, F., Ilyas, M., Van der Hoorn, R. A. & Kamoun, S. (2009) Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3. *PNAS*, 106 (5), 1654-1659.

Staal, J. & Dixelius, C. (2009) *Plant innate immunity*. Encyclopaedia of Life Sciences. Chichester, John Wiley & Sons Ltd.

Stam, R., Mantelin, S., McLellan, H. & Thilliez, G. (2014) The role of effectors in nonhost resistance to filamentous plant pathogens. *Frontiers in Plant Science*, 5 (582), 1-6.

Stassen, J.H.M. & Van den Ackerveken, G. (2011) How do oomycete effectors interfere with plant life? *Current Opinion in Plant Biology*, 14, 1-8.

Stassen, J.H.M., Seidl, M.F., Vergeer, P.W.J., Nijman, I.J., Snel, B., Cuppen, E. & Van den Ackerveken, G. (2012) Effector identification in the lettuce downy mildew *Bremia lactucae* by massively parallel transcriptome sequencing. *Molecular Plant Pathology*, 13 (7), 719-731.

Szuba, A., Kasprowicz-Maluśki, A. & Wojtaszek, P. (2015) Nitration of plant apoplastic proteins from cell suspension cultures. *Journal of Proteomics*, 29 (120), 158-168.

Thomma, B.P.H.J., Nürnberger, T & Joosten, M.H.A.J. (2011) Of PAMPs and Effectors: The Blurred PTI-ETI Dichotomy. *The Plant Cell*, 23 (1), 4-15.

Tian, M., Benedetti, B. & Kamoun, S. (2005) A second Kazal-like protease inhibitor from *Phytophthora infestans* inhibits and interacts with the apoplastic pathogenesis-related protease P69B of tomato. *Plant Physiology*, 138, 1785–1793.

Tian, M., Huitema, E., Da Cunha, L., Torto-Alalibo, T. & Kamoun, S. (2004) A Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. *Journal of Biological Chemistry*, 279, 26370–26377.

Tian, M. & Kamoun, S. (2005) A two disulfide bridge Kazal domain from *Phytophthora* exhibitsstable inhibitory activity against serine proteases of the subtilisin family. *BMC Biochemistry*, 6, 15.

Tian, M., Win, J., Song, J., Van der Hoorn, R., Van der Knaap, E. & Kamoun, S. (2007) A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-Like apoplastic protease. *Plant Physiology*, 143 (1), 364-367.

Torres, M.A., Jones, J.D.G. & Dangl, J.L. (2005) Pathogen-induced, NADPH oxidase-derived reactive oxygen intermediates suppress spread of cell death in *Arabidopsis thaliana*. *Nature Genetics*, 37 (10), 1130-1134.

Torto, T., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N.A., Van West, P. & Kamoun, S. (2003) EST mining and functional expression assays identify

extracellular effector proteins from the plant pathogen *Phytophthora. Genome Research*, 13, 1675-1685.

Torto-Alalibo, T.A., Tripathy, S., Smith, B.M., Arredondo, F.D., Zhou, L., Li, H., Chibucos, M.C., Qutob, D., Gijzen, M., Mao, C., Sobral, B.W., Waugh, M.E., Mitchell, T.K., Dean, R.A., Tyler, B.M. (2007) Expressed sequence tags from *Phytophthora sojae* reveal genes specific to development and infection. *Molecular Plant Microbe Interactions,* 20 (7), 781-793.

Tör, M. (2008) Tapping into molecular conversation between oomycete plant pathogens and their hosts. *European Journal of Plant Pathology*, 122, 57-69.

Tör, M., Lotze, M.T. & Holton, N. (2009) Receptor-mediated signalling in plants: molecular patterns and programmes. *Journal of Experimental Botany*, 16(13), 3645-3654.

Tsuda, K. & Katagiri, F. (2010) Comparing signalling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Current Opinion in Plant Biology*, 13, 459-465.

Tyler, B.M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R.H., Aerts, A., Arredondo, F.D., Baxter, L., Bensasson, D. & Beynon, J.L. (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science,* 313, 1261-1266.

United Nations World Food Programme http://www.wfp.org/hunger.

Van der Biezen, E.A., Freddie, C.T., Kahn, K., Parker, J.E. & Jones, J.D.G. (2002) *Arabidopsis RPP4* is a member of the *RPP5* multigene family of TIR-*NB-LRR* genes and confers downy mildew resistance through multiple signalling components. *The Plant Journal*, 29 (4), 439-451.

Van der Hoorn, R.A.L. & Kamoun, S. (2008) From guard to decoy: a new model for perception of plant pathogen effectors. *The Plant Cell*, 20, 2009-2017.

Van't Slot, K.A.E. & Knogge, W. (2002) A dual role for microbial pathogen-derived effector proteins in plant disease and resistance. *Critical Reviews in Plant Sciences*, 21, 229–271.

Vauthrin, S., Mikes, V., Milat, M.L., Ponchet, M., Maume, B., Osmana, H. & Blein, J.P. (1999) Elicitins trap and transfer sterols from micelles, liposomes and plant plasma membranes. *Biochimica et Biophysica Acta (BBA) – Biomembranes*, 1419 (2), 335–342.

Voegele, R.T. & Mendgen, K. (2003) Rust haustoria: nutrient uptake and beyond. *New Phytology*, 159, 93–100.

Wan, J., Zhang, X.C., Neece, D., Ramonell, K.M., Clough, S., Kim, S.Y., Stacey, M.G. & Stacey, G. (2008). A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*. *Plant Cell*, 20, 471–481.

Wang, Y.H. (2008) How effective is T-DNA insertional mutagenesis in Arabidopsis? *Journal of Biochemical Technology*, 1 (1), 11-20.

Wawra, S., Belmonte, R., Löbach, L., Saraiva, M., Willems, A. & Van West, P. (2012) Secretion, delivery and function of oomycete effector proteins. *Current Opinion in Microbiology*, 15 (6), 685-691.

Whisson, S.C., Boevink, P.C., Moleleki, L., Avrova, A.O., Morales, J.G., Gilroy, E.M., Armstrong, M.R., Grouffaud, S., Van West, P., Chapman, S., Hein, I., Toth, I.K., Pritchard, L. & Birch, P.R.J. (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, 450, 115–118.

Win, J. & Kamoun, S. (2008) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Signalling & Behaviour*, 3 (4), 251-253.

Wulff, B.B.H., Horvath, D.M. & Ward, E.R. (2011) Improving immunity in crops: new tactics in an old game. *Current Opinion in Plant Biology*, 14, 1-9.

Ye, W., Wang, X., Tao, K., Lu, Y., Dai, T., Dong, S., Dou, D., Gijzen, M. & Wang, Y. (2011) Digital gene expression profiling of the *Phytophthora sojae* transcriptome. *Molecular Plant-Microbe Interactions*, 24, 1530-1539.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T. & Felix, G. (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell*, 125 (4), 749-760.

Zupan. J.R. & Zambryski, P. (1995) Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiology*, 107, 1041-1047.

Appendix 1

SNPs detected on *Hpa* 804480 gene across *Hpa* isolates:

```
                50          60          70          80          90          100         11
1. EMOY2    ATTCTAAGCCTCACGCTCCTCTCAACTGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
2. CALA2    ATTCTGAGCCTCACGCTCCTCTCATCCGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
3. NOKS1    ATTCTGAGCCTCACGCTCCTCTCATCCGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
4. EMCO5    ATTCTAAGCCTCACGCTCCTCTCAACTGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
5. EMWA1    ATTCTAAGCCTCACGCTCCTCTCAACTGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
6. EDCO1    ATTCTGAGCCTCACGCTCCTCTCATCCGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
7. GOCO1    ATTCTGAGCCTCACGCTCCTCTCATCCGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
8. HIKS1    ATTCTGAGCCTCACGCTCCTCTCATCCGCTGAAGATGAGGAGCTGCCAGTGACGCCTGCGGT
9. MAKS9    ATTCTGAGCCTCACGCTCCTCTCATCCGCTGAAGATGAGGAGCTGCCAGTGACGTCTGCGGT

                190         200         210           460         470         480         490
1. EMOY2    AAAACACCGGAAGATTCTCGATC(    GGTCATGGTCATTCCCATTCGGTGGCTGATTTGTACG
2. CALA2    AAAACCCCGGAAAATTCTCGATC(    GGTCATGGTCATTCCCATTCGGTGGCTGATTTGTACG
3. NOKS1    AAAACCCCGGAAAATTCTCGATC(    GGTCATGGTCATTCCCATTCGGTGGCTGATTTGTACG
4. EMCO5    AAAACACCGGAAGATTCTCGATC(    GGTCATGGTCATTCCCATTCGGTGGCTGATTTGTACG
5. EMWA1    AAAACACCGGAAGATTCTCGATC(    GGTCATGGTCATTCCCATTCGGTGGCTGATTTGTACG
6. EDCO1    AAAACCCCGGAAAATTCTCGATC(    GGTCATAATCATTCCCATTCGGTGGCTGATTTGTACG
7. GOCO1    AAAACCCCGGAAAATTCTCGATC(    GGTCATAATCATTCCCATTCGGTGGCTGATTTGTACG
8. HIKS1    AAAACCCCGGGAAATTCTCGATC(    GGTCATAATCATTCCCATTCGGTGGCTGATTTGTACG
9. MAKS9    AAAACCCCGGAAAATTCTCGATC(    GGTCATAATCATTCCCATTCGGTGGCTGATTTGTACG

            )       500         510      )       710         720
1. EMOY2    TGTGGTTGTGGACACTTGCTGGGA    ACGAAACAGTGGTTGTACCAGAGC
2. CALA2    TGTGGTTGTGGACACTTGCTGGGA    ACGAAACAGTGGTTGTACCGGAGC
3. NOKS1    TGTGGTTGTGGACGCTTGCTGGGA    ACGGAACAGTGGTTGTACCGGAGC
4. EMCO5    TGTGGTTGTGGACACTTGCTGGGA    ACGAAACAGTGGTTGTACCGGAGC
5. EMWA1    TGTGGTTGTGGACACTTGCTGGGA    ACGAAACAGTGGTTGTACCAGAGC
6. EDCO1    TGTGGTTGTGGACGCTTGCTGGGA    ACGAAACAGTGGTTGTACCGGAGC
7. GOCO1    TGTGGTTGTGGACGCTTGCTGGGA    ACGAAACAGTGGTTGTACCGGAGC
8. HIKS1    TGTGGTTGTGGACGCTTGCTGGGA    ACGAAACAGTGGTTGTACCGGAGC
9. MAKS9    TGTGGTTGTGGACGCTTGCTGGGA    ACGAAACAGTGGTTGTACCGGAGC

                960         970         980       1,000       1,010
1. EMOY2    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
2. CALA2    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
3. NOKS1    GGAGCCATGGTTGGCACGGTCGT    AGGCGCTAGTGATTCTAGC
4. EMCO5    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
5. EMWA1    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
6. EDCO1    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
7. GOCO1    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
8. HIKS1    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
9. MAKS9    GGAGCCATGATTGGCACGGTCGT    AGGCGCTGGCGATTCTAGC
```

SNPs detected on *Hpa* 814014 gene across *Hpa* isolates:

```
                 30          40          50
                 |           |           |
1. EMOY2    ACAGTGGCGACCCTAGCTGCAACC.
2. CALA2    ACATTGGCGACCCTAGCTGCAACC.
3. NOKS1    ACAGTGGCGACCCTAGCTGCAACC.
4. EMCO5    ACAGTGGCGACCCTAGCTGCAACC.
5. EMWA1    ACATTGGCGACCCTAGCTGCAACC.
6. EDCO1    ACAGTGGCGACCCTAGCTGCAACC.
7. GOCO1    ACAGTGGCGACCCTAGCTGCAACC.
8. HIKS1    ACAGTGGCGACCCTAGCTGCAACC.
9. MAKS9    ACAGTGGCGACCCTAGCTGCAACC.

                120         130         140
                |           |           |
1. EMOY2    AATCAAGTTTGGAGGATTCTTCG
2. CALA2    AATCAAGTTTGGAGGATTCTTCG
3. NOKS1    AATCAAGTTTGGAGGATTCTTCG
4. EMCO5    AATCAAGTTTGGATGATTCTTCG
5. EMWA1    AATCAAGTTTGGAGGATTCTTCG
6. EDCO1    AATCAAGTTTGGATGATTCTTCG
7. GOCO1    AATCAAGTTTGGATGATTCTTCG
8. HIKS1    AATCAAGTTTGGAGGATTCTTCG
9. MAKS9    AATCAAGTTTGGAGGATTCTTCG
```

SNPs detected on *Hpa* 814231 gene across *Hpa* isolates:

```
                1         10          20              270         280
                |         |           |               |           |
1. EMOY2    ATGAGCACCCCTCTCCGTTC        GGCAACACGTGGGG
2. CALA2    ATGAGCACCCCTCTCCGTTC        GGCAACACGTGGGG
3. NOKS1    ATGAGCACCTCCTCTCCGTTC       GGCAACACGTGGGG
4. EMCO5    ATGAGCACCCCTCTCCGTTC        GGCAACACGTGGGG
5. EMWA1    ATGAGCACCCCTCTCCGTTC        GGCAACACGTGGGG
6. EDCO1    ATGAGCACCCCTCTCCGTTC        GGCAACACATGGGG
7. GOCO1    ATGAGCACCCCTCTCCGTTC        GGCAACACGTGGGG
8. HIKS1    ATGAGCACCCCTCTCCGTTC        GGCAACACGTGGGG
9. MAKS9    ATGAGCACCCCTCTCCGTTC        GGCAACACGTGGGG

                340         350             390         400
                |           |               |           |
1. EMOY2    AGCCGCCGACGTCGGAC       AGATGACGGACGCC
2. CALA2    AGCCGCCGACGTCGGAC       AGATGACGGACGCC
3. NOKS1    AGCCACCGACGTCGGAC       AGATGCCGGACGCC
4. EMCO5    AGCCGCCGACGTCGGAC       AGATGACGGACGCC
5. EMWA1    AGCCGCCGACGTCGGAC       AGATGACGGACGCC
6. EDCO1    AGCCACCGACGTCGGAC       AGATGCCGGACGCC
7. GOCO1    AGCCGCCGACGTCGGAC       AGATGACGGACGCC
8. HIKS1    AGCCGCCGACGTCGGAC       AGATGACGGACGCC
9. MAKS9    AGCCGCCGACGTCGGAC       AGATGACGGACGCC

                560         570
                |           |
1. EMOY2    AAATGCCCCACGCCGAGGGT
2. CALA2    AAATGCCCCACGCCGAGGGT
3. NOKS1    AAATGCCTCACGCCGAGGGT
4. EMCO5    AAATGCCCCACGCCGAGGGT
5. EMWA1    AAATGCCCCACGCCGAGGGT
6. EDCO1    AAATGCCTCACGCCGAGGGT
7. GOCO1    AAATGCCCCACGCCGAGGGT
8. HIKS1    AAATGCCCCACGCCGAGGGT
9. MAKS9    AAATGCCCCACGCCGAGGGT
```

SNPs detected on *Hpa* 813915 gene across *Hpa* isolates:

```
            280         290         300         310
1. EMOY2  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
2. CALA2  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
3. NOKS1  TACAGCGAGATTGTGACGCAGAGAGACGCAGTGTCTCCA
4. EMCO5  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
5. EMWA1  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
6. EDCO1  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
7. GOCO1  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
8. HIKS1  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
9. MAKS9  TACAGCAAGATTGTGACGCAGAGAGACGCGGTGTCTCCA
            410         420         430         460         470
1. EMOY2  TCAAGGACAGGAAGCGGCGATTCC. TCTCAAGCTCAGGA
2. CALA2  TCAAGGACAGGAAGAGGCGATTCC. TCTCAAGCTCAGGA
3. NOKS1  TCAAGGAGAGGAAGCGGCGATTCC. TCTCAAGCTCAGGA
4. EMCO5  TCAAGGACAGGAAGCGGCGATTCC. TCTCAAGCTCAGGA
5. EMWA1  TCAAGGACAGGAAGCGGCGATTCC. TCTCAAGCTCAGGA
6. EDCO1  TCAAGGACAGGAAGCGGCGATTCC. TCTAAAGCTCAGGA
7. GOCO1  TCAAGGACAGGAAGCGGCGATTCC. TCTAAAGCTCAGGA
8. HIKS1  TCAAGGACAGGAAGCGGCGATTCC. TCTCAAGCTCAGGA
9. MAKS9  TCAAGGACAGGAAGAGGCGATTCC. TCTCATGCTCAGGA
            490         500         510
1. EMOY2  CGTCGAGACTCGAGATGAGACGACAAACGTCGG
2. CALA2  CGTCGAGACTCGAGATGAGACGACAAACGCCGG
3. NOKS1  CGTCTAGACTCGAGATGAGACGACAAACGTCGG
4. EMCO5  CGTCGAGACTCGAGATGAGACGACAAACGTCGG
5. EMWA1  CGTCGAGACTCGAGATGAGACGACAAACGTCGG
6. EDCO1  CGTCGAGACTCGAGATGAGACGACAAACGTCGG
7. GOCO1  CGTCGAGACTCGAGATGAGACGACAAACGTCGG
8. HIKS1  CGTCGAGACTCGAGATGAGACGACAAACGTCGG
9. MAKS9  CGTCGAGACTCGAGATGAGACGACAAACGTCGG
```

Appendix 2

Amino acid sequences of the Hpa-sourced proteins detected via MALDI-TOF:

800098

MWTSWSAAVSLAVAVTYASLPGAAVGQDAFDARAQAIVDKFNISQVLGQMTQVDINFAMNLDTKTLNETSVRAYARRH
VGSYFNALFAGRQESAPYGWTASEFRALIKRIQEITMEENGGHPVIYGIDSVHGANYVEGAVMFPQQINMGASFNPELVY
KAGQITARDTEAAGIPWIFGPILEISYNPLWARTYETFSEDPYVASVMGDAIIRGLQSYNQTAACMKHFIGYSKTERGHDR
DNTHISDYDLLNYYMPSFKAAVDAGVMTAMETYTSINGKPVVASSRILNDLLRTDLGFKGLLVTDYAEINNLKDFHRVAKT
YDDAVAMAMMQTSIDMSMVPNNANFIDQALKMLEKHPEQEARLRESAKRVVKTKLKLGLYENPIPGEQYVSMVGNDDD
VQTALAMARESIVLLKNAENVLPLPKTASVFLTGHSADNVGYQCGGWSKAWQGYSGNAMFPNGVSVRQGFEGLVGNS
SFTYFNGLQVNGSISDADLATAVKQASEHEYTVAVIGEEPYTEKLGDIDDLALPAGQVRYIEALRATNTKMIVVLFGGRPR
LLGSIPDNSMAIINGMLACELGGKAMAEIIYGDVNPSGRLPLTYPKDPANIAIPYNRLVTTRCAQGPCRMQFDFGTGLSYT
QFNYTGLSLDTTRLTSASATVTATVTVTNVGARAGNETVMLFVIQPYRLISVPEVKQLKIFKKIELQPGQSMDVSFTLTTD
DLSVYDPQVGKGLKRVFEDSDYVVAIKPETNCDVYNDNSKHPLCAEFSIDTSGRPAPVQPSDEAAPVQTNTVPSNAIFG
ASPVNSDVMPPV

809344

MKIFLPISVASILLALYGSTDALNVKMPGINFTARKGPDWAPDSTKCKSAAQIQKDMHALKTATDKVRIYSLVDCNQAEMIL
PAAKNAGLKVHLGIWTTASHNYLLQEKAKLGALIDKGLYDDNVIGLHVGSETMYRKEINAETAISYMNEIRDFLRSRGIATP
VTITDVIDIYNANPQLIDAVDYVSVNQFAFWETAEVHVGAAVSLDRLRSLRILAASKNKKMVVSEVGWSSGGRSPAASIAS
PENQAKFFSDYFQMARSRNLEYYWYVAFDSKWRVTNGDKEVEADFGVFNEDDTMKSNFQQLNIGWREPRAVRNAGT
QLVLSENGGNVYMSGKSGDWLVQEQQVWFFDSATQQVRSKSSDRCLDAYQGWDGGIVHVFGCLDQEANQKWVFDS
ATGQLKHLKHQGFCLDTDPAQGNKVQLYGCSVNNANQKWSVIDPATI

814621

MVHSCLLKTAALVSSTLLASAEATYVSQGGVVTISHLSGSFAKMSTHGATVFSFSPAYDPYGDVLVMSNDSVYDGIKPLS
GGITIEFPNHDATKDYPVPELGFARVSKWTLEDVDLTNDKNSYSSAMFHLDSTDDTLQMWPFKFRLFYEVRLYSTWLET
ELTVMNTGSKGMVFEALLNNHFSVPDVRNNGVEVSGLQGVEYFDRVTGKTQNEARESFGIMSRVDSIYKDVKNDVTATI
RGDGFTETVVVEKSARLHTGYAPPVTPSTNCVVSNPWNQDTKSDFSDEEYIHLVAIGVGTVRDKIGIAPDDSYTLNQIIKA
SRYP

809613

MSTYLALGAAFTTVEYQNMQAQVSTEHVIVAINGAPARDGHQVHGSSFVISLDNGQQWVLYFFSNAGGNTSLLHQANKL
TTTSTFTGVVQAAFVSTSAVQAAAPQDDRKILQLYHASAGVYPKGATVQTLNSESFAFHWQLAAAVGSNPENRFLHFAL
THLQTLLDPSTVEAKHELVLHSHTRGPLFAYALVTRSGCNPSWLCHIPQAESQGVETCTAFYPPGAACVNAGEVKRLNL
CRIVTDEIHGDWALPHEGSYYFKGKALQKFGTMCLVARQLADTTSPELRNVAQHGISKLQQLLSEFADNRSAFPLVYDTI
YKGIISSEALAKNDMNVDFGNGVYNDHHYHYGYIITAAAMALYLDPHWRHSADAAKVQTFVDTLIRDVVTASPIGNDPYF
PRFRNFNWWLGHSYSHGVTPMIDGKDEESTSEDVNFLYGVALYGQVTENAAQEALAKLMLKVCVRAVNTYFLIRNNGP
LIHPAGFAKNKVTGVFFDNKCDYTTWFSPNQECIHGIQMIPVSPILEISRSSSFVREEWNEVLSKLPFVRDWKAHQSGWT
SLLFANYSVLNREAALEVLATCPMDDGLSRAWALYYAVTRPPPSC

812220

MGGAVNVYLIERRGTGRSTRLDCTAAQSTTTGSPSGGKIEPSEVPACALDLQNTYGNLAAFSTTSAAMDVSSFIAKLTNG
ASTFVYGRSYGTVVAQRLMQLNPTTVTGYILDSVTTPSLAPNQKFYYSARDAMFGEVAEHFMGLCALDSECSAHFKDTS

LSATLHDVLSKFDSNPDSTCATMVAEFTMTTPSVGLREQLGGLLDDLAGRSLIPSIVYRLHRCDANDVEALKVFFQVNAE
YASMPDESEAIDSELLRNLVLFSEFWEDPTPSQANLQARFATASMSTKPVHDVLPLYCAFSKEQSSQCAETGVKSYEAN
GIIYPYDQYSNAIPTISSQASVLLMSGGLDALTPSKYAKFLFEALDTTKKELVVFDYVPHSVMASALYGEDEKKCGLELLIS
FVSSKGDLAHLDKSCMAQMPAFDMKLTAEVANSFFGTDDAYNGVPAPAQQDAGATPA

808621

MLTSWLALGLLSVPSIAIANGIDEYDARAQAIVDSFSAAQVIGQMTQLTLSEVMNGTSRELNESAVREYAQLNVGSYLNTY
FEGPVNGVYGYNASEFRSLIERIQDITMEVNGGHPIIYGIDSVHGASYVVGPVLFPHQINSGASFNPDLVYEMARITARDT
EAAGIPWIFGPILDLSQNKMFARTHETYGEDPYLASVLGVAHVQGLQSYNQTAACIKHFIAYSKTPTGQDRDDVLVSDFE
LLNNHMPPFKAAIDAGALTVMENYISLNGDPVISSTKILNDLLRHDMGFDGVLLTDWNEIYNLYDFHRTASSREEAVAVSL
KQTSVDISMVAQDTDFIGYALNMLKENPEQETRLRESVKRVIKLKLKLGLYETPVPGEQYVSMVGNDKDREVALELARES
IVLLSNRDNVLPLPKNASVFITGPSSDNIGYQCGGWTYAWQGYSGNEMFPNGISIRKGVENLVGNSSVTYFNGLGIDGNI
NTTNLAQTVRLARQHEYTIAAIGEPNYTEKAGDINDPALPEGMEDFIEALAANGTKVIVVLLGGRPRLLGSIPSNAYAIINGF
LPCEVGGQAIAEILFGEVNPSGKLPITYPMHPASVAVPYNHLVTTRCRWDNCQMQWEFGTGLSYTQFNYSSVVIDKKVI
GTSVETLTASVTVTNTGTRAGKETIMLFLIQPFRRISVPNMKMLKKFKKIDLQPGESMEVSFTLSSADWGVYKPVIGSGLQ
WSVEDTEYVVAVKPDTRCDVYTNFTIPMQLSNPLCASFSVNLSGRNITSRNPSSQTPSSGIAMGAPVGPNSPLLTPVS

807735

MVLSSLFKTAALVSSTLLASSFPASLGATYAAPQIVHVTHMSGSSAEISTLGVQVISFYTSHNPDVNVLFMNNDQSAFDG
KNPIQGGATAVFPNYGPPKGHSAPTNGFARTMVWNFASMVPAPDATKPTYAVFDLTSSDATMQLWPFAFRLMYEVHL
WPSSLEITFTVINTHTTAISFDLLLQNYFAVRTIRDNNVMISGLKGYEFLDRITNKVQIDERDSFGIPIQVDIIHKNVNTDVVA
HVKGEQGNRNFIVQAAATTGDGAHTPAVPSTDRTVWNRWNTSTANLADFGDDKYLRTIAIGSGRVSEMEVLYPSLRYAL
NQVITVETV

808748

MVHFTSLSALPTKRVNRVLVVGSLETLQLQGAPSSFLHQTLLRASSSAVESSTEANALLLQYALNSLRPSTDSGATAALLL
LPTSPSTESPLHVTLHALPTQISRSNSQARPHAITSFLKRHTQLISTRDAEEKKAPEERDDDDDDVVLVVLVLPRHADTWL
AAGAAVARAASLYEHKSQRTTALPGLETRTIQSDRLHVVYETPLTRNEETLVQHTADAIQLAARLVDAPPNELHSDAFIAE
AQAVAARTSSRTTIVQGEELRKQGFGGIYGVGKAATHAPALVCLSHVPDGADEVPSVAMVGKGIIFDTGGLSIKTGGNM
VGMKRDMGGAAVLLAAFEAACHARNTTDKTPLHVVLCVAENAVGPDATRVDDVLVMYSGKTVEVNNTDAEGRLVLAD
GVAYAVKHLKPKVIVDMATLTGAQGIATGTRIGAIYTNDEELEKMAVKAGKVSGDLVHPMPYAPEFHRLEFKSTIADMKN
SVKTRTNGQVSCAGQFIANHLGEFERTGKWLHVDMAYPAFTSDDERATGFGVAFMQSLLKEIDNAGW

803598

MADPVEASIEVIDNPNGTLNCLDGPRTSIYAASAHYRGQSSLHFTKHQIAVDLHEDGALLGFPADRTFVLNGPTIDSSLLR
NHLAHWMYRCTDRYSPRSRHVVVFVRDRLSADDRSPRYRGIYVALEKIAYTRNRVALTQLNSACQSNDELSGGWAWQ
NNPLGYGDYSPNMVLNEATGLFGSSGARPVLTYPEPYVLTQSMRDYFVSPETGPLPRLYQYLYANMTDPDGLAEHIDIGS
FVDYFLHSELSENSDAYRRSTYFFKDRSQPINAGPVWDFNLAYGRGGRQTTWLYAGHTFWKRLVCNYKFASLLQQRW
VTLRTTTWSDESIQLFIQSSAEPIRRQLKNCNEWKSENLQCANVGANTKGTFDDAVTDLVNAVLARAQWMDSNMAGFY
KALDRTLCVAAGELPAYNCAADGNDKGCLSNPDAYSNAVEFPEPRKSATAKVCDISKLSSAGGVKMEEPTLDQCWLSA
GVYITDSSITPFCGGYGFCPEGPGAKCTCTSGRRPPTCARHDDVIVPHGGLFGSGVVKPESLLSNSADVVVGESTYLTS
PLFFLCVVSAMVLGAVLTVRQRRRNHHYSRLREVIATSSQNDVDGHSLFYGTS

804884

MVTLHKRLGRAVNVYTMDLRGTGRSSRLECVSSQATTSGSPSGTDIDPTEVAACARGLQNKYGDLASFSVTSAAMDVS
TFILKHTNGASNIVYGSSHIATLMLQRLMQLKPSGVTGYVLDSAAASVVPSGKKTYWSKWDVDFNEVSEHVLNLCSEDT
DCSSHFKTATLSATLHDVMSKFDKDPNSTCATLVRNFTEDEPSSGLRSTLGGLVRGPPMRSLIPPVVYRLNRCEPLDVE
TLTRFFTIITQDAGTAGDTLISNVAFQMTIFSEMWPKPSPSYAELEAQHKNLSISFLGVYESLPLYCAYTKDQSPVCEEYTF

GHYEANPIAYSKDHFWSEAAVVSSEASVLLLSGKLDVQAHHKYSEYIYEALETSKKELVVFDFSSHKTLEDTAFGESETT
CAMELLASFVSCNGDLACLDKSCMAEMPAFDMTVPADIAENCFGSMDAYDGALIENAGIAGGTPA

806582

MGALLAAATAIQSHVVESTGTTSLNGDGPHHHHAHDDGAADAPPLSDGPQDKEGAPDYKWYWHHGEDSIFAQDSGVF
TNFKPNDTDQTCSQDTHATPFNKQVRGVNLGGWLVLEPWITPTLFYQFLNTQQKYGDKAPEKTAMDMYTFCTALEKEE
ANRQLRIHFDKWVTEDDIAALAEAGVNSLRVPVGDWMFNPYEPFAGCTTGAVEALDRVADLALKYDMDILLDIHGLIGSQ
NGFDNSGKSTAVKWTSIASTQPVGTTTFEHWPIRQAEWAGTFDVENHNYTSINYANLNHSIVTVEAIINRYKGHEAIMGL
EPVNEPWELTPIKVLKEYYWKSYKRVKVLAPSWKFVIHDSFRFGLPLWADFLRGCPDIAMDTHIYQAWNPPGTMADFNS
NACQQKYVITKMENALMPVIVGEWSVGTDNCAMWLNGFNDNLPGFPKVQCHMVDCPVNSTYLGEGFPGTPLDTTKPIQ
GPYGTGMSGPSFGKCPVTSQSSFSQKDDAALTRSLTLKKLNGFAVGHGWYFWNFKTEFATKWSFLDLVRNGAFPKNV
SHYHESDGVQNVCVKEDEGDFVCRAKRGVQEYELKSGLAFACNFPDIDCSDIEQRFVTLEEQCDWAFDQYWHFNREK
GATCDFGGAGHLLNVPGAAPSKHPSSPTSQKLKAKQPDDSFAPVSEWGIAAILGVAGLALVVVGAAVFAVVRWRQQRA
RREYSPIAGRS

811928

MNLSLVLAAQKAATKRAASRLSHRVHFSTGKDIRFGVEGRAAMLKGADQLANAVQVTLGPKGRNVVIDQSYGAPKITKD
GVTVARSIEFADKFENMGAQLVRSVASSTNDAAGDGTTSATVLTRAIFSEGCKSVAAGMNPTDLRRGIQMAVDHVVDGL
QQLSMDVQDKEKVAQVATISANSETEIGSLISDAMERVGKEGVITVQDGKTLLNELEVVEGMKFDRGFISPYFVTDNKTQ
SCEMENPYILLVEKKISSLQAIIPMLETVVKQQRPLLIIAEDVESEALAALVINKIRGGVKVCAVKAPGFGDNRKASLQDMAV
LTGATVISEDLGHRLETATPEMLGSAKKVTVTKDDTLMLDGAGSTEAVEERSNLLRASIESTTSDYEKEKLQERLAKLSG
GVAVIKVGGASEVEVGEKKDRVVDALNATRAAVAEGIVPGGGAALLWASRSLSKLYDSCANLDQKVGVEIVERACRAPA
TQIAKNAGHEGAVVVGKMLENDSPELGFNAQTGEYVNLVDAGIIDPTKVVRTGLVDASSVASLMMTAEALIADLPAENGA
APPMGGMGGMGGMGGMMRRGPYLSFLMRTH

805873

MHFASLLAASLVVFTFIGGYSSIHVAAASIVRVMSFNVRTSNANDPCPSGCWRERKERFGMLLNLHPADLIGTQEGTPEQ
IEHFQTVLGFSYVGLCAGECNANERNAIMYKPDRLQLLWSKTFALSDTPDVLPSNSWGHEYLRAAVVARFLDPQSQHVI
CMLNTHFDVSIGHDESAITVARLLYSNCQPQDDVFVTGDLNTVPETAAVQYLLGNVQLNGTFTQLPLYETLTAVGQGGFT
FIGPAFDNNVNSAKIDYIFARKEKKTCLRSGKILTDLFGSYSVSDHAVLLSEFCLGDGCTDCVQ

808763

MASARRAATYGAKVLVIERGRQYDGMGLGGTCVNFGCVPKKIMFNTAAHVEHLNRAKDYSIHTAEFAFGDFDWSKMKT
KRDAYINRLTGIYQRNLGNARVDHVEGIAKFVAKNKVQVADRVFTGKHVLIAPGGVPLKPDLPGINHVIDSDDFFRLEQQP
QKVAVVGAGYIAVEMAGIFHALKSDTTLFCRHDQVLRRFEPIVRDLVNEEMEKAGVTFVRNMSAVGVEKQLDGKLTLVV
NVHNEEQKFTDFDTILFAVGRKPRTEDLGLEDIGVKLAEGGFIQVDAQENTSVPGVYAIGDATVTGWELTPVAIAAGRRL
ADRLFGGEKDACLNYHQIPSVIFSHPPIGTIGYTEPEAVAKYGEENVKVYTSKFVNLLHSMADPEHKGKTAMKLVTVGEQ
ETVVGVHVAGEGADEMIQGFGVAVKMNATKADFDNIVAIHPTASEELVTMAPWGMIKDKIVLPPKPARSGPTVGQ

807738

MVHLFLFKTAALVSSTLLASVRATDVGQGGVITLTHLSGSNLKVDTYGANIFSFCPSYDPLGDVLVVSRETVHDGIKPISG
GITVVFPNLEASKDHPIPELGFARVSRWTVEYVKLANDENTYSSALLRLESSDATHQMWPFKFLLMYTVRLYSTSLETEL
TVMTTGSGGMDFQALLNNHFYVNDVRENGVVISGLQNAEYLDRVTGMTQNDTRESFGIMSRVDSIYNDVKNDVIATIKG
NGFNKEVVIQKTARYNDGRISTFAAKTDCVVSNPWNEDTKSDFSNEEYIHMVAIGAGGVSQNLGVFLGVSYTVNQVITVS
KST

803710

MDRDSLVFLAKLAEQAERYDEMVDHMKAVANNHNVELTVEERNLLSVAYKNVIGSRRASWRVISSIENKGDSERSEHIK
AYRQKIEGELVDICNDILTIIENNLIPNSSSEEGKVFYYKMKGDYHRYLAEFQVDDERKESSDKALESYKQASTIAMAELPP
THPIRLGLALNFSVFYYEILNSPDRACNLAKQAFDDAIAELDTLSEESYKDSTLIMQLLRDNLTLWTSDQEPDTADANQGD
MNVQDVE

801564

MAQGSKVWSRARATIQRLLSADEPILRDNDALRQRALIPMDDVQMHLPAKVGDYTDFYSSREHATNVGRMFRGEDNAL
QPNWLHLPVGYHGRASSIVTSGTDVRRPSGQLQVDVKDPSKGSKHGPCRVLDYELEMASIGHEQRFVYQDHIFGVVLM
NDWSARDIQKWEYIPLGPFGSKNFATTISPWVVPLAALEPFRCEPSFGPAQKDPVPLAYISDPNYARGTYDINLKVSVKP
EDSEQAYTVTKSNFRNMYWNMKQQLVHHTVTGCNMQPGDLLGSGTISGSTDESMGSMLELSWQGTREVALGDSGLT
RKFLQDGDTVAITGFCQGHGYRIGFGSCEGKILPAHPCNT

811423

MTRYAGVEVGGTTWVAAIAEDHPENILEKFEVDTTTPDETMGAIIDWLKERTFDAIGIASFGPVDLNRKSPTYGYITSTPK
PNWGNTEVVGVFERAFPNIPIGFDTDVNAPALYEVAYGGHGDISSAVYITVGTGVGVGVCTNGNAIHGFMHPEGGHIIVP
PAPQDIETDFKGVCPFHGNCVEGMVASGSIAARTGVDRRDLASITDDDPVWDTIAHYLANLCLNVTFLTSPDVIVIGGGIA
RREKLFHLIREKFVARVNQYGQQPPVDKYIRASFHSAIGLVSSLHLARLELEQHK

810555

MVVFVSRTLSLVTAAVALSSAILASRAAPSKAGVPNAYLTQPVDPAALLHAFESCMDSSAADEHKGKTNESVTTNVNFFV
EESPLYEDFATGPKARIARKPLGVYLAKSEVDVVRAVKCSVSNNLAPVPRSGGHSYEVLSSMDGSLVIDMADMVDITLVS
ENKDERSALLTVQTGARLGWIHTELDRLGGYTFNAGTCPSVGIGGHISGGGYGMVGRHYGLAADQTTEMRVVLYDGSV
VTASPTENPDLFWAQRGGGAGSFGIVTLFTIKAYALPQVSVFNVHFNASVRAQILRSWMDFFPTADSKLTTQLVVDGDV
ARITGQYLGPMTELDAILDASGLFEHGGLIVQERLNNCSQLAIKAYAWKATCDDLSSLNVTHHMTIADKDYSKVKGGYSN
SVLDDEGVQTVVDWADRLPNTTWAYIQFEAYGGIFATQTNDMTPWAHRDAVWSVQIGVGAKKGESEDSTSYQWIRGIA
GALEKYFDGGNYQNYCDLDLGEDFGQRYWGADNFARLRQIKAHYDPLDVFKSAQSIPLP

806398

MVVTLDCTPNPNTVVSSFVKKTMTKNKTSTTTTSVGTPIPSPGLTTTTTAGPLGPVVLEDFTLLDRLAHFDHERIPERVVH
AKGGGAFGYFEVTHADITKYCSAKLFSNVGKRTPVALRFSTVGGEQGSADTVRDPRGFAIKFYTEEGNWDLVGNNTPIF
FLRDPMLFPSFIHTQKRLPHSHLKDPDMMWDFFSLRPETLHQQSFLFTDRGIPDGFRFMNGYGSHTFCNVNAKGETSY
VKYHFKTDQGIRNLSVDKAAELAGSNPDYAVQDLYEAIAKKQYPSWTLYIQVMSLDEVAKATFNPFDVTKTWPHRTYPLI
EVGRLALNRNPRNYFAEIEQLAFSPSFMVPGIEPSPDKMLQGRLFSYPDTQRHRLGTNYLQIPVNRPLKVPQTYQRDGA
MVVNGNMYEAPNYYPNSKGGPPEDTSQRHRAYRGDFSVVYKYSTYDDDNYSQVGEFYRKTLDAAGREHLTDNIATSL
VNASKPVQARAIANFTKCDEDYGRRLQEKVDALAAQKKHEDPTSLPAPSQLNPPRKPFEVAPPSDVMTPRL

808887

MSVQAQPGDKVVLAYSGGLDTSIILKWLTNKGFEVICYCANVGQQGEDYEKVRAKALSLGAKKVYIEDLREDFVQHYIFE
AVKANAIYESRYLLGTSLARPVIAKKQVEIAKAEHAKYVAHGATGKGNDQVRFELCAQALDAQLKTIAPWRDAEFIETFKG
RADLIQYAKEQEIPIDATPKAPYSVDENLYHTSYEAGMLEDPMTAPMVEMFKMTVDPKDAPDVPEQISIRFEKGIPVGVTN
LTAKTPELTGALELFLELNKLAGKHGVGRIDIVENRFVGIKSRGCYETPGGTILRHAHLDLEGLCMDREVMKVRDGLSAKF
AEFCYNGFWFAPEMDFVRHAVDYSQQNVTGYVNLELYKGNVTVIGRYSDEALYSADLASMDQEGGGDNFDYNPADAQ
GFIRINATRLKAWRCRPQPKK

814664

MVSSTFIFGISATALLSSLQIIHAETVTIEAKPGTDLFNQFRPVYHFQAREKWMNDPCAPYYDEATGLYHLFYQLNPNSTI
WGNMTWGHAVSKDQVTWKDYPDALRPFEDKWDHLGVFSGFAMNNAIDGKQTVFYTGVTALPIAWRKEYLFGEHVVYA

TTDDGGKTWQKGRKPLIERPPPGLDVTGWRDPMPFHSKSLDRSFGYTSGKNGSSNSSSSSNYLIVAGGLRDVGPRIFL
YHAEDYVNWEYKGYLLAPEKNTTFSPYSAGWGYNFETTIYREMVDDDNELHNVMLFAAEGETNRYAMWATGSFASSR
GYGSDFETDSEAGLFTPLMVGVSDNSDWYANSIYTDKEGKNVLIGWITEDNNFATGQPQGWDGILSLPREVGISIVRDIY
DRDRHLVGKGDWIVSDSKDVPCFHGSPKQSKTIKTLGVKPLRDLKRLRNVKAVEQVASVSMNGSTQVLTSTGASFELM
AEVAELTRGSKVGFEVRRSSSGDEVTTIVYDDAKKKVVIDRSKSSSADCAVFADNHVKPVSEPVWGHFYLYDLFTGTAK
NDVCEVAREKLSFHVFVDVSSVEVFLNGRFALSARIYPCASQTESDGIALIASRDATFENVQVWTDPRHAWVDTRVVPV
F

806254

MGDGDGEPGIYPSVPLDTRCSTAAIPCGKQDTSATDSCHGNTASRTPPQSILATGTQPQSMFEFRRCCLKQCPDAAVL
CSSAAEIPCGDRDESVSKVCREGKSIDDFHSCCVKNCPVLNPLGNTCSKAGYACGSSDESASLYCNKMDKTLSGFQQC
CANACSFNPTNLVCSKTRYKCDGGDEAVSGLCRDQKPFDKFESCCYRRCGSKEETSSPQFK

806861

MASSSSPSTRPLTLPEKILTHWAVGLSKPEVSPGQMLCVKAQWTLACEITWKSMDKTYQDMGRPRIWRNDRFWLAVD
HTVDPRVNHKPQQQMMIKASEDFAKEAELTNYQPPNTTILHTEFYRQRAQPGQIVVGADSHSCSTGGLGAFAIGLGAAD
VVMPLVTGETWIKVPETVLIEFVGTPPMGMGGKDVMLYTLGQLKCNTVAIGRCVEWGGNIAALSCDARFAISNMTAEFG
GIAGVFPADESTAAYISKRPDHNQDALYFRADKDAHYVEKFQINLDNLKPQVALFPSPDNVKPVSEVVGMKLDGCFIGAC
TTAEEDLVLGALVLEVCLKQGMKPVAQGQRRVTPGSQIIVDNLKKHGLIEVYEKAGFTVGAPGCSFCLGIAADVAGENEV
WLSSQNRNYRNRMGKGSIANLASAITVAASSFTMEVTDPTPFLKLIDNEKFDSAVPKRPAVPITIAEVKPELPVATSEAGD
NADAAAAAALETFSGSIKGRVQVFGDNIDTDAILPGEFLCENDMETLGKVAFLYTNPDFRDKVKQGQDVIVGGHGFGCG
SSREQAVTAIKGAGVKAVIARSFGYIFSRNYQNFSLLGIQLDDDRFYELAKENADITVNMTGRTIEVGGETFRFNMSLFEE
RLLAGGGIMPLYKKFGNRLFRVAVQDPAADKGSCGSGGGCSTKEETPASAGKDISW

803173

MGLLDIVQPGVLTGEDVMKVYKYAQVHKFAIPAVNVTSSSTANAALQAARDIKSPIIIQTSNGGAAFYAGKGIDNANQKAS
VLGAIAAAYHVRAMAKHYGIPVILHSDHCAKKLLPWFDGMLEADEEHFAKHGVPLWSSHMLDLSEEPMEENVAVSKQYF
ERMAKMGLILEVELGITGGEEDGVDNSEVDNASLYSQPEDILLAYNELGAVSPYFTIAAAFGNVHGVYKPGNVKLHPEILG
DFQKFVAKEKALSTDKPVFFVFHGGSGSTPQEIETAVGNGVVKMNIDTDTQWAYWDGLRKFYEGKKDYLQCQIGNPDG
PDVPNKKYYDPRVWVRKSEEEMIARLQQAFRDLNCVNVL

805299

MVCTSYKTVYAMLAALLVGSASAGQSVCYAPWHHPTVTKELLLKDILEIGIYFSSFRTFETRMNNLNVIQVAGEAGVKVA
VGVQMNDLAAVDAEIEAVCTGFGTNPEALEAVYVGNENLQNNGFGTVSADQLIGYIKRVKECTHGAVPVGTVQRINEWL
SAEGAMKVADASDLIGINVYPFFTLGDSSPVEKLQAQWEQMTGKYDAAKCRLTETGWPTKGEDSVGNHASLETTQLYF
DAFLTWAQDKGPSYWFMMYDTTKSYTGMEYEMNFGLFTATGTMKVKIPGSMGSKTGIPMNGSEIIPDPAIVPGTNSSNN
DHPENPVQETPVQGTPVQETPVQGTPGAPPATHSEPVYALIEPLYFDSNGKPVTGGAKQGPYEVNLPADATTGHPPVV
ESGNGEIPPVTEAGNGASTQGIPEENLPTDGTPVVQPGNNEVPLETSVGVPMYRDSNGKPVDGGVKQGIPEKDEPVTT
LGMPPVVDSGKDCAM

805450

MICQRSPLFHVLMVWSIATGLALGRRYMRPSPDPATSKLMSVGTYSRFVDNSRKEEIQQLERIYCRGDLLHEVQTQELF
FDSKTFVDMPIKKTSSVDEVLAKFQDLKASFGDNRLEGKAWKTSLAGFVDQHFDPPGAELVPIVPPDYQDGEMPLGIEQI
RNESLQDWAMELHKLWTVLARVPVSVSAAQKPKSSFLYALPIAPASDESVDAFARQFNGENVLVVPGGRFRESYYWDT
YWIVQGLLISGLHQTARGVVNHLLEYVAEFGFVPNGGRIYYLTRSQPPMLSDMVRLVAKLDGNNASDEDASVWDLHYLR
AALPLLEREYEFWMRQGDHGHAVEIEISGSTANETFVLNRYDASAGVPRPESYREDVSTASASVTDHPNTVDRSSVYNE
IIAAAESGWDFTSRWFGDYSTLKTVRTSRVIPVDLNSIMHRVELNLAMFHEILGDPVRSARFTEAANNRVRAMDAVLWSE
SDGCWKDYLLDSREHSKVVSASDYSPLWGGAFDASDSSRLDKIVASLETSGLLQVCGVQTTTFTSGEQWDAPNAWPP

LQDIIIEGLLVAGTARAQALAKHLVKTWVMASFVAWQNTGLMFEKYDAQQLGGVGDGGEYTPQFGFGWSNGAMLTFLT
KYQDHLDNSDGIVCGNGSTD

806256

MHVPSSLVLFMAAVASGSATEYAGGVSGSPSISPNTLCSAINVPCDGFDQDASSFCSWQKDPNVEFKDCCLGRCSDAR
IPRGYKCSDSGVPCDQLNVDASQSCGVYETRVGQFRDCCIRECRRQDQH

800014

MSTFPSIKLTYFSAAGRAEAARLIFYIGGIPFEDKRITREQFMAMKDTFPLGQVPTLEINGDVMTQSYAMMRYAGRLSGLY
PSTRPVLALKIDEVLSVMGEIEARMVPSFYESDVAKKKAMREELTTVTLPRYAGLLEACLQKMKQTPFFQSDRIYVHEVAI
YIWVKSLRSGVIDFIPTTVLSSYQLLNEMHDKVANHPKVQEWYRLEHNAPKLKLTYFPAQGRGEPIRLAFYIGDVEFEDER
VPLEEMLARKPSLPFNQFPVLEVNGEVMSQALAILRYAGTMSGLYPVTDPIAAYRIDELLSIMDDMFNNPLWGASHREKN
MEKKEELRAELSTVVPKTLGFLEKRIAKNKGPYAAGPNLSVVDLAIYSALMDFTMNLPKLMSTVQDSYPNVYRVYDQVLK
HPKVVEWNAAHNQ

800104

MNPGADFFRSAEFGALNPNRFVPVIKDEDFVLTEGMAILQYLGDRFNWTGPGDLYPNDIRVRAKINEYMHWHHTNTRLF
TINIFRPATAQRNNSATGKDLEALAGTDDLIARVFGLLEVFLVNDYIAHTNFPTIADFAAYCEIDQLGMMGFQFSQYPKVSA
WIGRMKMIAHYEEVHQKLEVYIDERGLRNFHS

811239

MHLINIVDAWKTGDRDRKLSRGEAKELMPMYDYEPSLSLDLGDRVSPSTNDGSGGFYPGTPESGASNRELDDRSVVVV
SADGVVEIYGDLEAEHDDVLTLGTRAEEEEELSPEARALLPAPNSCPRSVPYQIRLDDYLCACKHCPRALTIQRAFQLYG
FALDHPHSDALLRVLRACANYDETIVEVDPILVLHAEKMLEELNFNENETFCLLCRTYCIQDTHNAGLESITALSSDSSVDN
TSEEMSGSIHVILERSPPTQSNNNCSKMTRERTYIMIKPDGVQRHLVGEIIKRFETKGYKLVALKLARPSVEHLEAHYADL
SGRPFFPALIKYMSSGPVTCMVWEGTNVVLEGRKMLGATKPTDSALGTIRGDFCVDVGRNVCHGSDSVDSAEKEIALW
FPEGLVDWSAFDDEWVYEK

802945

MVLKIRHDAPGFTAEAVVDGEFKTVSLSDYKGKYVVLFFYPLDFTFVCPTEIIAFSEKAAEFRKLGCEVLGCSVDSKFSHL
AWINTPRKQGGLGELHIPLIADFNKEIATAYDVLIDAGDDIGATFRGLFIIDGEGKLRQSTVNDCPVGRNVDEILRLVEAFQ
YTDEHGEVCPANWKKGSKTIKPSVAESKEYFEAAH

813525

MPTATADYKLFKPLRLGHDLELKNRVVFGPLTRGRANADRVPSADNELYYEQRAGAGLIITEATAISEQGYGWYHAAAC
YNEAHVAGWKRVTDRVHKKKGTIFLQMWHMGRQSHSSFQSTREIVSASALRLESGHTRDAKYVQAEYETPRALETHEL
SGVVESYRHGAELAMKAGFDGVEIHGANGYLIDQFLQSCSNKRTDKYGGSFENRARLLLEIVEAVKTAVPAHRIGVRLAP
NGAFGGMGSEDNYEMFKYTMEQLGTHGLGYLAILDGPGFGYTDKCRHTTPFDAKTAFKGVVMANNNYSRDLAEGALR
TGAADCVGFGRLYISNPDLAERFQNDWPVAPEARREVYYDSSLGAKGYNDFPFYQASTDKSNK

802651

MTGRRLAKKLTYPTVRRADIVEHLHGVPVSDPYRWLEDPDAAETRTFITQQNDMTQELLRHTVPYVAETKARLTDLFNY
EKYSAPRKYGTAKYVFSKNDGLQNQSVLYIQDTLQGQPRVLLDPNQLAADGTAALSSRAFSKSKLNDGKLFFAYGISRG
GSDWQTIHVMDVQDEGTKMLEDTVEWVKFSSISWTHDDKGFFYSRYPPLEKDEKDSIQLEKGTETDANLNHQIWFHAL
HTSQTQDKLVYAYPSKPTFNVGAEVSDDGNKLLLYVQDGCKNANMVHVADLSDFDSFLKGPRDSTIAVTKLVDEMDAA
YSYVLNDGDYYYFKTNADAPRERVVRTKVVEGCAAPVWEEVIPEQPDAIVIEDVHPVAPNLLVVQIVKDVHNELHVYNLD
GVYQYKIPLPSVGTVGVASKRTESELFYHFVSFLYPGSIFRIDLSKAEAGSTTAPETEVFRETKVKGFDPSQFETEQVFYP
SRDGTKIPMFLVKRKNAPKNGQLPVYLYGYGGFNISLTPAFSVSRLVFVQHFNGMLALPNLRGGGEYGEQWHQDGMLH

KKQNVFDDFHGAAEYLIAEGYTNPEKIAIHGGSNGGLLVAATSNQRPDLYRCVVGAVGVMDMLRFHKFTIGHAWRTDYG
DPEVADDFHYIRQYSPLHNVPRADSPIMQKLAERGGFPSVLLTTGDHDDRVVPLHSYKLIAELQHQLGGSEAQTNPLLIRI
DTKSGHGAGKPTAKIIEETSEVFAFIAWNLGATFVA

808043

MVYTIVVHLYAKEGKEVEEKLRNKLIEASQVYSKDTETVGWHVMQDHEDPRKWCIVERYEHESSQKYHLGNPYWKTFD
PYVMPLLDKPMDLRRFNELDTSKPVHVE

803591

MDFGSAAFTSAEFLALNPNGQIPVLKDGDFSIFEGSAILVYLAEKLGWTDLYPTDVQAHAKVNQYLHWHHTHTREITSKV
LLPLMHTKMNIGTPEEAVMIKDSTAMITKLADTTEKLLVKDFVAETDHPTIADVAAYCEFVQVELMGIFDFSKHPKLSAWL
GRMKEVPHHDEMHADLDGFCTSLGLKTKASA

808107

MDYTKNIRNMSVIAHVDHGKSTLTDSLVSKAGIISAKHAGEARFTDTRADEQERCITIKSTGISMFFEYDMDAGEQATADA
IAKETAEEKAANEESVQISKNSYLINLIDSPGHVDFSSEVTAALRVTDGALVVVDCIEGVCVQTETVLRQSISERVKPVLMV
NKVDRALLELHLEPEDCYQSFTRAIETVNVVIATYFDEKLGDVQVYPEKGTVAFGSGLHQWGFTLKKFARLYSKKFGIGE
DKMMQKLWGDWYFDAANKKWTSKNNAEGTLKRAFCQFIMDPIIKMFDAIMNDKKAKYEKMMKAVGVELKSDEKELTGK
PLLKRVMQRWLPAADAVLEMIVVHLPSPVTAQRYRVDTLYEGPQDDECAEAIRKCDVNGPLVMYVSKMVPTSDKGRFY
AFGRVFAGKIATGQKVRMLGPNYVPGKKTDLWVKNIQRTVIMMGRYVEQTPDIPAGNTCALVGVDQYLLKSGTITTSET
GHTIRTMKFSVSPVVRVAVEPKTAADLPKLVEGMKRLSKSDPMVLCYTEESGEHIIAGAGELHLEICLKDLQEEFMGTEV
KISEPVVSYRETITGNSSKTCLSKSPNKHNRLFCEATPLSEELTLEIEDGKDEVSPRFDQKLRARYLADNHEWDVTDARKI
WGYGPDGTGANLFVDVTKGVSYLNEIKESVLGGFNWATKDGVLCEEVVRGMRVNLLDVVLHADAIHRGMGQILPTTRR
VVYACQLVSEPALMEPVFLADIQVPQDAVGGVYGVLTRRRGHVFAEEQRPGTPMMQLKAYLPVNESFGFTADLRQATG
GKAFPQCVFDHYQVVGGDPTDTNNMAGKLVNGVRVRKGLTPEVPPFDRFYDRL

810945

MVATMSDTRRLIFALCSFIAAQSNIAAAANVKGGVAIPQSSTADVGACLKKAGIENSVPTTSTWTVDTEAWNSRVTPKPS
AVAFPKTEEEVSAALKCAAANGVKVTTLGGNRSFSSMGFGRNDGALVMNLKYFKHLEYDESTQLLSYGGPVMISEAANF
MWNKYQRTLPHGRCPDVGMTGVAAAGFGTISRAAGTVLDNIVSVRVALANSTIVNADAKHNAPLFWAVQGAASSMAVV
LDFKIKTIPPPSEFAVNYTISFDKSYKPTQQDNVDALIGTQKWAQSKDNNDLLSIRFGLKKSSTLQGFFYGSSAEAAKVFA
SLKKYLPAAMNVSTTENDFWASEEITTPGLVKQTLTPRRYFYITSVTIPSNAPLTNVTAWELFSNTAFAPKLPDASASGFV
DFWAGEFAKTVLPGATAWKHDDNLLLIRWDMRSSDFNVSFAESSIDMLRKNFYKFVNAYKASGGKPGGFTTYRDEKWT
MAEVAEYLYGGGNYAKLQKIKTEYDPLELFNTDAQAVRALS

801500

MPLTQKHELAPFIQSMTSCPKHGLHNPLKNVERYILYGQVCAKAGAMFREEKVADYFAKLVQRTSPKSYYQARNEVRVT
GKFLENFSGDQVRFLARSFGVPGDHKGQASHGYRWPYGPVALITPFNFPIEIPLLQMMGALFMGNKVLLKVDSKVSIVM
QETLRLLHECGMPMTDVDFVHSDGGVMNELLLKIKPRNTLFTGYMHIRHHCKEFCVRIPANTSAPCQDAYACSGQKCSA
QSALFMHKNWVKAGLEKKLAELASRRKLSDLTIGPVLTVTTKRMLEHADSLLKIPGARVAFGAEELENHTIPERYGAVKPT
AIFVPLKEFVKPENFELVTTEIFGPFQILTEYDDNELPLVLDALERMEAHLTAAVVSNDLEFTHEVLAATVNGTTYSGIRAR
TTGAPQNHWFGPAGDPCAGGIGTPEAIKMVWSCHREIINDVGPVPAHWAIPEAT

804167

MASSDFLNEPFTVNSKNVVYTDDEITSQYTYTTTRVEGTVVTPLEEKYTLKTQRRVPRLGVMIVGLGGNNGSTLVASVLA
NKHQLTWTTKEGVQTPNYFGSVTQASTIRLGTTAQGEGVYIPFHNLLPMVSPNELVLGGWDISSLNLADAMQRAQVLDF
DLQRQLVPYLEKMSPLPSIYYPDFIAANQADRADNILSGSKQDHLDAVRQHIRDFKQSHKLDKVIVLWSANTERFSDIVDG
VNDTSEHLLAAIQANEAEVSPSTIFAVASILEGCSYINGSPQNTFVPGVLDLAEEKKVFVGGDDFKSGQTKIKSVLVDFLVS

AGIKPTSIVSYNHLGNNDGKNLSAPQQFRSKEISKSNVVDDMVASNRLLYKENEHPDHVVVIKYVPYVGDSKRALDEYTS
KIFMNGTNTISMHNTCEDSLLASPLILDLVLICELAERITLKKEGVSDFEHLHSVLSLLSYMLKAPLVPRGTPVVNALFAQRE
CMINLFRACVGLAPENHMMLENKLFSEIDGRQ

801540

MAQTMSDYSVGIDVGTMYSRVGVWQNDSVEIIAKDQGNRTSAVVAMNAHNTVFDGKRLIGRKFSDAVVQADVKHWPF
KLTSGAGDKPQITVQFKGESKTFQSEEISSMVLIKMREVAEAFMGKEVKKAVITVPAYFNDSQRQRRRMQGPLLDSTCS
VSLTSRRLRPSPTVWTRRAASAMCSSSTWATAGDTHLGGEDFDNRLVDHFTQEFKRKHRKDITENQRALRRLRTACER
AKRTLSSSAQAYIEIDSLFDGIDFNSTITRARFEDMCGDYFRKTMEPVEKVLRDAKLSKSQVHEVVLVGGSTRIPKVQQLL
SDFFNGKEPNKSINPDEAVAYGATVQAAILSGNDSSEKLQDLLLLDVTPLSLGLETAGGVMTTLIARNTTVPTKKSQTFST
YADNQPGVLIQVFEGERSMTRDNNLLGKFNLDGIPPMPRGVPQIDVTFDIDANGILNVSAVEKSTGKENKITITNDKGRLS
QADIDRMVSEAEKYKSEDEANKIRIEAKNALENYAYNLRNTLNDEKLKDQIPEDDRKAIDDKVTETINWLDANQSAEKEEY
EAKQKEVESVANPILQKMYAAAGGAPGAGGMPGGMPGGMPGGMPGGMPGGMPGGGAPPAGAADQGPKIEE
VD

808599

MKILVPAIPLALVATSSLTAAQRFCGQYDLKVVPPYTVYNNLWGQADDSNGTQCTEVTRISDESIAWTTDFKWAGSRYQ
VKSFANAALTFTPVKMSQVASMPTVIEYKYESVDDTLITNVAYDMLLSPHPEGEFTYELMVWLATFGGAAPLARSYDPM
VPVKANVTVAGVNFNLYEGMNGNVTVLTYLATGSINRFSGNLQDFVEKLPNPKLLDDQYLVKAETGTEPFQGDAKLIVSK
FSLEIIQKPSNAS

801658

MSTTTITPAPPPAFQRLPSSAVPQKYHINYERIDLLNFRFEGSERVLLHVTESTCVLTCHAVELHVFHVFVDDVASPWRT
QEATAIHSNRLDDTIAFHFAEPLRANASVTLRLQFRGVLNDQLRGFYRTEYEAKGEQRALAVTQFQACDARRAFVCWDE
PALKATFEISMVTDTDVVALSNTHVVETFTRPRANAHLRHETRAKDKVDGTLEKVVWTFAETPVMSTYLVAMVVGEFDVL
SDVTNEGVTVSVYTAPGQSARGSFALDVAVKALSFFTESFKIPYPLQKLDMVAIPDFLGAMENWGLVTYTETYLLVDPKL
SSHEIKADATRAICHELSHQWFGNLVTMTWWTGLWLNEGFAQFMEFDAADHLFPQWKLWETFVQDITLRSAFVRDAMV
SSHPIEVVVNHPDEADEIFDAISYNKGSSVVRMLSEYLGRDVFYRGVHDYLLKYSYKNAVTENLWEALETTSGQKLKSMA
DTWTKQVGFPLVTLTQHDDGKCTLVQQRFFADTSLDKTDNTLWDIPLTLCTSEDPSAIKRLGIWSAKTSSLQRSTTVETP
LVAGDDINSQINVPAGPKGWIKLNPNQTGFYLVNYSPALWKRLEVPVREQILGVPDRVSLLNSVFAFARTGVLSLPDALD
FTSAYVDETASLCWKEISRNMGYYSNLFSDEPFFLMFQQYVRTLFAHVMSRLGWDADPSKQADADEGEFRKIVIHCLGL
ADDKDVINEAKRRFHAYIGGDFTALSGDLRRAVFDIEVTSGDASSAKLLQELHNKSDFAEERDDCLAAMGSVTDASAKLQ
VLEWAVENVRSQGIHHPFISVASDKIGAEVAWQYLQDKWDFLVKKYSATTLGMIVSGAVSCFPTVAMTKEVEAFMVDKD
TSGYKRRLDVALEAVRLKSATFGRDREPLAKWLENREAR

802132

MPPAILRLTVFSTMLLSSANAALPHENSTRLLEAESTPNWLSLSIVVRPDVNDRPNAGYNMLAVPVLSEGGTKVLYDVSA
RFDQGRESCIDKSVISYTQTNGYALVVIPSQKEKPTSCLETESGAYMIPRINFILEAVARLEDGDPVDKTSTCNKQYTITLS
ETDYAVCVTKPDAEMTISGEGMLLTVTYVRDVAEINNGGEATGCWRVSSPFEVKEILGRSLLSPLH

807640

MMSLVRVTKTRTFAKAVRNVHVERKLEELGYTLPAVAEPKGNYRTSVRSGNTIYLAGHLPQPAGGDLIVGKIGKDLTLEQ
GYDAAHYVALSLMATLKNEVGDLDKIKRIVKLVGFVNCVDGFSQQPAVINGASDTIGKVFGDKGIHARSAVGTNALPLNV
AVEIEAVVEVKDKKHQHDSVGLFAMNRGCMASHVYRSCAMHFLLNRFLMLPNAK

809280

MKIFVPSAVALISIMTMIVANAITCSGPHARVEPPAGALVVDAAANPKYKNSFRTLAGAVNKLDLSSGNQQTIFILSGLYKE
KVTIPFLNGPLVLQGATCDATSYAKNQVTIAQATAQKNLPNDVTNDRNALTSTVLFKSNNVKVYNLNIANTAGNVGQAVA

VTVDGENYGFYGCDLRGYQDTLLTKKGKQLYAKSRITGAVDFIFGLEAAVWCERCDIESIGAGCITANGRSSNGSNSYYV
FNHARVYGSKGSLVGKTFLGRPWRPYARVVFQNSELSNVVNAAGWSKWNGQSPDHVHFREFKNVGPGAAKRERAAF
SQQLTQAVSINQILGDNYKTQSWVDLAYL

<mark>813024</mark>

<mark>MHVKTSAVLLLATVASGPATFSSAWDTPGVPPQYTLPKYCGTAVPCDGWDQSASSICTYNKDANVEFKDCCLKNCVHA
TIRPGDKCSDVGVHCDQLNEVASQTCGEQYYRVGDFARCCTLKCQSQWQH</mark>

811332

MSVNIKVHSSHNMKFQEVSIIEDDAVCPNGGDVLMCKSVGYECQEGDDGVQRCLRRDSSFLDKVDNTTATYWGVCSLT
DPSKPSKCLGKFQCIALDVANKNARCYPPDVWRSGRGIAKTCTTSKGKQNDCDKGQYCRTHDDKQECVPMPYPPSGT
SYLSDCTSDGKCDEGLTCEHHPNFSTCTDKED

805669

MASLGDYRVADISLADFGRKEIEMAEIEMPGLMQSRTEFGPQQALQGARVAGSLHMTIQTAVLIETLKALGGDIRWCSCN
IFSTQDHAAAAIARDGSAAVFAWKGETLEEYWECTLNAITWPSDDGKGDGPDIIVDDGGDMTLLIHEGYKAEKAFAANGS
VPDPSSTDNAEMKCVLSIIARLLKTNPTKWTQIAQRCKGVSEETTTGVHRLYQMEKNGTLLFPAINVNDSVTKSKFDNLY
GCRHSLPDGIMRATDVMLAGKRVVICGFGDVGKGSAQAMKAAGAVVYVTEVDPICALQACMEGFQVVRLESIVSKADIFI
TTTGNKDIIMAKDMLKMKNNAIVGNIGHFDNEIDMAGVTTIAKRQNIKPQVDRFIFEDGHAVIILAEGRLLNLGCATGHPSF
VMSNSFTNQTLAQIELWKERDTGKYATGKVYVLPKELDEKVARLHLDNLGAELTVLSDEQAEYIGVKSTGPYKPAAYRY*
806408

MTPPLLKQITPAEVASFAYDAVDAVALEDARRIVHDVRDNVKALMRHAARLGDVPSEDAPLLIRPAELQKAFETLPKDQQ
QVLERTVQRVTRFAQAQRASIRNFEQPIDGGIAGQDVCRAFLDFVLKTAGCYAPGGRYPLPSSVIMTVVTARVAGVKTV
VVSSPRPAQVTLAASYLAGADVFVAAGGAQSIAAMAFGVSGIPVCDIIVGPGNKWVTAAKSLVYGKCAIDMLAGPSECLV
LADETADAAVIAADLLAQAEHDTAAVPILVTTSQLLIDAVNDQLSVQLETLSTAETARASVTTNGFAVLCPDIETCVNVSDV
LAPEHLEILTDEARQVAEKVTNYGGLFIGGRAAEVFGDYGAGPNHVLPTGGTAKYTGGLSVHTFLRIRTWMRIDDAQES
QILVQDSARLARMEGLEGHARAAERRLL

813371

MHEATQAAARAYKTWKEVGVQHRQRVMLKLQHLIRAHTDELALAITQEQGKTLADARGDVFRGLEVVEHTCGAATLMM
GETAENLATSLDTYSYKQPLGVCAGICPFNFPAMIPLWMFPTGTVTGNTYVLKPSEKDPGATMILARLAQEAGLPDGVLN
VIHGAHDTVKFICDAPEIKAISFVGGNQAGEFIHSRGSANGKRVQANLGAKNHAVIMPDCDKEQAIGALAGAAFGAAGQR
CMALSVVVFVGKSKEWVHDIVEKAKEMKVNGGMEPDTDVGPLITVAAKERAERLIQEGVDNGADLLLDGRNVKVPKYP
NGNFVGPTVLNNVRADNPAYVEEIFAPVLVCTSVETLEEAIEIINRNPYGNGTSIFTSSGAAARKFQHEVDVGQVGINVPIP
VPLPMFSFTGSRASIRGDLNFYGKAGINFYTQIKTVTSLWDYNEKTRYSAVMPTLGQK

810356

MKFLVRAAAVAIVVAGRSARATVDDPDVGEGFNYRDKTVEAATQKLVHASIPPSTGSVLHQGPIVQGPMASSPYSTSGG
YFAPFGFDAPSETLYKRNGSLARLAPIINVAPEALNQPIPTNEWWGNLIHATATNTSENFPAWAQPYAFMLPKVPPFGLQ
TYYPYTYREMAPEINGTVRYYEHGIHNDLMLSAREFNESRPTYEVYAWSDMGIKLRLCSRGSINCIDSALSSGMAFVSGT
YKGLTPRIETEYNITFVEESVPGKFIVYLNNNQTWVLYASDVSLSLRVEESVNFSMNASGSSLVADDTYAGTIRVAILPEG
AADTLYDDYVSCIVRGGTVSMESRTGYSLTWDAEGSSCDTVGLLHFALPHQVESMIGSPITAETPGAIMLNSTTRGLMVA
QVSMRWSFVVPEADIAIDYYPDRKRLSLYLIRETDMLRTLQDDIYANWTLATNSWYFNGKALQKYASLCLMAADSLVVGP
DTNGLLAYCIEKLETLVEPLLNNTLSPPLVYETLYGGVVSGLIFDTGQLYMDFGSGIYNDHHYHFGYYVTSAAIIKHLDPTW
SRMPDLEDLIWMLLRDVTNPSLEDPIYPRFRHFSWYHGHSYSRGVTLLDHGKDQESTSEDLNFYYGMTLWGKETGHKA
LEDLGSLMLRLNAYTVREYFLLSSDNKIHPPEIVRNRVTGIFFDNKVYYNTWFLDEKYAIHGIQMIPVSPANALARTPAFVT
EEWNEILSKEPIVTMKNSNNTWLSLLLVNAAVVNKMDSLYKLMNATMDDGLTRSWALYNAASA

808503

MNKKDGYNECTYTTSPDFTLHSNIFHATLIQSRRNQYLRLSLPMYPSIKLLYLFIPQSSQCHTPKNTRSFPLSPPRPPTMT
KIAIIYYSTYGHIAKMAESIKAGVESVAGATADVYQVQETLSDEVLSKMHAPSKQDHPVATPDVLKNADGVLLGIPTRFGS
MPAQVKALFDACGGLWTAGALVGKPAGIFFSTGTQGGGQETTAFTTVTFLTHQGMTFVPLGYRSPLLFNMDEIHGGSP
WGAGNLAGTDGSRQPSQLEKDVAKVQGESFAQVAKKLSSA

814283

MSSTAPTATHDVRHDLILYVNGKRIQVAEKDVRPEQTLLQFLRNDLLLTGTKLGCGEGGCGACTVMVSRFHVSSGRVRH
MSVNSCLSPLCAMDMCAVTTVEGVGAITGPGGDASGLHEVQKALAESHASQCGYCTPGFVMALYAMVKQRETGEELT
MDDIEHNLDGNLCRCTGYRPILDAAKSFGDDASEAQCKGMCLGCPHAVENGNAEVEIEDLYGNLPEVVTSCSSRKIREL
AQQRQHREGEDVGGTKSEQMEGAAQMGWSFPTELVETARTPKVLHIDSDRIKWFAPVTLAHLLEIKSQHPDAKISVGNT
EMGIETKFKGCEYRHLINISRISELVASGDVTSTDRINQTVFTDAETFEGVKFGAAVTLTEVKQHLSKLIGILPPYQTRAFEA
IVQILKWFASTHIRNVACIAGNLVTASPISDMNPLLAAMNAYIELQSTRGVRFARVRDFFISYRKVGMVPDEVITGVYVPYT
KKWEYVLPFKQARRREDDISIVTAGIRVRLECSRRTGDWTIEEASAVYGGMAPITKSASETEKYFVGKLFNASTFDEACD
TLLLSDFELPDGVPGGMAKYRESLCASFLYKFFIASSLRLKHDLQTIMETALVRSEAPSIDDNLQSAGKSFLHQARPVSHG
AQSFGLETGGLQDSKHQPLGDENSTRGPVGDPLVHKSAYVQVSGEAQYTDDIPSTLGTLHGALVLSTCAHGLIRSIDAN
EALVMDGVHRFFDASVFETEKLGSNKIGPVLKDEECFASKEVVCVGQPIGIIVADSHELAMKAADRVKVVYEELPSVMTIQ
DAIREKSFILPVHTIDSGNVEKGMAESDIVLEGEVHMGGQEQFYFETNVSVCTPKEGGMKIISSTQAATKAQVLAASVLGI
DSNRITSTTKRIGGGFGGKETRSVFVTCAAAVAAYVMKRPVKCMLERHVDMLTTGGRHPFYTRYKVGIKRDGTILALDA
DIYSNAGYSMDLSLAVMDRALFHCDNAYSISNLRCNGTLCRTNLPTNTAFRGFGGPQGLFVAETYIDHIARTLTIPPEEVR
SRNMYHEGQSTHFGQVLEDFNLRTLWQHTIDRSRFESKKTDVEAFNQKNRWKKRGISILPSKFGISFTAKFMNQGGSLV
HVYADGSVLVSHGGVEMGQGLHTKVIQVAALAFGIPHEQIHIDDTSTDKVPNSQPSAASMSTDLYGMATLDACEQILARL
APVRERLGPDAPFCDVTNAAYMARINMSAQGFYIVSNDRCGYDFTKSVAENVAIGTAFNYFTTGVACTVVELDVLTGDF
HMLSVDILMDLGASINPAIDIGQIEGAFMQGFGLFALEELVWGDDEHPWVKRGNLFTRGPGAYKIPSANDVPLDFHVWLE
SNRKNKFAVHSSKAVGEPPLFLGSSAFFAVKQAIYAARKETGHHEYFELHSPVTPERARMACADEMLKKVFTARAGDMI
SYRPSGSM

802113

MLLSSSNAAPLPENSARVLKAEPATPGLSILVTLDVNGQSGPSYEMVARPKYFEDKSKVLFDVSARFDQGSPVPGIDKSV
ISYKQTDGYAIVVIPSNGEKPTTCLTTESSTEKETYMIPPLNAMFEKIMGITEDGDGVSTPDCAAGHGISAEGRGYLVCET
RLGQEITIEGNSMLLNVTYHEDVTEINNGGKESGCLQVAKPYEVKGSLGNTLLNLSP

813874

MSFELPKISYAYNALEPFVDTDTMNIHHTKHHQAYVNNINKYIASDKGAALQGKSILEVVQSATEAPVRNNGGGHYNHSL
FWTWMTAPGTTNTAPHGALKTRIDEDFGSLDQLKQEFNAAAASRFGSGWAWLGVKTDGKLAITSTPNQDNPLMPSVD
QPLIPILGLDVWEHAYYLKYQNRRPEYISAFWNVANWDKVVEYYDDFASKGKPVDV

803071

MSDIAAIINVVNKFALAYSLEDADASIKTLLSLFTADATFVEELVGGKCSGKSEIEAALKELAELKFVADTRHLPSGHVVELV
DAENATVSSHTTVFWKCTPVMVVAWTDVLVKSEEGQWLFQQRAADAVQKNLEMIGEMQLRGKKQYAHKDES

800270

MTLFKALLLQRTVSFFTCVALVSSLLANYPTRALIKQNTHTVKTGATLQSYVNGSGYSVTSTSESNGVLTLHLALNPTFTF
TPYGSDLAALVVTVTKTESDSVRVKIVDKNTKRWEVPLSLFTSGTLGINTTKTTAHADPMYSFSYTHNPFTFKVVRKSDG
STLFDSSGIPLVVKDQYLQISTVLSSDVSVYGIGESTRDNLEMVPGDKQTLWARDQPSNTAEVNTYGSHPFFLGLNGAG
QAHGVLLLNSNGMDVTMDSGHLVYQTIGGVLDFTIVVGPTPTNVVSQYTKLIGRPKLMPYWSYGFHQCRWGYESVDEL
RNVVQKYAHNDLPLDVIWADIDYMKDFHDFKPDPINFPHDEMAAFLDVIHSSGQKFVPIIDPGIPDDTDDYTYTRGLSMDI
FIKDTSGKPYLGQVWPGPTVFPDFFHPSATSFWGEQIQLMYKSLAFDGLWIDMNELANFCPGTTCVRQASVTCPNTSSL

TTCCLSCSNDGNKYDNPPFKINNVNSREPIHNKGISTSALHFGNVLQYDAHNLYGMTESIATNSIQEELTNKRSFVLSRST
FPGSGVHVAHWTGDNAATWNDLRWSIITILKFGLFGIPMAGADICGFFGVSSMELCARWTALGSFYPFARNHNNIDAPS
QETYIWPEVTKVGRKFIGMRYQLLPYIYTLGYHAHVNGLPIARPLFMEFPTDVATHEIDHQFMFGDALLVTPVVTQGATSV
TGYFPAGTWFNIFDCSKIFSSGVSFTTKVTLYDMPVHIRGGSILPMHQSALTSAAARLTPFDILVALDSNGSASGDLYLDD
GETITNPNATIVQFSASVGTFISTVAQNNYVGARTSRVDKITVLGVTSSPSRVSLGSISKYDSTTQCLEISLSGVKQTIDSD
YRITWS

813080

MSQASSRVTSLEQAIEALDVRVESLETQITTASSSLSSVETWPAARIRKTFIDFFESHEQLPHTFFKSCPVVPLDDPTLLFI
NAGMNQFKPIFLGQVDPLQPMAKLTRACNSQKCIRAGGKHNDLDDVGKDVYHHTFFEMLGNWSFGNYFKQEAVHLSM
TLLTEVFQVDKEWLYATYFGGDKQLGLEPDEETRLIWLQYLPSERVLPFGCKDNFWEMGDVGPCGPCTEIHYDRIGNR
DASQLVNADVPDVIEIWNNVFIQFNREQDGKLVPLPHKHVDTGMGFERLASILQGKDSNYDTDVFTPLFAVIQKLTKAAP
YTGKLGEEDPDKKDMAYRVVADHVRTLTFAITDGAVPSNDGRGYVLRRILRRAVRYGQQFLNAPSGFLTQLVPCVVEML
GEAYPELVDKQKKVEDVILDEEKSFGRTLNKGIERFKKIAQSIREASAGSNKPLVVPGEDAFFLYDSMGFPFDLTEIMAEE
EGMTVDKKGYEACMRLQSERSKMDRKKGGSNGARPLVLEARETSALADKHVDVTDDTAKYEWHVKTPAKVVALFTTT
ESSSDFVDEVNAGDFERVGVILDKTSFYAQAGGQIYDTGVLSSGNFKLDVDSVESYAGYVMHMGPIAFGSIQVGDAVEC
EVDYVRRAKVAPNHTMTHVLNFALRKVLGESVDQRGSLVDELRLRFDFTNNNALKPNQLAEVEAICDDVIKQQLEVFTQ
NSAQAEATRIEGLRAVFGETYPDFVRVVSIGQPVALMLEDPENKLWSNFSVEFCGGTHLKNTREAKRFVLFEECAIAKGI
RRVSAYTCDLAIEAEERGEKLQAELDAIDKLSGIEFVERVSSFKPVLDQAVISLPLKDKLRKHVDGLVSRVKTIKKEAAAAR
AANGVHDATAEATKAKDAGQEIVVVKFDVGTDSKLGREMLDAMSAIIPTGSFMIFSTDSDTNKTAAFTRVSQQHADSKQL
DARKWVNHAMVVMNGKGGGKDVLNATGQAKTVEKVDEAVTLAKGFIQ

804029

MSGKTYKIAVLPGDGIGPEVCDQAAQVLTTIGELFQHQFSFTKALCGGAAFDEHQVHLPQSTIDTVAASDAVLFGSVGGP
TDAQEDPKWKDAEKNCLLGLRKNFQLAVNIRPAKIYSMLPDLSPLKPSIIANGVDMVIVRELVSGIYFGEHSTNGDTAVDV
MKYTEAEITKPMKFAFETAMGLIQNPSQFDVIVTGNMFGDILSDAASVLPGSLGLMPSASLGDKVHLFEPIGGTAPDIAGK
DLANPIAQILSAALLLRYSFQMETEAQLIEKAVEQVLLDEVEQITMQKLGIRIPQSLH

812913

MELKDLAPRLLLKERANGDIDPNVLTNILRGGKVANDRRKHLIQTIASHPVLSDRDMMFRNHTERYEFGLKKAFYYVKLL
QEGNYSDEEDQKILYKALGEPLGVDVHRAMFIPTLENQGDESQRAKWLPLAKQYKILGAYAQTELGHGSNVQGIETVAT
FDKATQEFILDSPTLTSRKWWPGGLGKTANHAIVHARLTLDGKDVGVQAFLVQIRSMHDHMPLPGIKVGDIGPKVGFNA
VDNGYCAFHKIRVPRENMLMRYAKVLSDGTFVKPQSDKLVYLTMVRIRAYLIETLATLMGAAATICTRFSAARIQGRTPDL
KGEFQVLDYQNQQHKLFPLLAISYAAQFAGSSLKELHDSALDAIKKGEPRFGTKLAELHAVSSGLKAWIAEHVSDGIETC
RRLCGGHGFTQSSNLSHIFAEIVGANTYEGTFDVLVQQHARYLLKTLALLPANENSTKFLSQVTRYSDSELRCDAETRED
FDNLDLLLEAFQARAARIILMLASHMRAANNDGNACMVLMARASTSHAELILLESFIHGVRNIPSGREQQAVAQLCSLFGA
WLITKSLGDFRQHDYLSSDQADLVRKQVVHLLPVIRKNCVLLTDAWDFSDFELNSTIGRYDGDIYRALVKRAADEPLNKT
QVPEGYTKYLKPLIQSDL

811591

MSLVSLLTRRLAPRSSPVPFRCRHLMSQLASSSPSASAVHAQLASLQGQNFLSTMDLNTAQLTGLLSTAQELKRTYSGP
DKASAPKPLTGESCSMLFQKRSTRTRLSSETGMYLLGGKGIFLSSEDIQMGVNETLKDTAHVLSRFNSLILARVYAHKDIH
ELAQLASVPVINALSDKFHPLQMLADYLTVHEHFGKIEGLTFAWVGDGNNVLHDYMLAAPKLGANIQIATPKGYEPDQDV
VDETKRLAALAGTTVVMTTDPVEAVKGANVVATDTWISMGQEEEAKKRLEAFAGYQITKKMLKHAATDHVFLHCLPRHQ
DEVDDEVFYGPRSLVFDEAENRLWTVMAVFSSLLGKF

800080

MWTNWIAAISLAVAVTVPRLPGAEAAVSPDSYDAQAQAIVDKFSILEVIGQMTQVDISTVMNPANNTLDEAKVRAHARLYI
GSYLNSYFSERRGIGTGWTATEFRELVQRIQEISMEENGGHPILYGLDSVHGANYVGGAVIFPQQINMGASFNPELVYQA
GQITARDTQAAGIPWIFGPILEISYNPLWARTYETFSEDPYVVSVMGDALVRGLQSYEQTAACIKHFIGYSKTSTGHDRDG
VSLSDFDLLNYYMPPFKAAIDAGAMTMMENYVSINGEPVAASSRILNDLLRTDLGFDGLLVTDWAEINNLKDFHRVVKTQ
EEAVALSMMQTSIDMSMVPTDTNFIGHVKKMLEKHPEEEARLRESAKRIVRTKLQLALYENPVPGEQYVSMVGNDDDLH
AALTMARESIVLLKNADNVLPLQKDASVFLTGHSADNVGFQCGGWSRSWQGYSGNELFPHGISVRQGFEDLVGNNSFI
HFNGLLADGDVSDADLTAVKHASEHEYTVAVIGEATYTEKPGDINDLALPAGQVRYIEALSKTSTKVIVVLFEGRPRLLG
SIPKNAVAIINGLLPCELGGQAMAEIIFGEVNPSGRLPLTYPKDPANIAIPYNHLVSTRCANDNCEMEFEFGTGLSYTQFSY
TGLSLDTTRLTSASATVTATVTVTNTGSRAGKETVMLFVIQPYRLISVPEMKLLKKFKKIELQAGQSMDVSFTLTTDDLSV
YDPQIGHGLKRVFEDSDYVVALKPETNCDVYNDNSKHPLCAKFSVDTSGRATRVNTKGGDGPGFFDASDDGGLPLLTF
V

807220

MTLNAPDTALRAAIECDNGKTLLADADVPVDVGGASSTQHQVAQDDGQPLARSEDLQVSARAAPPPSVKHKKSNVSWT
VGSLGMLVLAVAVPIVRGGPLYSPTKPQAPAYNAANWIELLPHDILSHMDTQVDPCDDLYAFSCGSWLKQAQIPDGESR
VSLTLTPMRDDNEKVLQEVMQQGWPLLGELYDSCMNFSNTSSATENVASVAFLMPAIKQIAATMSKSELFQLAGTLSRS
GPYFLTDLVVEPDQRESNTYALHASQTGLSLPSPEYYLDPKEFDSVSAAFHAYIVELFSLVGWDSHEAASQASSVIAFEQ
TLAPLFVPEEKLQDPVATYNRMSVAQAAEKYPLLVAQVMRGVGVLEDLVARNAYVVVTTLAYFKRVEELVTGDSVTLDT
LKAVLTYQYISAQANFLSEPFAQASFTFLEQTLKGEKSRAPRWKVCQKQVTDTFPDLVGKYYALLRFDEDSERVANDLV
MRIRASFEEDLTHVDWLDGPTRQSALEKLGNLALLVGYSNRSEHFPYHLTRDAPLADNLRIINEHVCDQTMDLIGRPVDR
YKWDGSGAALDVYHQPTANLVGLPAGFFQPPDFTPAQQSVRNFGACGSFIGHKITHHFDTEGKSYDKDGTLGNWWSN
FTATEFSQRADCFVKQYSEFPVMSSEDPSKVLGHVDGELTLSENIADNGGVKLSFQAYRTYVAKQAKEPSKINEGGATSI
SRSQPEPALPADAADKLFFISFAQSWCSKRSDAAMIQRLASPRAPGQWLVNGAARNSHDFARVFSCPVGSPMNPTTKC
QVW

809066

MAIKHVAHALVGALALAACATAFEEEDNVLVLKESDFADAVTGHDALLVEFYAPWCGHCKQLAFDYAAAATSLKKKNLPV
RLAKVDATVETKLAEQFAIRGFPTLKFFKGDVNAVKDFDGGRTAAEIEKWVVRKSGPAVVIVETVQELEELKEANDVVVF
AVVDAVEGEARALLEKLADADDVAVYVASTQLDLVADAKAVKNVVLYKKFDEGKVVVYDGEWENEALAAFIKANSLPLVIT
FSQDKAPMIFGGDMKEHVLAFVDMSKDYVGGLEAAVKQSAETNKGKLLHVMMPSTEERILDYFGLTPDDLPAVMLVNM
GGSMKKYGFEHKGDALVARLNGDLAVELVAFESEYFEGKLTPQLKSAEPEDDSDEVVKVIVGKDFQERVIDNEKDVLLE
FYAPWCGHCKALAPKYEELGEAFADVDSIMIAKIDAAANEVDHAGVEVQGFPTILFFPAMDKQNPIVYEGSRDVDGFTEF
LKANAQKFELDGVQYGTNEIGKGDEVEQTSEEDVKQKKEEGAEHEEL

802420

MSKLDHAVYDQYMSLDTGNVTLAEYVWIGGSGQDLRCKTKTLTHDVASVADLPVWNFDGSSTGQAPGEDSEVLLRPV
AMFKDPFRRGKHLLVLCDCLKPDMTPIANNTRADAARVMEAAAVEEPWFGIEQEYTLFEKDGVTPYGWPQGGFPGPQ
GPYYCGAGAHSVFGRMIVDAHYRASLYAGINVSGINAEVMPGQWEYQVGPCTGIESGDHLWMSRYILLRVCEDFGVNV
SWDPKPIPGDWNGAGCHTNFSTKAMREDGGIATIIEAIEKLKLKHKEHIAAYGTGNERRLTGRHETASMDTFSYGVANR
GASIRIPRVAEAEGKGYFEDRRPASNMDPYVVTGRIVKTTILNEA

803975

MSSAVPKYIDEHHGFGTTAIHAGQDPDEHTGAVAVPISLASTFAQSSPGVVTGRGQANSFGKGWEYSRTGNPTRGALE
RALAACEKGRFAVCYSSGMAATTAVTHLLKHGDHVLCIDDVYGGTQRYFRQTVQPTYGIDFDFTDLSDLSHVQTLLRPG
KTKLLWVESPTNPTLKITDLRAVAEFAKANELLLVVDNTFMSPYFQNPLMLGADLVVHSITKYINGHSDVVGGVVVTNSEE
LNDKLRFIQNGIGAVPAPFDCYMTLRGLKTLHVRMATHARNAQAVAEYLESHSEVEKVCYPGLPSHPQHAIAKTQASGF
GGMVTFYVRGGLDKARAFLENLHVFTLAESLGAVESLAESPAIMTHASVPPAVRKELGISDNLIRLSVGIEGLPDIVADLER

ALAAEPKASS

814239

MLRLVAARPVAPAATKRLFSAAAGSEVIGIDLGTTNSCVAVMEGKTARVIENSEGARTTPSVVAILDNDERLVGMPAKRQ
AVTNSENTFYAVKRLIGRKFDDKETQEVAKVVSYKIVKGTNGKDAWVEAKGKKYSPSQIGSMVLTKMKETADGFLGKPV
TQAVVTVPAYFNDSQRQATKDAGKIAGLDVLRIINEPTAAALAYGMDKADGKVIAVYDLGGGTFDVSILEISGGVFEVKST
NGDTLLGGEDFDEELLRYLHLNMKITRSTFEKLVGKLIERTMGPCKKCVKDAGLEKSEINEVILVGGMSRMPKVQTTVEA
FFGKKPSKGVNPDEVVAMGAAIQGGVLRGDVKDILLLDVTPLSLGIETLGGVFTKLIPRNTTIPTKKSQVFSTAADSQTQV
GIKVLQGEREMASDNKLLGNFDLVGIPPAPRGVPQIEVSFDIDANGIVNVGARDKATGKEQNIVIQSSGGLSDAEIEKMVA
DAEANAVADQQRKELIEAKNEADSVLYTTEKTLEEHKEKIDADTLEKVKVALSDLRQKSESENTEEIKAALKEVQEMAMKI
GEAVYKSQQSDSAAGSGDSKDENVHDAEFKDKKD

802267

MKLSTFTSIAPWMASPVVNLVLCSIPSVRGEPITDTVVTRLGDAWATLTHSHPAFEGWQRYAAVDSIKPTASSGLMTARH
LEPERSDVARLEAFFGTTMESNYNVLNDLYPRAVSDELPWSGDNWPTYKDGINAVWKANEPSPAQKYASAFNLHVDTF
LSSLSLSNGVLSESSSSSSCRTDADCVTAQDNDNDDRRCGLRANTSSGYCIPAWHGLGHARAAAALLEDEPQCDVSKN
NVTFHAVDIKALLTQLYDDAAITIVLTGARFNGPDAPEELDRYGRYTDAARRDVNAGFFHMAMTNILGKHRQSFIVDVSAN
AQVWNEPVQAFEILEASVVNASMLSLEHFDSDTYPFNDAATFLAKCTTRLTRTVESMDDGGELTSTERIKVHTVYEDYEY
CLELDGNYTIIGGEWLGASRKNHPDFLWLPTGKPRQDTVTATGISYANVLELLKESRQCEHTTGTLPVEETPSISTQRPK
CDNSWVHDKKRPANAVASQTPIVTSELNETYPPATDMPATYALGTSVPSTPSSKSLQSDSSTKQWNEPLSWDMSTLSQ
WYRHLADTTLWW

808833

MKSNRDLLKGSIILNTVLLFTFTEYSYSVSQFSSARTFEDSIDSRVVDVPSEHFTGSSDEILTLVGKPSTYTGRLFSGYLPL
ENGGQAFYFLAPSQSATSRTDPVLLWLNGGPGASSLLGCFSENGPLLVNEDGRTLRVNNFAWSQRANLLCIESPVGVG
FSYNISGVYEADDLSQAQDLYDALHKFFGKFPWLQENDFVIAGESYGGVYVPTTALAILKGNAAATQPSERINLKKYAVG
NGVNEYSGLATTMFAYYHGLLSTEKYQNIRSVCSNLHEFEKSALGVPGIGRASSDCTSALVDVITTLVYDRINIYDVYGSC
VGSVEEDIQRLVNRTSIPPTPGKLPHPVGNTIEMCLDTKHIESYFNLVEVRDALHADPALAYWSGNALTTASLDLLSAGFG
LDRSLFQHSLTLKYTPSLNFEVTKLWRQLLDAGLEGVIYHGDADLMCNAMGGLWAVESLGLPRLAPRSAWTYEEKGTN
QTGGFVELFEGIKYVTVKGAGHLVPASKPAEAKQMLDLFVLSDLEL

812621

MWRPIFSIPRRSTTTRCLVPQRALFSSSSSSSSSDYDLVVIGGGPGGYVAAIKAAQLGLKTACVESRGKLGGTCLNVGCIP
SKALLHSTHLLHTAQHEFKSYGIDAPNVTANFPQMMKAKEKAVRTLTAGIESLFSKNKVTYIKGVGKLAAKGQVSVTLAD
RQGNGTLTAKNIILATGSEVTPLPPVPVDNDAGKIVDSTGALALTRVPEHLVVVGAGVIGLELGSVYKRLGAKVTVVEYLD
SACSGMDKGAVKEFLKLLKKQGLEFQFGTKVTASEVNGDVVTLTTEPANGGDGSTIKCDTVLVATGRRAFTAGLGLEQM
GIQTDKQGRIEVDHAFQTQVPGVFAIGDVIKGTMLAHKAEEEGIACVENIAGKHGHVNYGAIPNVIYTFPEFASVGKTEEE
LKAEGVEYNVGKFPMMANSRARTVAEADGMVKVLADKKTDKLLGVHIIASNAGEMISEGVIGIEYGAASEDLARTCHAHP
TLSEAFKEACLAAHDKSINF

801302

MRRHVLHRLTRAISCSSVAVKSSTVTLLPRHFSVEVPPVKQTRLFINNKFVPSSSGATFDTFNPATEEKIASIAEATRADV
DAAVDAARTAFNGPWRTIDASERGRLLYRLADLIEQHGDELARLEALDAGKPTTLARVVDVQATLNVLRYYAGWADKIH
GITVPISGPYLCYTKQEPVGVCGQLVPWNFPMIMMSWKLGPALAAGNTVVLKPATQTSLTALRIAELIVEAGFPEGVVNIV
TGGGKTVGAYLSQHPNVDKIAFTGSTRTGMAIMRSSHVDNLKRVSLELGGKSANIILNDADIDAAIEQSQKGLFFNSGQL
CISGSRVYVQEGIYDEFVRRSVEAARTIKIGDPTDPRTQQGPQIDGGQLRRILDFIELGKQAGARLMCGGKRWGDKGFF
VEPTVFTDVTDDMAIAREEIFGPVMSILKFKTIDEVIERANNSEYGLGAGVVTSSLDNAIKISNGVRVGTVYVNCYAVIEPN
APFGGFKNSGFGREQGEQGLRNYLESKTVIIKHPDDSMP

809865

MAESLSKKPRVTGPLTALNLLNGAFVPPVEEKYIDVISPSTGQVIGKCALSSDTDVQQAVAKGQEAFAQWRRTTVKARA
AIMFKFHTLLKQYQDAIEHRINYGHLFIYFIFHPDELVDLVVLENGKNRTEALASIMKGNETVEYACSLPQVVQGRTLQVSR
GITCQEVRDPLGVVACIVPFNFPIMVPMWTIPIALTMGNCVILKPSEKIINGTADPVNSLCDHPGIAAVTFVGSSHVAQLVA
RRCRALDKRVLALGGAKNHLVALPDADADVASKDIVASYAGCAGQRCMAASVLLLVGDCSKLLDTIVAKSKELIRGTGAG
QVGALIDDASKTRVLKYINQAEATGVKVLVDGRSWAQQSTGFWVGPTVLLHTNPKDAAMTDEIFGPVLSVYCVNSFDEA
LRIENENEYGNAAAIYTSNGAYAEYFQSRFRAGMIGVNIGVPVPREPFSFGGMYGTKSKFGDMDITGDGCVEFFSTRRKI
TSKWSAGNTSGDAANFSGQL

808413

MSSTDALCCSDFPPMNDAICSSLFQPIEGCAATSVYPRCSVVKKISRPQATNPKDLFFSLNPMSSSYSTTVVASGAVLLS
TLSQRAMHIQALLQQHVSERVPSKLRVRFAQAVLAAPSAVGSNTDVQGLVTIVGDEIAQLLSALQLLQLWIQLQVPKVED
GNNFGVEVQKYAYVHLKESHEKWQKTWDSLAEYCSLRATAVEKLNMKASSESSTTTTVTSSKGGKEGDEEKSVTAKVE
KQSHTGSQVAADALAAVVELDVKWYMNLVRALEGVRDQYAVTDDVISKNKEKIELPRGKNDRAGFNMF

806837

MRLLIAAFNSLFAASCLLTQVLSWDQQPDGYEVIDGPNNLKALFSWRQDWTIWKKMELSSNRYNPTDACNIYNIPETQW
TQRNFVQTFIMLNDRAIYDRETQQYTVDKYVDEMEQRVGTIDSVLIWSAYPNIGIDDRNQFDLLRDLPNGLEGVKKVVAD
FHRRGIRVIIPYNPWDIGTRDESGLEDTVRMYTADIISLTEIVSELQADGFNGDTMYGVPKSFFNCSKPLVASPEGGVPTA
YLSHNPMSWGYFWGYSHFPPVARAKFLDSRHMVQICARWSLDRLTELQTAFFNGAGYVVWENVWGIWNAMTEREDE
TAKRMFAILRKFGHVVSTGVWTPYYAMKDDGLFASAFDVQENGESLYTVISTVPEDMTFELPLTKEQSGDGFRVYDVYH
GVELQKQQSGVNGSVVTLTLEPRAFGAIYITKATDDMTQFCGAMRAMTTKALADYSTTRNLLQQQLIRSDHAEAEAKSS
VDGMVRVSGMDNWWFNVSGVQIEPVLAWTPNSVQFGTGVQFPWENRPWNNHSTRLSIQDFFLDVHPVTNAQFRTFL
QSSGYHPKSMERFLLHWENRQGPLASWSIPTALEQSPVVNVAHEDAEAYAKFYGKRLPHDWEWQYVASNGENYDAY
PWGPEMDRNKVPKVFHGKELPALSPVGSYVNSRSTKFQVEDLVGYVWQMTDHFCDSHTCGLLFRGGSNYHPISATHS
DPNWYFPQTLDAQHHNRFLMISEGYDRSPMVGFRCAKSIAPSQAMN

810322

MINDSPAWKVLEAHAAEIQATHLRDLLSNDARNAAMRTHQEGIHLDFSRQNATPKTLNLLLDLAEAAEVTKKLQAMASGE
HINSTEDRAVMHMALRAPKTCSLLVDGHDVVPQVHQVLDNIRIFATNVRSGAFVGATGKPLKHVISVGIGGSYLGPEFVF
EALRHEPVAKRAAAGRTLRFLANVDPVDVARATSGLDPEETLVVVVSKTFTTAETMLNARTLRKWIVDDLVAKGSSETEA
VAKHMVAASSAVPLVQDFGIDRDNIFGFWDWVGGRYSVTSAVGILPLALQYGYDVMEQFLAGAHAMDKHLLEAPLRSN
LPVLMGLFGVWNSTFLGHSSRALLPYSQALLRFAAHIQQVDMESNGKRVSIEGVELPFQTGEVNFGEPGTNGQHSFYQ
LIHQGRVVPCDFLGFCESQNPVQLAGQPVSNHDELMSNFFAQPDALARGKSLEDLKAENVAEDLCNHKLFPGNRPSISL
LFKKLDAYSAGQLLALYEHRTVVQGAIWGINSFDQWGVELGKVLAKQVRAQLSASRNENAPVEGFNSATTSMLEAYLAY
KS

807402

MSVLIKSIAARSGSAIAPAAARAMSTSSTSLLDELKQQLPVRQEALKKLKAEHGHKSLGEVTVEQTLGGARGIKSLLWETS
LLDAEEGIRFRGHTIPDLQKVLPTKIQGGEPLPEGLLWLLMTGEVPTDDQAASVTAELHARAKVPEHVTKLIRDLKHAHPM
TQLSAAVTVMNTESVFAKKYAEGIHKSTYWEHTYEDSMNLIARLPEIAALIYRNTYYSESNHAYDASLDYSANFNRMLGY
DNAEFDELMRLYLVIHSDHEGGNASAHATHLVGSTLSDPYLALAGGLNALAGPLHGLANQEVLGWILDLQAEFHSKGLE
VNKETITQFAWDTLNAGKVIPGYGHAVLRKTDPRYTCQREFGLKYMPNDELFRIVDTIYQVMPGVLTEHGKTKNPYPNV
DSHSGVLLQYYGMTQKNYYTVLFGVSRAIGVLSQLFWDRALGLPLERPKSVTSEWINNHFANKK

808045

MLLPALLAPPTTASIGHILHFSDVHLNISKSESSSSEDVPIRYFADAPLTLLESALRYAKEHVVANPKLFLYTGDHAVHGSL
TDEFIAKVVETNVHTMESYYPSKRGMMDVTAIIGNADGNPDYHMEVTNPKTETNPSIELISSAWEDSLSAGHMNVLNRR

GYLSYALNDRLHVITLNTVPYSFAWLDKTLAELQDAGKVAYIAGHIAPIVDSYGGDPQWHTKYIVKYKSIVGKYANVIKAQF
FGHVHSLEFRVPVVSLDGADAESDTFQLLPMFMSGSISPLFGNNPSFMVWEYDTDTHEVRDYAVYASDIRNSKPRLEW
KLLFKASDAYGLNSLSLTELSAFVRRAEQNVSLVEDYYWNMKARSPNAGACKDTVCHAKTLCSLKWWTTKGEYLACLD
TIRLTMASHVVDYDNTKMSGSAAVIAEAPLNDSRMNGVLMTIGATALATAVAVVCVALTVYALRRSGILKYSTTYESVTTS

803469

MADKQAQLVFVYGSLKTGFNNYPVFIQPAVEQGKASFVGAARTTNDEFHMVLNDQAFFPCLYRVPKGDGYKVSGDVFS
LDRDTIAALDAFERVDDKLHVRGVLDVDLVDGERKGETVSCQVYFMSVYEDLARLERITEYTRAMNAKFAQFMADPLPD
FVESILDKMPKQQQT

804440

MLARQLLASKRPLRSVVLRVSATTHARFLSAPVFPEDHEEDEEVMSGGEPSGSEYDTLDTYLMTLKKYNPPIKTGLKGL
NLVHDPLYNKGTGFPHVERDRLGLRGLVPPRRLTVKAQLTKLYDTFSQEKDTLSKSRFLTDLHDRNETLFFRLLIKHMKE
MAPIVYTPTVGLVCQKFGHLFRRPRGMFFSTRDRGQFGAMVHNWPSDDVEVIVVTDGSRILGLGDLGANGMGIPIGKLA
LYTAAGGIDPRKVLPVMLDTGTNNQKLLEDPYYLGVQQPRLTGEAFWSMVDEFMRAVRHRWPKVLVQFEDFSSEHAA
DVLNAYRLKQLCFNDDIQGTGATVLAGALSACERVQIPLKDQRIVILGAGSAGLGVATTLLQGMLREGMTADEARKQFYIL
DQYGLLGESREMLNSGQQFFSRSDIADKTSLVDVIKQAKPTMLMGLSAAAGAFTQEAVEEMAKHTRQPIVFPLSNPTSV
AECTAEQAYEWTKGKCVFASGSPFPPVMYDGVQNIISQCNNMFIFPGVGLAASVIQATRVTDGMLYSAAKALSQCMKPE
EIASGQVFPSVENIRDVSLKVATAVCETALEEDVAGLRPTIRRGATLEDFVASKMYYPTYHALVE

806975

MNTHFAIAAIALAVATSVNGQGDCSPEVTKAAYTSMSSLLKRAELMSCGDRSHYNFMTAEQPANHEQELAMCGVAECH
TLIAEVKELNPPDCVISIPGRFPINIKAMADAFEGKCKSPNPARSAIEEPTESAMLTAAAPSPDVSEETDVIQQDNDFTEKN
TTAYTPGKALDSVCGNTIVSW

809424

MRRSLRLAESEHIVMAFQQQALHLAAVRLDPSLSDFQSRRYEINMGKEKVHISLVVIGHVDAGKSTTTGHLIYKCGGIDKR
TIEKFEKEAAELGKTSFKYAWVLDNLKAERERGITIDIALWKFESPKFFFTVIDAPGHRDFIKNMITGTSQADCAILVVASGV
GEFEAGISKEGQTREHALLAFTLGVKQMVVAINKMDDSSVMYGQARYEEIKSEVTTYLKKVGYKPAKIPFVPISGWEGDN
MIERSPNMAWYKGPYLLEALDNLNPPKRPSDKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVATFGPVGLSTEVKSV
EMHHESLPEALPGDNVGFNVKNVSVKELRRGFVASDSKNDPAKGTQDFTAQVIVLNHPGQIGNGYSPVLDCHTAHVAC
KFKEITEKMDRRSGKVLETAPKFVKSGDACMVILEPSKPMTVESFAEYPPLGRFAVRDMRQTVAVGVIKSVNKKEASAK
GGAKKK

809723

MADLSDIGLYGLAVMGQNFALNMASHGFKVSVCNRSPDKVDATVQRAHDEGELPLVGFKDLEAFVASLARPRKIIILVVA
GKPVDLTIAALATFMEPGDILIDGGNEWFPNSVRRASELEPKGIHFIGMGVSGGEEGARNGPSLMPGGPREAFDAIEPILT
KCAAQVDDGACTTYLGPIGAGNYVKMVHNGIEYGDMQLIAEAYDILKLAGGLTNDELASVFDDWNKGELESFLIEITAQIF
AKKDDLEDDGYVLDKILDKTGMKGTGRWTVQEAAERSIAAPTITASLDARYLSARKEERVFASKILAGPSEIPAVDKQQLI
DDVRQALYASKICSYAQGLNLIREAGVQMGWNVNLGECARIWKGGCIIRAKFLDRIKAAYSKDATLISLLVDPDFAAELQA
RQYSWRRVVSLAVASGIPTPSFSASLNYYDTFRRERLPANLTQAQRDFFGGHTYERTDREGLFHCAWSAAHHSIGNVQ
DRIRGNL

812115

MASSVRTFFFSRGLLAACFFVMLLSAAIDQYPGEIEPRPLEYDYDLHSREQESRYNETRSLIRNTELFDASILEHLFINQSL
SAEDLVVGPDGLAYVGLADGRLASFQVQATELRNFSRTGQDVAGCGALDMEPTCGRPLGLAFAAARPFAKFINRIPAAR
LFPGDQVLLVADAYKGVLLFDANGQRTLLFSRVGEQHVNFLNGLAVVQETGEIYVTESSRRFQRNRVVVDFLERRSTGY
LLRFDPRDESVQVEASDLGFPNGLTLDKDGTGLLIALMFQNKIVHFDRTTKQMKDFAFLPGEPDNISIEKVGAGETETHVL

MVGLVSRNDGGAFDYVKQSVKTRKMLSVLPTWATVLFIQRLGVFASLDLETGEIRHVYEASQGQTPIVSGAKRFGDHIYL
TSWSRSSLTRIPAAMVQ

808637

MAAPKKEAIAIEQQELAPMEKSIWYKEALTEDLYLSVALKTITFEAKSKFQTMQIIETHSFGKTLVLDGKTQSTKSDEAIYHE
TLVHPALLAHPNPKTVYIGGGGEFATAREVLKHKSVEKVVMVDIDELVCNMCRKEMPEWADGAFEDPRLEVHYTDAHAF
LKQYDGTFDVIIMDIADPIEAGPGYVLYTEEFYQYAVTKLNRGGFMVTQSGPGAVYNWHECFSSIYRTLKTNFNTVMPYT
VDIPSFGCPWAFNMATNLDDGGDSKAAVASIRERSIVTTDELIETRIGKPLVSLDGVSHLGLFGLPKTVREALQNETRIITV
DNPVFMY

803265

MLAPNKLRELRALMSRTNVRLLPLRHRLGKTEKEEAAVEPTIQSLQAFIVDTADAHQSEYVGDAHKRREFLTGFTGSSGT
ALVTMNEALMWTDGRYFLQAEQELSEDWTLMKSEEPGVPTIEDTKIELVALHETENLIDRVWTDRPAISPSQITFLSEKYT
GRSVADKLTRLREAVKEEGADAVVLTALDDIAWLFNIRGNEVKFNPVVTSYAIVTSDSAILFLNASNEDEVTQHLGSSGIV
CKPYSNMLKELSAFTSANKGKQILVDPAQCNVAVFLAIPAVNRKEVMSVVMAQKAIKNATEIEGMRQAHIRDGAALVKYF
CWLENEMDSSHKDRWDEVMVADKQEQLRQQVKDFVSLSFDTISSIGANGAIIHYSPKRGNCAKMSTSAMYLNDSGAQY
LDGTTDVTRTLHFGHPTSYEKACFTYVLKAHIALASTVFPDKIEGVRLDAVARAPLWKAGLDYRHGTGHGVGAFLNVHEK
GVLMSFRLNPNGLKIQDGMILSNEPGYYEDGKFGIRIESIMVAQKVSCSSCASLSVVMCIRDALLTACLYLYV

804695

MTTEPTVYGTSSLASDERLSARQTRRNVVFKWLLLFVSVCGLAAIVAVFGFSTRSRPEASVASIKTTTSPTAAARGSDSD
LSCFQSSYVVNVTVLMEPIAGLKWTRAVHSDAASVITVDVATQFQEIMGFGGAFTEAAALQFQRLSPEKQEELLTLYFDP
DKGSAYTFGRVPMGSSDFSRASYSLDDVVNDTTLVHFDHDVKHDQEVLIPFIQRALERQPALKLFLAPWSPPAWMKRK
ATNYVPSMLNSAKPIGLKDDVRATWALYFSKFITAYKKHGIAFWGLTPQNEPQAEVPWESCMYTAEYQAEFIGNFLGPTL
ERDHPELVLMVFDHNRGSVHQWATEIYNHPTAAQYVDGMAFHWYDEQRYMDGVEYHERLNDTHFVDQNRFMLATES
CNCPGVAHGDLAWFRAQRYGHDIMTDLNNYVAGWVDWNLLLDHEGGPNHLANKCDAPIILTQDGTDFYIQPMYYFIQH
FSKFIPVGSRRVKTQVAAKFTRPGDPQLYVDYPSMLAACDGSSRQQWRPTDDGKLEVTGTGFCIDLKEIPWQGHQGLL
VECRYTQQHWTYELDTGRIRMGDYCISLNHGSTENGVRISADLCEDAVVPHQQWRFNVEDGTMRSFASTTDQCVTAG
YAFVQAAAFVTPENRKVLVVLNENTEPIDFELQVSGNALQTTVPRGAIRTFTWD

812337

MNLQRLLLVASSSLACAASKFTITPGTTQAATSTVDNVKFVKQWTLTAATTTDLINSIDLDLAGRVYVSYRSGLPTGVLGY
VNVSGDSQTVVDAVGVAHDDANDNDNDDNDDENEANDPDGDLNVHINSNASASLTGYLLTEVILASTGIVMDLQSKRSA
VVVVEDGVLMTSSTTAELQVEASGSSAIFVSAADAAVSMRQLQLDASGSASLQFEVASVSVSNEAQVDAQGSGQVVVL
VPTVEAAMLEIDTENTGAICFSGQQVTATTYEGKDESRISMPNAADKHGSTGTAPCEKATVPPREPACVGTGCASGTST
TPGAAGTSTTPGAAGTATTPGAAGGNNAGISASDGTSASSTPQLAVLVAFAAVGIAMATLL

807649

MSLDFSKVENNPLPVIAEKHNPNVVFVVYRNKNKNVVVYAARLLEDGTLDPENPLDVHWIMFEQDGEPREDLNMIERTT
AYGATVKPREGHPGEYVVTLTSLKDRVIYLSVVDGKPVGRGSINGQDGCTLERVFVQQTSSWGMPKVQHIEIHGRDAS
GNATSEKKIP

804180

MSDHVDKKPRIGEELCINTIRMLSADQPSAGKSGHPGAPMGCAPMAHVLFGKTMKFNPKNPKWTNRDRFVLSNGHAC
ALQYSMLHLTGYDLSIDDLKQFRQIGSKAPGHPENFCTPGVEVSTGPLGQGISNAVGLAIAEKHLAAEFNKDGLDIVDHFT
YVICGDGCLQEGVSSETSSLAGHLGLGKLIVLYDDNRITIDGSTSLSFTEDVQQRYEAYGWHVQSVEEGNFDHAAIEKAV
EAAKAVTDKPSLIKIHTTIGFGSNIENTHSVHGAPLKPEDLAATKEKFGLKGTESFFVPEQVKKFYDKTELGADLENQWNE
LFAKYAEAHPKDAAEFTRRMEGKLPKDWKKDMPKYTPDDAVKATRQYSEIALNAVATALPELVGGSADLTPSNLTLLSM

SGDFQKDTPIGRYIRYGVREHGMAAISNGIFAHGGLRPFCATFYNFIGYALGAVRLSALSRFGVIYVATHDSIFLGEDGPT
HQPIEMNASLRAMPNMYVYRPADGNETVGAYIAAVENQHKTSVLALTRQGLPNLANSTAEAVTKGAYVVANVANGAEV
ESLESEPDLVLVASGSEVSLSIDAAKLLTGHKVRVVSMPCRDLFDEQPIEYKKEVFGEGIPSMSVEAAATFGWSTYSHSQ
FGLDRFGTSATIAQLKKQFGFNPDAVADEARKLLKFYDGRSAPDLFDHPAPRVLKAIHH

807075

MAQFDVFRAKMQAAGLSTEAIKAFEYSYEALVSGETGMITEASIKSVDDLPYLENKAGSIRESIKADPALLKETVVLKLNG
GLGTSMGLDKAKSLLTVKGDDTFLDIMAKQVTELRSAHKSHVRFVLMNSFSTSADTLEYLQKYPELVEDETLELLQNKVP
KVDVSNMEPALYSSNPSKEWCPPGHGDLYASLAGSGKLDKLVADGVKYMFVSNSDNLGATLDLDLLTYFAQSGKPFLM
ECCERTENDKKGGHLAERTADGRLVLRESAQCADEDEKEFQDIAKHRYFNTNNLWIRLDKLQEELKQQGGVIRLPMIKN
SKTVDPKDSSSTPVFQLETAMGAAIECFDGAGAVCVPRTRFAPVKKCDDLILLRSDAYVITEDYRPIIAPEREGVAPIVSLD
SKHFKLVQQLEAAVRGNVPSLIKCDRLKITGNVGFAAGVVFEGSVEVVNKSSEQKTVLAGTYKDTVVDLSKQKGLGKLK
VTTVKTTPFQDQKPGTSGLRKKTKTFMSDNYLQNFVASVFDALPAKDLNGGTLVVSGDGRYFNKDAIQIVIKMAVAYGV
DRLWIGKDGLLSTPCVSAVVREREGGSVAFGAFILSASHNPGGPDEDFGIKYNCENGGPASEKVTEEVYALSKVITSYKI
ASEFPTIDIAKIGTTSVVADDESRTITVEVFDSAEHHVDLLKQIFDFHAIKKLVSRSDFSFVVDSMSGVNGPYARRVFVEEL
GCDESCLLNATPMEDFNGCHADPNLTYAKTLIKAMGVDAKGLPVLGQEQDPPSFGAAWDGDADRNMILGSRFFVTPSD
SLAIIAANCTVIPFFKNGLRGVARSMPTSGAVDLVAKKLNVPFFEVPTGWKFFGNLMDSNVVFGKEDYTPFICGEESFGT
GSNHIREKDGMWAVLSWLSILASKQVDGAPLVTVEDVVRDHWKKFGRNYYCRYDYENVDKAAAEVMFADMTKFDGVA
GKEIDGFKIDKADEFEYVDPVDGSVSSHQGIRYLFEGGSRVIFRLSGTGVAGATVRMYIEKYEEPTGNLGQNAAEALEKLI
NVGLKLSDLQKKTGRSTPTVIT

814377

MKILVPAIPLALVVTFSLTAAQRLCGQYDEIVLPPYTVYNNLWGQGDDPNGIQCTEATRILDKAIAWTTDFNWAGNPNQA
KSFAIVALSFTPVQMSQVTSMPTEIEYEYESVDKTLVANVAYDMFLSSSSDGETYTYEVKVWLTTFGIIAPFVEDPMNPIL
ANVTVAGVNFKLYRGMDSNVTVFTYVATANINRFNGDFKEFFTNLPDAKRLNDQYLVDAMAGTEPSHGKAKLIVSKYSL
AIIS

804992

MFVKTVRSVLPPARRGFASSSLKKIKVHQPIVELDGDEMTRVIWSQIKDKYIHPYLDLPIEYFDLGLPHRDATNDQVTVDA
AHAIQEHHVGIKCATITPDEQRVEEFKLQKMWRSPNGTIRNILNGTVFREPIVISNVPRLVPGWQKPIVVGRHAFGDQYKS
TDFLAPGPGKFEVVYTPADGGDKMTLEVYNFESAGIGLAMYNTDESIYGFAKSCLSFALSKNQDLYLSTKNTILKKYDGR
FKDIFEEVYQSEFKAKYDAAGISYTHRLIDDMVAQALKSEGGFVWACKNYDGDVQSDIVAQGYGSLGLMTSVLVAPDGK
TVEAEAAHGTVTRHWRQYQQGKKTSTNPIASIYAWTRGLAHRGKLDDNQELVDFCLGLEDAVIKTVEAGHMTKDLAICIH
GSNVTPDHYLYTEDFMDKVKDTFDAARN

808954

MEEAARQTNEIIAREFATPSQSPKIRDAAAVAFDPLSLGLEEDFVLTNFSALKGCGCKLPQAKLLGYLDNVMPNDVKPNE
TPGMDSSVVKISHGSGLYLVSTTDFFFPSVEDPYVQGGQIACANVLSDVYAMGVTEVDTMLMILGVCRDMTEKQRDVVTT
EMIRGFNDLARQAQTNVTGGQTVMNPWPIVGGVAMSIIRPENAVVGDVIILTKPLGTQVAVNVFQWKKKPEQWQRVNQ
VVTPHDADVAFQMASESMGRLNLNAAKMMHKHGAHSATDVTGFGLLAHARNQAKSQLEDVSFELHTLPIIKNMVKVNE
VIGNSFKLLDGFAAETSGGLLLCLPAENAEAFIKELRELDEKPAWIVGRVIAGLKDACIVPNPTIVEVSPED

804529

MAPVPHYVKSVALKNSTSHHVKVTATFGSDEFEADGKAKIHETRELAPGAEAKLDEREYDMGGWTAVAGLYSLEVEHS
TDRHVLGKTLYTPSVGGIVDVLHVDIGADENAKSFKVAAVREA

809219

MEVAGGKGDAAKGAKIFKTKCAQCHNTEKGAVHKQGPNLHGLMGRQSGRAPGYSYSAANKNSGVVWTDATLFDYLLA

PKKYIVGTKMVFAGLKKPQERRDLIAYLIESTSK

802703

MSRRPEIRTRDVRLNSIGVRAKIVLKNEMAKIFLHLPIDPVPSTSIEMTNNCSTLGTATPINESVVESEKQWNVQPSAFSKL
CSNPIREIVDNIKKPATSTKKLIPLSLGDPTVFGNLHCPDVLVEAIVRNTRSMQHNGYIHSAGSEAARTAIAEHFGKPEAPL
TMEDVIITSGCSGAIEIALQGLLNPGDNILLPNPGFPLYRALCAAYKIECRFYTLKPKNNWEVDLEHMQALVDDQTKAILVN
NPSNPCGSVFSKSHSEKILEIAELNKIPIITDEIYGDMVFGHNSFFPMATLTTTVPVVAVGGLAKQFLIPGWRVGWITVHDR
NNVLNDVRTAYFKLSQNILGANSLVQSAIPDLLTPVPGSVEEQSLAAFKKRLFATLDDNAKYTIGKLSKIPGLEVVVPQGA
MYAMVKIQTDVLTKIKDDLDFTQKLLDEESVFVLPGQCFGMTAYFRIVFSAPHEVLADAYARLEDFCHRHQ

810520

MTAKTVLITGSNRGIGLSLATHYKNEGWQVIGCARNVDAATDLQRLEPWKLVTLDTSSEESIAAVASTLQNEPIHLLINNA
GVMDAHDLQSTTKADLLRHYEVNAVGPFLMTRALLPNLRLAASAEAPAFVAQLTSRMGCIADNDSGGFHGYRASKSAL
NMLSKCLAIDLMKENIGCLLLHPGYVKTAMVGFHGHVSPEDSVKGLTTIIARAKLQDSAKVFHFEKGDVLPW

813189

MSDPIGELKELFIDNEWVPPVNKKYMDVLNPASEEVIQQVAVASAVDVGLAVQAAKRAFATWGTTSGTERAVYLRRMG
ELVAKRMDALSRLETLDNGKPLAEAVGDMEDVSSCFNYYANAAEALDARQYEKVALPLKDFQGALRYEPVGVVAAVIP
WNYPALMALWKLAPALAAGCTVVLKPSEVTPLSALQLAQIAVDAGLPAGVLNVVNGLGGEAGCPLVSHPDVHKVAFTGS
VPTGRTIMTEAAKEIKKITLELGGKSPAIVFDRANLERTVEWVMFGCFGNNGQICSATSRLLVHKDIADEFLEKLIEETKKL
AMGNPLDSKVQMGPLVSKAQQEKVLGYIQSAKDEGAIIVQGTLPSGNKGYFVPPTIITGVTKHMRVWKEEIFGPVLSVMT
FETEEVAIALSNDSDFGLAAAVFTSNDTQLKRVTMALRAGIVWNNCSQPCFVELPWGGVKKSGIGRELGPFGLNAYLEP
KQICTYVADKPFGWYLKS

806307

MMSFRTAALFAGLALTSSAAQEQQPVHDLGLNAYDQARDVTLNEVAFLTTTACHPSLYNADVTSRVCFTEFTTVTTQTS
GGGTYYKFQVKGCPVDTEKQLGYCREGACSTTSLYEVAIYSQPRTSAVFLTSIKEVV

811467

MKFSSLTLLAAAGVASTVSTVTSQITGDATTTATTTTTAMNTSSTTPSTTPSSSSPSGSMAGVSVGGGGSGGASATITFK
GGDVFADLETTKHPDDSYHMPPVRVIQARVQSDNPINVKGVFVSSFGDGDLEAGYLSAMDSVNTASVEGALMYVQAEG
ININVRAEEERCERKSGMANIVFYEILVVQTNETLAQFQSSWGKTPEYGPMLPMDSGRCTPLSGDDDFPAGCLQFNGDK
GQPNVGPFVGAGIKDDDVRAPYPENYWFSYPGTCPLEAWGDKTDECRASTRKGLCDYGKGPDGVECTFAYNILGWVT
IDDIVGITAIENPSTGSPYANFTEWCSADTNNTEFAADAETGEYQTGLPFWEDPLNTTANAARAQAVIAKYEEVLTSGSTQ
IESTVLAAFRPLPTPEELAAMNPPCYMTVEACGSGNGCKRVGYSQICTECEADEGCETGGNGFVYPTLAKAKTELSEEE
TTTKVADGSTTGGATDSKTDSTTSGAAAPMMFTAASVALGFVAVVLAL

808734

MTVKSAANGKFLVTPQLEKGKVCLSRGDDQLLHFQWVDRQTGASPEDFIIFPDDAHFDKVDTGRPDDRVYILQYKNSSR
RFFFWMQVTGGILKHFVSFFLLLILIGLLYQNKDASRDEELVKKVNDAMNNAQAASSNDGGRVSGNNVQLDHNAIMQML
GAMSAGDQGRGGASGGGVGGGGQAVQMSELQNILQNMGLPAEAQSTASSPATSAVSSSQASQNGGGISTAATTAST
QHDHDVSGMEVDEMDEDELLRLAIEESMRDGGSTNPDASGDGAAGGNSDSSSSNRPAGHGGHSDSDAVSAVVPPEP
VASSQFTSTAAVAASAPAGTNVAAGGTLTTADLQRAMASFQGLTQPKPVSLTKLLSADNMESVLDDPACVDALLPHLPE
GSQTLAELRATVCHPKLFEPILPCAVNSYDFALSPWLSIDCIHSVALSSVASKHRLARECSTERKFGRGHGQFRTGSCGR
CCETRIRRRRGQTSKAPQSLLEL

810603

MSESKPLWAPVKWAQRKEALYVTVDLPDVKDEKVTLSNTNLTFKGTSNGQEYEVTLDFLKEVDAESKESIWAKTDRNLH

FYIVKKNQDEEFWPRLLTDKHLEKTNVKVDWSKFVDEDEDEGQSSFDMSALNGGGGFDINQMMAAQNASGMMGDEG
SDSEDEDMPDLDPNEQ

809042

MTVGGAPVGRITFELFADKVPKTAENFRALCTGEKGVGRSGKPLHYKGSAFHRVIPNFMCQGGDFTRGNGTGGESIYG
EKFPDENFLLKHKGEGILSMANAGPDTNGSQFFVCTVETSWLDGKHVVFGRVVEGMDVVKNIEAVGTQAGQTTKPVVV
ANCGEL

810208

MTATKFSFIISLCLTALALVVQADSTPTITNQVYFDVEIGGKPEGRIVMGLYGEVVPKTTENFRALCTGEKGVGKSGKPLH
YKGSIFHRIIPNFMVQGGDFTDFNGRGGESIYGTKFKDENFDLKHAGKGTLSMANSGEDTNGSQFFICTVKTSWLDGRH
VVFGRVVQGMDVLNAMEKVGTQGGTPSKVVTIADSGELAPDGFVEEH

811164

MFTHATKFLRASSLRRTYSSSVTGQQKVAVLGAAGGIGQPLSLLLKDCDHIKHLSLFDVVHTPGVAADIGHINTHATVTGH
VGMEQVGDALKDADVVVIPAGVPRKPGMTRDDLFNTNAGIVQSLAQAAAKHCPKAMMLIIANPVNSTVPIVAETFKKAGV
YDPKRLFGVTTLDVVRAATFVAEHQKWNPRDTNVTVIGGHAGTTILPLLSQLEHAKFSDEDIGKLTHRIQFGGDEVVQAK
NGTGSATLSMAYAGARFTTRLLDAMNGAKDVVECSYTQNDVTKLPFFSVPVTLGPNGVEELHHFGNLSTAEQANFDEMI
VALEAQIKKGVEFANQN

Appendix 3

A.thaliana-sourced up- and down-regulated protein predictions of

MALDI-TOF:

Accession # (Up-regulated)	Annotations (TAIR)
AT1G26390.1	| Symbols: | FAD-binding Berberine family protein | chr1:9130164-9131756 REVERSE LENGTH=530
AT5G09590.1	| Symbols: MTHSC70-2, HSC70-5 | mitochondrial HSO70 2 | chr5:2975721-2978508 FORWARD LENGTH=682
AT4G12290.1	| Symbols: | Copper amine oxidase family protein | chr4:7304434-7306973 FORWARD LENGTH=741
AT1G26380.1	| Symbols: | FAD-binding Berberine family protein | chr1:9126901-9128508 REVERSE LENGTH=535
AT4G26970.1	| Symbols: ACO2 | aconitase 2 | chr4:13543077-13548427 FORWARD LENGTH=995
AT4G01870.1	| Symbols: | tolB protein-related | chr4:808473-810431 REVERSE LENGTH=652
AT3G48000.1	| Symbols: ALDH2B4, ALDH2, ALDH2A | aldehyde dehydrogenase 2B4 | chr3:17717082-17719843 REVERSE LENGTH=538
AT4G37520.1	| Symbols: | Peroxidase superfamily protein | chr4:17631704-17633060 FORWARD LENGTH=329
AT2G19500.1	| Symbols: CKX2, ATCKX2 | cytokinin oxidase 2 | chr2:8444311-8447301 REVERSE LENGTH=501
AT1G66700.1	| Symbols: PXMT1 | S-adenosyl-L-methionine-dependent methyltransferases superfamily protein | chr1:24873460-24874690 REVERSE LENGTH=353
AT1G30730.1	| Symbols: | FAD-binding Berberine family protein | chr1:10900854-10902434 FORWARD LENGTH=526
AT2G45220.1	| Symbols: | Plant invertase/pectin methylesterase inhibitor superfamily | chr2:18644281-18646394 REVERSE LENGTH=511
AT1G13750.1	| Symbols: | Purple acid phosphatases superfamily protein | chr1:4715490-4718091 REVERSE LENGTH=613
AT5G04360.1	| Symbols: ATPU1, ATLDA, PU1, LDA | limit dextrinase | chr5:1221566-1228399 FORWARD LENGTH=965
AT3G23600.1 (+1)	| Symbols: | alpha/beta-Hydrolases superfamily protein | chr3:8473833-8475655 FORWARD LENGTH=239
AT2G06050.1 (+2)	| Symbols: OPR3, DDE1 | oxophytodienoate-reductase 3 | chr2:2359240-

	2361971 REVERSE LENGTH=391
AT3G57260.1	| Symbols: BGL2, PR2, BG2, PR-2 | beta-1,3-glucanase 2 | chr3:21188709-21189822 REVERSE LENGTH=339
AT5G20960.1 (+1)	| Symbols: AAO1, AO1, ATAO, AT-AO1, AOalpha, AtAO1 | aldehyde oxidase 1 | chr5:7116783-7122338 FORWARD LENGTH=1368
AT5G39580.1	| Symbols: | Peroxidase superfamily protein | chr5:15847281-15849027 REVERSE LENGTH=319
AT2G20420.1	| Symbols: | ATP citrate lyase (ACL) family protein | chr2:8805574-8807858 FORWARD LENGTH=421
AT5G38900.1	| Symbols: | Thioredoxin superfamily protein | chr5:15573745-15574945 REVERSE LENGTH=217
AT3G16400.1 (+1)	| Symbols: ATMLP-470, NSP1, ATNSP1 | nitrile specifier protein 1 | chr3:5566516-5568330 FORWARD LENGTH=470
AT5G11920.1	| Symbols: AtcwINV6, cwINV6 | 6-&1-fructan exohydrolase | chr5:3839490-3842206 FORWARD LENGTH=550
AT4G01850.1 (+1)	| Symbols: SAM-2, MAT2, SAM2, AtSAM2 | S-adenosylmethionine synthetase 2 | chr4:796298-797479 REVERSE LENGTH=393
AT3G22200.1 (+1)	| Symbols: POP2, GABA-T, HER1 | Pyridoxal phosphate (PLP)-dependent transferases superfamily protein | chr3:7835286-7838863 FORWARD LENGTH=504
AT5G07440.1 (+1)	| Symbols: GDH2 | glutamate dehydrogenase 2 | chr5:2356153-2358012 FORWARD LENGTH=411
AT4G14630.1	| Symbols: GLP9 | germin-like protein 9 | chr4:8392920-8393680 FORWARD LENGTH=222
AT5G18470.1	| Symbols: | Curculin-like (mannose-binding) lectin family protein | chr5:6127952-6129193 FORWARD LENGTH=413
AT2G37040.1	| Symbols: PAL1, ATPAL1 | PHE ammonia lyase 1 | chr2:15557602-15560237 REVERSE LENGTH=725
AT3G19010.1	| Symbols: | 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein | chr3:6556306-6557862 REVERSE LENGTH=349
AT1G06290.1	| Symbols: ACX3, ATACX3 | acyl-CoA oxidase 3 | chr1:1922423-1926002 FORWARD LENGTH=675
AT1G76680.1 (+1)	| Symbols: OPR1, ATOPR1 | 12-oxophytodienoate reductase 1 | chr1:28776982-28778271 FORWARD LENGTH=372
AT2G30110.1	| Symbols: ATUBA1, MOS5, UBA1 | ubiquitin-activating enzyme 1 | chr2:12852632-12857369 REVERSE LENGTH=1080
AT1G13060.1	| Symbols: PBE1 | 20S proteasome beta subunit E1 | chr1:4452641-4454663 FORWARD LENGTH=274
AT1G27020.1	| Symbols: | unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G27030.1)
AT4G19810.1	| Symbols: | Glycosyl hydrolase family protein with chitinase insertion domain | chr4:10764151-10765753 REVERSE LENGTH=379

AT1G78460.1	| Symbols: | SOUL heme-binding family protein | chr1:29518547-29519296 REVERSE LENGTH=219
AT2G27860.1	| Symbols: AXS1 | UDP-D-apiose/UDP-D-xylose synthase 1 | chr2:11864684-11866843 REVERSE LENGTH=389
AT4G05180.1	| Symbols: PSBQ, PSBQ-2, PSII-Q | photosystem II subunit Q-2 | chr4:2672093-2673170 REVERSE LENGTH=230
AT5G35100.1	| Symbols: | Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein | chr5:13360459-13361377 REVERSE LENGTH=281
AT4G12490.1 (+1)	| Symbols: | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein | chr4:7409830-7410378 REVERSE LENGTH=182
AT2G22970.1 (+1)	| Symbols: SCPL11 | serine carboxypeptidase-like 11 | chr2:9774875-9778224 FORWARD LENGTH=433
AT4G29260.1	| Symbols: | HAD superfamily, subfamily IIIB acid phosphatase | chr4:14422310-14423409 REVERSE LENGTH=255
AT4G21280.1 (+1)	| Symbols: PSBQ, PSBQA, PSBQ-1 | photosystem II subunit QA | chr4:11334446-11335587 FORWARD LENGTH=223
AT2G26010.1 (+3)	| Symbols: PDF1.3 | plant defensin 1.3 | chr2:11087411-11087755 FORWARD LENGTH=80
AT4G16660.1	| Symbols: | heat shock protein 70 (Hsp 70) family protein | chr4:9377225-9381232 FORWARD LENGTH=867
AT1G50380.1	| Symbols: | Prolyl oligopeptidase family protein | chr1:18662480-18666185 FORWARD LENGTH=710
AT5G08300.1	| Symbols: | Succinyl-CoA ligase, alpha subunit | chr5:2667579-2669672 FORWARD LENGTH=347
AT1G26420.1	| Symbols: | FAD-binding Berberine family protein | chr1:9141715-9143304 REVERSE LENGTH=529
AT1G32960.1	| Symbols: ATSBT3.3, SBT3.3 | Subtilase family protein | chr1:11945351-11948429 FORWARD LENGTH=777
AT3G56650.1	| Symbols: | Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein | chr3:20984807-20985913 FORWARD LENGTH=262
AT5G12040.1	| Symbols: | Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase family protein | chr5:3885162-3887772 FORWARD LENGTH=369
AT2G21620.1 (+1)	| Symbols: RD2 | Adenine nucleotide alpha hydrolases-like superfamily protein | chr2:9248749-9249986 FORWARD LENGTH=187
AT3G16460.1 (+1)	| Symbols: | Mannose-binding lectin superfamily protein | chr3:5593029-5595522 FORWARD LENGTH=705
AT4G25100.1 (+4)	| Symbols: FSD1, ATFSD1 | Fe superoxide dismutase 1 | chr4:12884649-12886501 REVERSE LENGTH=212
AT5G41870.1	| Symbols: | Pectin lyase-like superfamily protein | chr5:16758817-16760490 REVERSE LENGTH=449

AT3G07390.1	| Symbols: AIR12 | auxin-responsive family protein | chr3:2365452-2366273 FORWARD LENGTH=273
AT1G30900.1	| Symbols: VSR6, VSR3;3, BP80-3;3 | VACUOLAR SORTING RECEPTOR 6 | chr1:10997275-11000543 FORWARD LENGTH=631
AT1G53850.1 (+1)	| Symbols: PAE1, ATPAE1 | 20S proteasome alpha subunit E1 | chr1:20104131-20105792 REVERSE LENGTH=237
AT1G15000.1	| Symbols: scpl50 | serine carboxypeptidase-like 50 | chr1:5168613-5169947 FORWARD LENGTH=444
AT4G16155.1	| Symbols: | dihydrolipoyl dehydrogenases | chr4:9153570-9157322 REVERSE LENGTH=630
AT5G20830.1 (+1)	| Symbols: SUS1, ASUS1, atsus1 | sucrose synthase 1 | chr5:7050599-7054032 REVERSE LENGTH=808
AT1G67550.1	| Symbols: URE | urease | chr1:25312842-25316911 FORWARD LENGTH=838
AT4G08770.1	| Symbols: Prx37 | Peroxidase superfamily protein | chr4:5598259-5600262 REVERSE LENGTH=346
AT1G08980.1	| Symbols: ATAMI1, AMI1, ATTOC64-I, TOC64-I | amidase 1 | chr1:2884455-2886430 FORWARD LENGTH=425
AT4G34480.1	| Symbols: | O-Glycosyl hydrolases family 17 protein | chr4:16481147-16483988 REVERSE LENGTH=504
AT1G11580.1	| Symbols: ATPMEPCRA, PMEPCRA | methylesterase PCR A | chr1:3888730-3890649 FORWARD LENGTH=557
AT3G03640.1	| Symbols: GLUC, BGLU25 | beta glucosidase 25 | chr3:881028-884028 FORWARD LENGTH=531
AT1G20510.1 (+1)	| Symbols: OPCL1 | OPC-8:0 CoA ligase1 | chr1:7103645-7105856 REVERSE LENGTH=546
AT5G54080.1 (+1)	| Symbols: HGO | homogentisate 1,2-dioxygenase | chr5:21945920-21948070 FORWARD LENGTH=461
AT5G05600.1	| Symbols: | 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein | chr5:1672266-1674602 FORWARD LENGTH=371
AT3G06770.2	| Symbols: | Pectin lyase-like superfamily protein | chr3:2135119-2137033 REVERSE LENGTH=446
AT5G55480.1	| Symbols: SVL1 | SHV3-like 1 | chr5:22474277-22477819 FORWARD LENGTH=766
AT3G61440.1 (+2)	| Symbols: ATCYSC1, ARATH;BSAS3;1, CYSC1 | cysteine synthase C1 | chr3:22735885-22737792 FORWARD LENGTH=368
AT2G02390.1 (+1)	| Symbols: ATGSTZ1, GST18, GSTZ1 | glutathione S-transferase zeta 1 | chr2:629015-630955 FORWARD LENGTH=221
AT3G27890.1	| Symbols: NQR | NADPH:quinone oxidoreductase | chr3:10350807-10351938 REVERSE LENGTH=196
AT5G60640.1 (+1)	| Symbols: ATPDIL1-4, PDI2, ATPDI2, PDIL1-4 | PDI-like 1-4 | chr5:24371141-

	24373993 REVERSE LENGTH=597
AT4G30910.1	| Symbols: | Cytosol aminopeptidase family protein | chr4:15042621-15045248 REVERSE LENGTH=581
AT3G10060.1	| Symbols: | FKBP-like peptidyl-prolyl cis-trans isomerase family protein | chr3:3102291-3103801 FORWARD LENGTH=230
AT3G11700.1	| Symbols: FLA18 | FASCICLIN-like arabinogalactan protein 18 precursor | chr3:3698992-3700971 FORWARD LENGTH=462
AT4G18250.1	| Symbols: | receptor serine/threonine kinase, putative | chr4:10087343-10091963 REVERSE LENGTH=853
AT1G02930.1 (+1)	| Symbols: ATGSTF6, GST1, ERD11, ATGSTF3, GSTF6, ATGST1 | glutathione S-transferase 6 | chr1:661363-662191 REVERSE LENGTH=208
AT5G64120.1	| Symbols: | Peroxidase superfamily protein | chr5:25659551-25660946 REVERSE LENGTH=328
AT1G30700.1	| Symbols: | FAD-binding Berberine family protein | chr1:10892623-10894437 FORWARD LENGTH=527
AT1G02920.1	| Symbols: ATGSTF7, GST11, ATGSTF8, GSTF7, ATGST11 | glutathione S-transferase 7 | chr1:658886-659705 REVERSE LENGTH=209
AT2G32240.1	| Symbols: | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to cadmium ion; LOCATED IN: plasma membrane;
AT4G16260.1	| Symbols: | Glycosyl hydrolase superfamily protein | chr4:9200180-9201441 REVERSE LENGTH=344
AT3G24503.1	| Symbols: ALDH2C4, ALDH1A, REF1 | aldehyde dehydrogenase 2C4 | chr3:8919732-8923029 REVERSE LENGTH=501
AT2G05710.1	| Symbols: ACO3 | aconitase 3 | chr2:2141591-2146350 FORWARD LENGTH=990
AT1G21750.1 (+1)	| Symbols: ATPDIL1-1, ATPDI5, PDI5, PDIL1-1 | PDI-like 1-1 | chr1:7645767-7648514 FORWARD LENGTH=501
AT5G03630.1	| Symbols: ATMDAR2 | Pyridine nucleotide-disulphide oxidoreductase family protein | chr5:922378-924616 REVERSE LENGTH=435
AT3G08590.1 (+1)	| Symbols: | Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent | chr3:2608683-2611237 REVERSE LENGTH=560
AT3G15730.1	| Symbols: PLDALPHA1, PLD | phospholipase D alpha 1 | chr3:5330835-5333474 FORWARD LENGTH=810
AT3G49120.1	| Symbols: ATPERX34, PERX34, PRXCB, ATPCB, PRX34 | peroxidase CB | chr3:18207819-18210041 FORWARD LENGTH=353
AT4G15530.1 (+4)	| Symbols: PPDK | pyruvate orthophosphate dikinase | chr4:8864828-8870727 REVERSE LENGTH=956
AT5G54500.1	| Symbols: FQR1 | flavodoxin-like quinone reductase 1 | chr5:22124674-22126256 FORWARD LENGTH=204
AT4G34890.1	| Symbols: ATXDH1, XDH1 | xanthine dehydrogenase 1 | chr4:16618736-16624983 REVERSE LENGTH=1361
AT4G20850.1	| Symbols: TPP2 | tripeptidyl peptidase ii | chr4:11160935-11169889

	REVERSE LENGTH=1380
AT3G14310.1	| Symbols: ATPME3, PME3 | pectin methylesterase 3 | chr3:4772214-4775095 REVERSE LENGTH=592
AT5G54960.1	| Symbols: PDC2 | pyruvate decarboxylase-2 | chr5:22310858-22312681 REVERSE LENGTH=607
AT3G16530.1	| Symbols: | Legume lectin family protein | chr3:5624586-5625416 REVERSE LENGTH=276
AT5G17530.1 (+2)	| Symbols: | phosphoglucosamine mutase family protein | chr5:5778168-5781863 FORWARD LENGTH=581
AT1G22840.1	| Symbols: CYTC-1, ATCYTC-A | CYTOCHROME C-1 | chr1:8079384-8080286 FORWARD LENGTH=114
AT2G41220.1	| Symbols: GLU2 | glutamate synthase 2 | chr2:17177934-17188388 FORWARD LENGTH=1629
AT4G22470.1	| Symbols: | protease inhibitor/seed storage/lipid transfer protein (LTP) family protein | chr4:11840316-11841443 REVERSE LENGTH=375
AT2G43560.1	| Symbols: | FKBP-like peptidyl-prolyl cis-trans isomerase family protein | chr2:18073995-18075385 REVERSE LENGTH=223
AT4G20860.1	| Symbols: | FAD-binding Berberine family protein | chr4:11172726-11174318 FORWARD LENGTH=530
AT2G43590.1	| Symbols: | Chitinase family protein | chr2:18081592-18082749 REVERSE LENGTH=264
AT3G04720.1	| Symbols: PR4, HEL, PR-4 | pathogenesis-related 4 | chr3:1285691-1286531 REVERSE LENGTH=212
AT4G16760.1	| Symbols: ACX1, ATACX1 | acyl-CoA oxidase 1 | chr4:9424930-9428689 REVERSE LENGTH=664
AT2G27190.1	| Symbols: PAP1, ATPAP1, PAP12, ATPAP12 | purple acid phosphatase 12 | chr2:11621400-11623438 REVERSE LENGTH=469
AT5G13120.1 (+1)	| Symbols: ATCYP20-2, CYP20-2 | cyclophilin 20-2 | chr5:4162714-4164720 REVERSE LENGTH=259
AT5G13420.1	| Symbols: | Aldolase-type TIM barrel family protein | chr5:4302080-4304212 REVERSE LENGTH=438
AT5G11670.1	| Symbols: ATNADP-ME2, NADP-ME2 | NADP-malic enzyme 2 | chr5:3754456-3758040 FORWARD LENGTH=588
AT2G14610.1	| Symbols: PR1, PR 1, ATPR1 | pathogenesis-related gene 1 | chr2:6241944-6242429 REVERSE LENGTH=161
AT3G15356.1	| Symbols: | Legume lectin family protein | chr3:5174603-5175418 REVERSE LENGTH=271
AT5G62530.1	| Symbols: ALDH12A1, ATP5CDH, P5CDH | aldehyde dehydrogenase 12A1 | chr5:25099768-25103159 REVERSE LENGTH=556
AT5G50250.1	| Symbols: CP31B | chloroplast RNA-binding protein 31B | chr5:20452677-20453965 REVERSE LENGTH=289
AT3G28940.1	| Symbols: | AIG2-like (avirulence induced gene) family protein |

	chr3:10968324-10969311 REVERSE LENGTH=169
AT2G29350.1 (+1)	| Symbols: SAG13 | senescence-associated gene 13 | chr2:12601036-12602222 FORWARD LENGTH=269
AT2G15220.1	| Symbols: | Plant basic secretory protein (BSP) family protein | chr2:6608689-6609366 FORWARD LENGTH=225
AT3G28930.1	| Symbols: AIG2 | AIG2-like (avirulence induced gene) family protein | chr3:10959890-10960728 REVERSE LENGTH=170
AT3G62250.1	| Symbols: UBQ5 | ubiquitin 5 | chr3:23037138-23037611 FORWARD LENGTH=157
AT1G33610.1	| Symbols: | Leucine-rich repeat (LRR) family protein | chr1:12188910-12190346 FORWARD LENGTH=478
AT4G26690.1	| Symbols: SHV3, MRH5, GPDL2 | PLC-like phosphodiesterase family protein | chr4:13456793-13459890 REVERSE LENGTH=759
AT2G43910.1	| Symbols: ATHOL1, HOL1 | HARMLESS TO OZONE LAYER 1 | chr2:18184658-18186951 REVERSE LENGTH=227
AT1G21680.1	| Symbols: | DPP6 N-terminal domain-like protein | chr1:7613028-7615148 FORWARD LENGTH=706
AT3G09440.1 (+1)	| Symbols: | Heat shock protein 70 (Hsp 70) family protein | chr3:2903434-2905632 REVERSE LENGTH=649
AT4G37870.1	| Symbols: PCK1, PEPCK | phosphoenolpyruvate carboxykinase 1 | chr4:17802974-17806332 REVERSE LENGTH=671
AT1G16470.1 (+1)	| Symbols: PAB1 | proteasome subunit PAB1 | chr1:5623122-5625439 FORWARD LENGTH=235
AT4G00570.1	| Symbols: NAD-ME2 | NAD-dependent malic enzyme 2 | chr4:242817-246522 REVERSE LENGTH=607
AT1G13900.1	| Symbols: | Purple acid phosphatases superfamily protein | chr1:4753494-4755554 REVERSE LENGTH=656
AT3G54960.1	| Symbols: ATPDIL1-3, PDI1, ATPDI1, PDIL1-3 | PDI-like 1-3 | chr3:20363514-20366822 REVERSE LENGTH=579
AT4G37530.1	| Symbols: | Peroxidase superfamily protein | chr4:17634786-17636082 FORWARD LENGTH=329
AT1G75040.1	| Symbols: PR5, PR-5 | pathogenesis-related gene 5 | chr1:28177754-28178731 FORWARD LENGTH=239
AT2G44920.2	| Symbols: | Tetratricopeptide repeat (TPR)-like superfamily protein | chr2:18524419-18526502 FORWARD LENGTH=224
AT4G20830.2	| Symbols: | FAD-binding Berberine family protein | chr4:11155486-11157108 FORWARD LENGTH=540
AT5G56590.1	| Symbols: | O-Glycosyl hydrolases family 17 protein | chr5:22907521-22909436 FORWARD LENGTH=506
AT1G53310.1 (+2)	| Symbols: ATPPC1, PEPC1, ATPEPC1, PPC1 | phosphoenolpyruvate carboxylase 1 | chr1:19884261-19888070 REVERSE LENGTH=967
AT2G27150.1 (+1)	| Symbols: AAO3, At-AO3, AOdelta, AtAAO3 | abscisic aldehyde oxidase 3 |

	chr2:11601952-11607014 FORWARD LENGTH=1332
AT5G17710.1 (+1)	| Symbols: EMB1241 | Co-chaperone GrpE family protein | chr5:5839560-5841639 REVERSE LENGTH=324
AT4G03200.1 (+1)	| Symbols: | catalytics | chr4:1408296-1412566 FORWARD LENGTH=818
AT3G51260.1 (+1)	| Symbols: PAD1 | 20S proteasome alpha subunit PAD1 | chr3:19031086-19032746 FORWARD LENGTH=250
AT4G35830.1	| Symbols: ACO1 | aconitase 1 | chr4:16973007-16977949 REVERSE LENGTH=898
AT5G37600.1	| Symbols: ATGSR1, GLN1;1, GSR 1, ATGLN1;1 | glutamine synthase clone R1 | chr5:14933574-14935656 REVERSE LENGTH=356
AT1G17290.1	| Symbols: AlaAT1 | alanine aminotransferas | chr1:5922771-5926093 FORWARD LENGTH=543
AT5G06860.1	| Symbols: PGIP1, ATPGIP1 | polygalacturonase inhibiting protein 1 | chr5:2132373-2133434 FORWARD LENGTH=330
AT1G07890.1 (+6)	| Symbols: APX1, MEE6, CS1, ATAPX1, ATAPX01 | ascorbate peroxidase 1 | chr1:2438005-2439435 FORWARD LENGTH=250
AT5G04590.1	| Symbols: SIR | sulfite reductase | chr5:1319404-1322298 FORWARD LENGTH=642
AT4G12730.1	| Symbols: FLA2 | FASCICLIN-like arabinogalactan 2 | chr4:7491598-7492809 REVERSE LENGTH=403
AT2G13560.1	| Symbols: NAD-ME1 | NAD-dependent malic enzyme 1 | chr2:5650089-5655103 FORWARD LENGTH=623
AT2G33120.1 (+1)	| Symbols: SAR1, VAMP722, ATVAMP722 | synaptobrevin-related protein 1 | chr2:14043785-14045337 REVERSE LENGTH=221
AT5G07470.1	| Symbols: PMSR3, ATMSRA3 | peptidemethionine sulfoxide reductase 3 | chr5:2362760-2364286 REVERSE LENGTH=202
AT5G45680.1	| Symbols: ATFKBP13, FKBP13 | FK506-binding protein 13 | chr5:18530894-18532128 FORWARD LENGTH=208
AT5G60360.1 (+2)	| Symbols: SAG2, AALP, ALP | aleurain-like protease | chr5:24280044-24282152 FORWARD LENGTH=358
AT1G27130.1	| Symbols: ATGSTU13, GST12, GSTU13 | glutathione S-transferase tau 13 | chr1:9425582-9426597 FORWARD LENGTH=227
AT5G40370.1 (+1)	| Symbols: | Glutaredoxin family protein | chr5:16147826-16149052 REVERSE LENGTH=111
AT1G79720.1	| Symbols: | Eukaryotic aspartyl protease family protein | chr1:29997259-29998951 REVERSE LENGTH=484
AT5G55450.1	| Symbols: | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein | chr5:22467560-22467874 FORWARD LENGTH=104

Accession # (Down-regulated)	Annotations (database)
AT1G70730.1 (+1)	| Symbols: PGM2 | Phosphoglucomutase/phosphomannomutase family protein | chr1:26669020-26672726 REVERSE LENGTH=585
AT1G72150.1	| Symbols: PATL1 | PATELLIN 1 | chr1:27148558-27150652 FORWARD LENGTH=573
AT5G56010.1	| Symbols: HSP81-3, Hsp81.3, AtHsp90-3, AtHsp90.3 | heat shock protein 81-3 | chr5:22681410-22683911 FORWARD LENGTH=699
AT5G60600.1 (+1)	| Symbols: GCPE, ISPG, CSB3, CLB4, HDS | 4-hydroxy-3-methylbut-2-enyl diphosphate synthase | chr5:24359447-24363274 FORWARD LENGTH=741
AT1G23190.1	| Symbols: PGM3 | Phosphoglucomutase/phosphomannomutase family protein | chr1:8219946-8224186 FORWARD LENGTH=583
ATCG00120.1	| Symbols: ATPA | ATP synthase subunit alpha | chrC:9938-11461 REVERSE LENGTH=507
AT1G64190.1	| Symbols: | 6-phosphogluconate dehydrogenase family protein | chr1:23825549-23827012 REVERSE LENGTH=487
AT3G20820.1	| Symbols: | Leucine-rich repeat (LRR) family protein | chr3:7280930-7282027 FORWARD LENGTH=365
AT5G65010.1 (+1)	| Symbols: ASN2 | asparagine synthetase 2 | chr5:25969224-25972278 FORWARD LENGTH=578
AT2G40840.1	| Symbols: DPE2 | disproportionating enzyme 2 | chr2:17045368-17050779 FORWARD LENGTH=955
AT1G10760.1	| Symbols: SEX1, SOP1, SOP, GWD1, GWD | Pyruvate phosphate dikinase, PEP/pyruvate binding domain | chr1:3581210-3590043 REVERSE LENGTH=1399
AT4G22670.1	| Symbols: AtHip1, HIP1, TPR11 | HSP70-interacting protein 1 | chr4:11918236-11920671 FORWARD LENGTH=441
AT5G19770.1 (+1)	| Symbols: TUA3 | tubulin alpha-3 | chr5:6682761-6684474 REVERSE LENGTH=450
AT2G36390.1	| Symbols: SBE2.1, BE3 | starch branching enzyme 2.1 | chr2:15264283-15269940 FORWARD LENGTH=858
AT3G22960.1	| Symbols: PKP1, PKP-ALPHA | Pyruvate kinase family protein | chr3:8139369-8141771 FORWARD LENGTH=596
AT1G28600.1	| Symbols: | GDSL-like Lipase/Acylhydrolase superfamily protein | chr1:10051228-10053073 REVERSE LENGTH=393
AT5G40450.1	| Symbols: | unknown protein; FUNCTIONS IN: molecular_function unknown;

	INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast, plasma membrane
AT5G19820.1	| Symbols: emb2734 | ARM repeat superfamily protein | chr5:6695731-6701247 REVERSE LENGTH=1116
AT4G39330.1 (+1)	| Symbols: ATCAD9, CAD9 | cinnamyl alcohol dehydrogenase 9 | chr4:18291268-18292772 FORWARD LENGTH=360
AT4G29840.1	| Symbols: MTO2, TS | Pyridoxal-5'-phosphate-dependent enzyme family protein | chr4:14599434-14601014 REVERSE LENGTH=526
AT5G35360.1 (+1)	| Symbols: CAC2 | acetyl Co-enzyme a carboxylase biotin carboxylase subunit | chr5:13584300-13588268 FORWARD LENGTH=537
AT5G45390.1	| Symbols: CLPP4, NCLPP4 | CLP protease P4 | chr5:18396351-18397586 FORWARD LENGTH=292
AT5G57560.1	| Symbols: TCH4, XTH22 | Xyloglucan endotransglucosylase/hydrolase family protein | chr5:23307296-23308235 REVERSE LENGTH=284
AT3G25230.1 (+1)	| Symbols: ROF1, ATFKBP62, FKBP62 | rotamase FKBP 1 | chr3:9188257-9191137 FORWARD LENGTH=551
AT2G46520.1	| Symbols: | cellular apoptosis susceptibility protein, putative / importin-alpha re-exporter, putative | chr2:19096867-19099785 FORWARD LENGTH=972
AT3G09820.1 (+1)	| Symbols: ADK1, ATADK1 | adenosine kinase 1 | chr3:3012122-3014624 FORWARD LENGTH=344
AT5G16390.1 (+1)	| Symbols: CAC1, CAC1A, BCCP, BCCP1 | chloroplastic acetylcoenzyme A carboxylase 1 | chr5:5361098-5363020 REVERSE LENGTH=280
AT3G05350.1	| Symbols: | Metallopeptidase M24 family protein | chr3:1527103-1533843 REVERSE LENGTH=710
AT3G18060.1	| Symbols: | transducin family protein / WD-40 repeat family protein | chr3:6183880-6186788 FORWARD LENGTH=609
AT3G26380.1	| Symbols: | Melibiase family protein | chr3:9660140-9663145 FORWARD LENGTH=647
AT1G20440.1	| Symbols: COR47, RD17, AtCOR47 | cold-regulated 47 | chr1:7084722-7085664 REVERSE LENGTH=265
AT1G49240.1 (+2)	| Symbols: ACT8 | actin 8 | chr1:18216539-18217947 FORWARD LENGTH=377
AT3G48420.1	| Symbols: | Haloacid dehalogenase-like hydrolase (HAD) superfamily protein | chr3:17929743-17931551 FORWARD LENGTH=319
AT1G22530.1	| Symbols: PATL2 | PATELLIN 2 | chr1:7955773-7958326 REVERSE LENGTH=683
AT1G29900.1	| Symbols: CARB | carbamoyl phosphate synthetase B | chr1:10468164-10471976 FORWARD LENGTH=1187
AT5G26830.1	| Symbols: | Threonyl-tRNA synthetase | chr5:9437351-9441568 FORWARD LENGTH=709
AT5G01410.1	| Symbols: PDX1, ATPDX1.3, RSR4, PDX1.3, ATPDX1 | Aldolase-type TIM barrel family protein | chr5:172576-173505 REVERSE LENGTH=309

AT2G41530.1	| Symbols: ATSFGH, SFGH | S-formylglutathione hydrolase | chr2:17323656-17325430 REVERSE LENGTH=284
AT5G03650.1	| Symbols: SBE2.2 | starch branching enzyme 2.2 | chr5:931924-937470 FORWARD LENGTH=805
AT2G30200.1 (+1)	| Symbols: | catalytics;transferases;[acyl-carrier-protein] S-malonyltransferases;binding | chr2:12883162-12885482 REVERSE LENGTH=393
AT5G64050.1	| Symbols: ATERS, OVA3, ERS | glutamate tRNA synthetase | chr5:25630196-25633099 REVERSE LENGTH=570
AT4G25370.1	| Symbols: | Double Clp-N motif protein | chr4:12972747-12974580 FORWARD LENGTH=238
AT1G20950.1	| Symbols: | Phosphofructokinase family protein | chr1:7297467-7301336 REVERSE LENGTH=614
AT2G35840.1 (+2)	| Symbols: | Sucrose-6F-phosphate phosphohydrolase family protein | chr2:15053952-15055776 FORWARD LENGTH=422
AT5G65730.1	| Symbols: XTH6 | xyloglucan endotransglucosylase/hydrolase 6 | chr5:26299080-26300290 FORWARD LENGTH=292
AT2G21660.1	| Symbols: ATGRP7, CCR2, GR-RBP7, GRP7 | cold, circadian rhythm, and rna binding 2 | chr2:9265477-9266316 REVERSE LENGTH=176
AT5G23120.1	| Symbols: HCF136 | photosystem II stability/assembly factor, chloroplast (HCF136) | chr5:7778154-7780463 FORWARD LENGTH=403
AT1G15140.1	| Symbols: | FAD/NAD(P)-binding oxidoreductase | chr1:5210403-5212137 REVERSE LENGTH=295
AT4G18810.1 (+1)	| Symbols: | NAD(P)-binding Rossmann-fold superfamily protein | chr4:10322622-10325735 REVERSE LENGTH=596
AT5G57550.1	| Symbols: XTR3, XTH25 | xyloglucan endotransglucosylase/hydrolase 25 | chr5:23305055-23306384 REVERSE LENGTH=284
AT5G53480.1	| Symbols: | ARM repeat superfamily protein | chr5:21714016-21716709 FORWARD LENGTH=870
AT5G41670.1 (+1)	| Symbols: | 6-phosphogluconate dehydrogenase family protein | chr5:16665647-16667110 REVERSE LENGTH=487
AT1G49750.1	| Symbols: | Leucine-rich repeat (LRR) family protein | chr1:18411177-18412779 REVERSE LENGTH=494
AT2G34810.1	| Symbols: | FAD-binding Berberine family protein | chr2:14685292-14686914 FORWARD LENGTH=540
AT4G09670.1	| Symbols: | Oxidoreductase family protein | chr4:6107382-6109049 REVERSE LENGTH=362
AT2G27920.1	| Symbols: SCPL51 | serine carboxypeptidase-like 51 | chr2:11885777-11889043 REVERSE LENGTH=461
AT4G27450.1	| Symbols: | Aluminium induced protein with YGL and LRDR motifs |

	chr4:13727665-13728683 REVERSE LENGTH=250
AT1G27400.1 (+1)	| Symbols: | Ribosomal protein L22p/L17e family protein | chr1:9515230-9516725 FORWARD LENGTH=176
AT1G18270.1 (+2)	| Symbols: | ketose-bisphosphate aldolase class-II family protein | chr1:6283634-6293772 REVERSE LENGTH=1373
AT1G11750.1 (+1)	| Symbols: CLPP6, NCLPP1, NCLPP6 | CLP protease proteolytic subunit 6 | chr1:3967609-3969535 FORWARD LENGTH=271
AT1G12410.1	| Symbols: CLPR2, NCLPP2, CLP2 | CLP protease proteolytic subunit 2 | chr1:4223099-4224954 FORWARD LENGTH=279
AT5G55730.1 (+1)	| Symbols: FLA1 | FASCICLIN-like arabinogalactan 1 | chr5:22558375-22560392 REVERSE LENGTH=424
AT4G23500.1	| Symbols: | Pectin lyase-like superfamily protein | chr4:12264640-12267074 FORWARD LENGTH=495
AT2G25080.1	| Symbols: ATGPX1, GPX1 | glutathione peroxidase 1 | chr2:10668134-10669828 FORWARD LENGTH=236
AT1G06570.1 (+1)	| Symbols: PDS1, HPD | phytoene desaturation 1 | chr1:2012015-2013543 REVERSE LENGTH=473
AT2G20890.1	| Symbols: PSB29, THF1 | photosystem II reaction center PSB29 protein | chr2:8987783-8989185 FORWARD LENGTH=300
AT1G13930.1 (+2)	| Symbols: | Involved in response to salt stress. Knockout mutants are hypersensitive to salt stress. | chr1:4761091-4761558 FORWARD LENGTH=155
AT2G06850.1	| Symbols: EXGT-A1, EXT, XTH4 | xyloglucan endotransglucosylase/hydrolase 4 | chr2:2763619-2765490 FORWARD LENGTH=296
AT2G44160.1	| Symbols: MTHFR2 | methylenetetrahydrofolate reductase 2 | chr2:18262301-18265185 FORWARD LENGTH=594
AT5G64260.1	| Symbols: EXL2 | EXORDIUM like 2 | chr5:25703980-25704897 FORWARD LENGTH=305
AT1G05560.1	| Symbols: UGT1, UGT75B1 | UDP-glucosyltransferase 75B1 | chr1:1645674-1647083 REVERSE LENGTH=469
AT1G03230.1	| Symbols: | Eukaryotic aspartyl protease family protein | chr1:790110-791414 FORWARD LENGTH=434
AT1G12230.1 (+1)	| Symbols: | Aldolase superfamily protein | chr1:4148050-4150708 FORWARD LENGTH=405
AT5G58250.1	| Symbols: | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: thylakoid, chloroplast;
AT4G01130.1	| Symbols: | GDSL-like Lipase/Acylhydrolase superfamily protein | chr4:485868-488007 FORWARD LENGTH=382
AT1G09210.1	| Symbols: CRT1b, AtCRT1b | calreticulin 1b | chr1:2973217-2976655 REVERSE LENGTH=424

AT1G13280.1	| Symbols: AOC4 | allene oxide cyclase 4 | chr1:4547624-4548552 FORWARD LENGTH=254
AT2G02560.1 (+1)	| Symbols: CAND1, ATCAND1, ETA2, TIP120, HVE | cullin-associated and neddylation dissociated | chr2:690345-697342 FORWARD LENGTH=1219
AT2G34480.1	| Symbols: | Ribosomal protein L18ae/LX family protein | chr2:14532916-14534161 REVERSE LENGTH=178
AT1G69830.1	| Symbols: ATAMY3, AMY3 | alpha-amylase-like 3 | chr1:26288518-26293003 REVERSE LENGTH=887
AT3G12145.1	| Symbols: FLR1, FLOR1 | Leucine-rich repeat (LRR) family protein | chr3:3874764-3876075 REVERSE LENGTH=325
AT3G12390.1	| Symbols: | Nascent polypeptide-associated complex (NAC), alpha subunit family protein | chr3:3942344-3943595 FORWARD LENGTH=203
AT3G14210.1	| Symbols: ESM1 | epithiospecifier modifier 1 | chr3:4729886-4731562 FORWARD LENGTH=392
AT5G59320.1	| Symbols: LTP3 | lipid transfer protein 3 | chr5:23929051-23929492 FORWARD LENGTH=115

Involvement of the Electrophilic Isothiocyanate Sulforaphane in Arabidopsis Local Defense Responses[1]

Mats X. Andersson[2], Anders K. Nilsson[2], Oskar N. Johansson, Gülin Boztaş, Lisa E. Adolfsson, Francesco Pinosa, Christel Garcia Petit, Henrik Aronsson, David Mackey, Mahmut Tör, Mats Hamberg, and Mats Ellerström*

Department of Biological and Environmental Sciences, University of Gothenburg, SE–405 30 Gothenburg, Sweden (M.X.A., A.K.N., O.N.J., L.E.A., F.P., C.G.P., H.A., M.E.); National Pollen and Aerobiology Research Unit, Institute of Science and the Environment, University of Worcester, Worcester WR2 6AJ, United Kingdom (G.B., M.T.); Departments of Horticulture and Crop Science and Molecular Genetics, Ohio State University, Columbus, Ohio 43210 (D.M.); and Division of Chemistry II, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE–17177 Stockholm, Sweden (M.H.)

Plants defend themselves against microbial pathogens through a range of highly sophisticated and integrated molecular systems. Recognition of pathogen-secreted effector proteins often triggers the hypersensitive response (HR), a complex multicellular defense reaction where programmed cell death of cells surrounding the primary site of infection is a prominent feature. Even though the HR was described almost a century ago, cell-to-cell factors acting at the local level generating the full defense reaction have remained obscure. In this study, we sought to identify diffusible molecules produced during the HR that could induce cell death in naive tissue. We found that 4-methylsulfinylbutyl isothiocyanate (sulforaphane) is released by Arabidopsis (*Arabidopsis thaliana*) leaf tissue undergoing the HR and that this compound induces cell death as well as primes defense in naive tissue. Two different mutants impaired in the pathogen-induced accumulation of sulforaphane displayed attenuated programmed cell death upon bacterial and oomycete effector recognition as well as decreased resistance to several isolates of the plant pathogen *Hyaloperonospora arabidopsidis*. Treatment with sulforaphane provided protection against a virulent *H. arabidopsidis* isolate. Glucosinolate breakdown products are recognized as antifeeding compounds toward insects and recently also as intracellular signaling and bacteriostatic molecules in Arabidopsis. The data presented here indicate that these compounds also trigger local defense responses in Arabidopsis tissue.

Plants are constantly challenged by pathogenic microorganisms and have developed several detection and defense systems to protect themselves against the invaders. Preformed defenses include the waxy cuticle, thick cell walls, and antimicrobial compounds. After recognition of microbe-associated patterns, defense responses are induced, which include the fortification of cell walls and the production of phytoalexins (Monaghan and Zipfel, 2012). Overcoming the preformed and induced defenses of the plant hosts requires adaptation by the pathogen. Pathogenic bacteria use type III secretion to inject proteins (so-called effectors) into the host cytosol in order to overcome plant defense responses (Bent and Mackey, 2007). In turn, plants have developed systems to recognize the pathogenic effectors and mount defense. Recognition of type III effectors by plant resistance (R) proteins induces robust defense responses that frequently include the hypersensitive response (HR).

The HR is a complex defense reaction characterized by the induction of programmed cell death (PCD) in the local host tissue as well as the activation of other defense responses in both local and systemic tissue (Mur et al., 2008; Shah, 2009). Oomycetes and true fungi also secrete proteinaceous effectors that can be recognized by host R proteins (Coates and Beynon, 2010; Hückelhoven and Panstruga, 2011; Feng and Zhou, 2012). The lesions formed during the HR vary in size between different host-pathogen pairs; however, a lesion induced at one or a few cells can spread to surrounding cells (Mur et al., 2008). Since pathogens inducing HR typically fail to proliferate, the first infected cell likely releases a compound that promotes PCD in surrounding cells. This is especially clear in models with oomycete and fungal pathogens, where the localization of the pathogen and the spread of cell death around the infection site can be clearly visualized (Mur et al., 2008; Coates

and Beynon, 2010). Trailing necrosis is an incomplete resistance phenotype characterized by cell death that trails, but fails to contain, the filamentous growth of the pathogen. One explanation for trailing necrosis is a failure of infected cells to produce a putative mobile defense signal required to enhance defense in neighboring cells. Farther from the site of PCD, other defense pathways are activated and systemic tissue is primed for defense.

The hunt for systemically acting compounds has been intense, and several candidates for this signal have been presented (Dempsey and Klessig, 2012). In contrast, even though the phenomenon of HR as a defense reaction was described almost a century ago (Stakman, 1915; Mur et al., 2008), compounds acting on the local tissue scale of the HR have attracted little attention. We set out to find substances released from cells undergoing the HR that could induce cell death in naive tissue. We report that leaf tissue of the model plant Arabidopsis (*Arabidopsis thaliana*) releases the reactive electrophilic compound sulforaphane after bacterial effector recognition. Mutants affected in sulforaphane production as well as other glucosinolate breakdown products showed delayed or reduced cell death after the recognition of pathogenic effectors and decreased resistance to an oomycete pathogen. Moreover, pretreatment of plants with sulforaphane enhanced resistance against a virulent oomycete isolate. Thus, we interpret this as that sulforaphane and likely similar compounds might both possess direct antimicrobial properties and, through a cytotoxic mechanism, act directly on plant cells to trigger defense responses.

RESULTS

Isolation and Identification of a Cell Death-Inducing Compound Released from Arabidopsis Tissue Undergoing the HR

To investigate if cells undergoing the HR release chemical signals to induce or promote cell death in noninfected tissue, a transgenic system to scale up the HR was used. The transgenic system consists of Arabidopsis harboring a dexamethasone (DEX)-inducible copy of the *Pseudomonas syringae* type III effector Avirulent Resistance to *Pseudomonas Maculicola1* protein (AvrRpm1; Mackey et al., 2002, 2003; Andersson et al., 2006). Leaf tissue expressing the bacterial effector AvrRpm1 (DEX:AvrRpm1/Columbia-0 [Col-0]) was incubated in water with DEX. An isogenic line in the *rpm1-3* background (a protein null for RPM1 and thus unable to recognize AvrRpm1) and the untransformed wild type (Col-0) were used as controls. The bathing solution was filtered, and the filtrate was run through a C18 solid-phase extraction cartridge to capture small molecules (Supplemental Fig. S1). The obtained fraction was dissolved in water and infiltrated into nontransgenic Arabidopsis leaves. The fraction obtained from DEX:AvrRpm1/Col-0 caused cell death when infiltrated into Arabidopsis wild-type leaves (Fig. 1A).

In contrast, the fraction from DEX:AvrRpm1/*rpm1-3* and from untransformed wild-type material had no apparent effect on plant tissue (Fig. 1B). Thus, it is apparent that Arabidopsis tissue undergoing the HR releases one or several soluble compounds that can induce cell death in naive leaf tissue.

The active material was further fractionated by HPLC (Fig. 1C), and a fraction with the ability to induce cell death evident as visual lesions and by trypan blue staining was obtained. Gas chromatography-mass spectrometry (GC-MS) analysis of this active fraction revealed a single peak (Fig. 1D). Comparison with publicly available mass spectra libraries identified the compound as 4-methylsulfinylbutyl isothiocyanate, a compound commonly known by its trivial name, sulforaphane. The mass spectrum (Fig. 1E) showed the expected molecular ion at mass-to-charge ratio (*m/z*) 177 and a prominent peak at *m/z* 160 caused by the loss of the sulfoxide oxygen plus one hydrogen. The mass and UV absorption spectra (Fig. 1F) were identical to those reported for purified sulforaphane (Tierens et al., 2001). A sulforaphane standard coeluted with the active fraction in the semipreparative HPLC system. Thus, the purified compound isolated from Arabidopsis leaf tissue undergoing the HR was unambiguously identified as the isothiocyanate sulforaphane (Fig. 1G).

Sulforaphane Released by Arabidopsis Tissue Undergoing the HR Can Be Degraded by *P. syringae* SURVIVAL IN ARABIDOPSIS EXTRACTS Genes

Quantification of sulforaphane released into the bathing solution from leaf discs of DEX:AvrRpm1/Col-0 plants revealed that during 12 h post induction (hpi), about 200 nmol of sulforaphane was released per 1 g of tissue (Fig. 2A). A near-peak concentration was reached after 6 h and remained high for at least 24 h. The corresponding transgenic line in the *rpm1-3* background released less than 5 nmol of sulforaphane per 1 g at 12 h after induction. The release of sulforaphane from wild-type and *rpm1-3* leaf tissue infiltrated with *P. syringae* expressing AvrRpm1 (DC3000:AvrRpm1) was similar to that of plants harboring DEX-inducible AvrRpm1, except that the concentration of released sulforaphane decreased dramatically between 12 and 24 hpi (Fig. 2B).

It has been reported that sulforaphane formation is induced in Arabidopsis in response to several nonhost bacterial pathogens (Fan et al., 2011). Furthermore, *P. syringae* strains adapted to Arabidopsis were shown to harbor so-called SURVIVAL IN ARABIDOPSIS EXTRACTS (SAX) genes, which enabled the bacteria to detoxify the sulforaphane produced by the host. This spurred us to test if the presence of SAX genes in the bacteria could explain the difference in the amount of sulforaphane present in the DEX-inducible and the bacteria-infiltrated material at the 24-h time point. To this end, Δsax mutant *P. syringae* (Fan et al., 2011) was transformed with a vector encoding the AvrRpm1 effector.
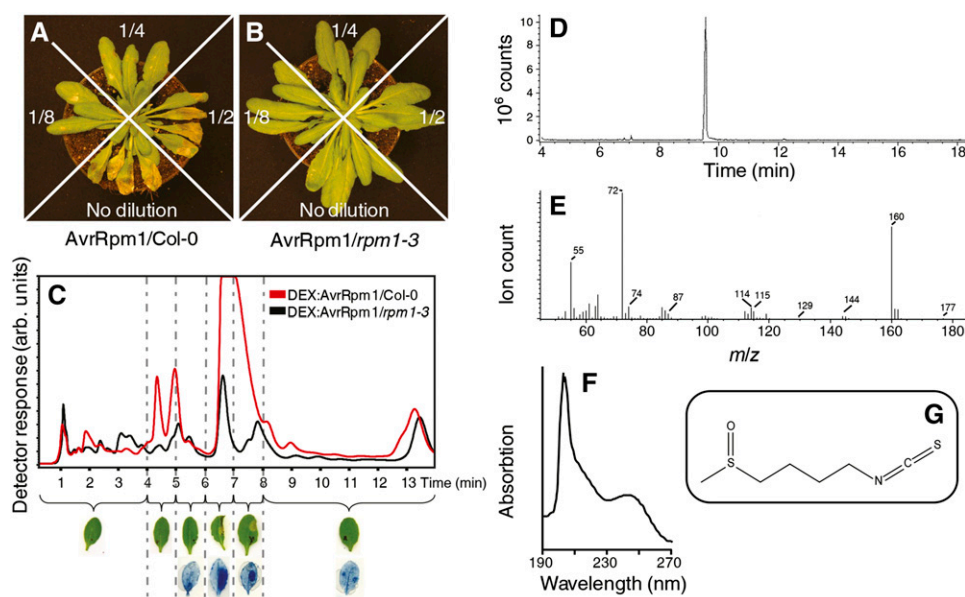
Figure 1. Induction of cell death by an aqueous extract from Arabidopsis tissue undergoing the HR and identification of sulforaphane. A and B, Transgenic Arabidopsis plants expressing the bacterial *P. syringae* effector AvrRpm1 (DEX:AvrRpm1/Col-0 and DEX:AvrRpm1/*rpm1-3*) were incubated in water with the inducer DEX. Small molecules recovered from the bathing solution of the Col-0 (A) and *rpm1-3* (B) plants were infiltrated into wild-type (nontransgenic) plants at the indicated dilutions. C, The material obtained was further analyzed by HPLC. Fractions were collected, dried, dissolved in water, and infiltrated into wild-type leaves. Visible effects and Trypan Blue staining of leaves receiving the fractions from the DEX:AvrRpm1/Col-0 extract are shown at bottom. D, The HPLC-purified fraction from the DEX:AvrRpm1/Col-0 fraction was subjected to GC-MS with electron-impact ionization. E, Mass spectrum for the major peak. F, The fraction was also dissolved in methanol, and the UV absorption spectrum was recorded. G, Structure of the identified compound, sulforaphane. The experiments depicted in A to C were performed twice with identical results.

The resulting Δ*sax*:AvrRpm1 bacteria were infiltrated into wild-type (Col-0) leaf discs and incubated in water, and the amount of sulforaphane released to the bathing solution was measured at different time points (Fig. 2C). Again, DC3000:AvrRpm1 inoculation caused a rapid induction of sulforaphane, which decreased back to low levels at 24 hpi. In contrast, inoculation with the Δ*sax*:AvrRpm1 strain resulted in the accumulation of sulforaphane, which peaked at 12 hpi and remained high at 24 hpi.

Sulforaphane and Other Isothiocyanates Induce Cell Death in Naive Leaf Tissue

Leaf infiltration experiments were conducted in order to determine the concentration of sulforaphane needed to induce cell death. As seen in Figure 3A, visible cell death occurred after infiltration of 0.5 and 1 mM sulforaphane. In some experiments, lower concentrations were sufficient, but 0.5 mM always caused large visible lesions. This concentration also produced large necrotic lesions on leaves of several other plant species; broad bean (*Vicia faba*) and sunflower (*Helianthus annuus*) are shown in Figure 3, B and C.

For a more precise measurement of cell death, electrolyte leakage after infiltration with sulforaphane was monitored. As seen in Figure 3D, 100 μM sulforaphane was sufficient to cause significant electrolyte leakage, and 200 μM caused a large loss of cellular electrolytes 48 h after infiltration. Infiltration with benzyl, butyl, and isopropyl isothiocyanate at 100, 200, and 400 μM was also tested for their ability to induce cell death. The strongest effect was observed with benzyl isothiocyanate, which at 100 μM already induced robust electrolyte leakage by 6 h. An effect of butyl isothiocyanate was apparent but weaker than that of benzyl isothiocyanate and sulforaphane, whereas isopropyl isothiocyanate seemed to lack any effect over mock infiltration.

It is believed that the major mode of action of sulforaphane in mammalian cells is to decrease the cellular glutathione pool (Valgimigli and Iori, 2009). This would be consistent with its reactive electrophile properties, and it is well known that plant defense triggers oxidative stress that contributes to cell death (Overmyer et al., 2003; Torres, 2010). Total and oxidized glutathione, therefore, were measured in leaf discs 30 min after infiltration with a sulforaphane solution (Fig. 4). Treatment with sulforaphane led to a severe depletion of the total glutathione pool and, consequently, to a large calculated increase in the redox potential of the glutathione pool.
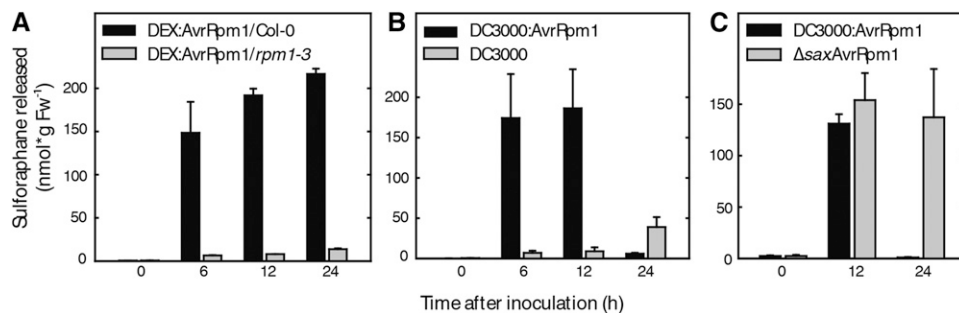
Figure 2. Release of sulforaphane during effector-induced HR. A, Transgenic Arabidopsis plants expressing the bacterial *P. syringae* effector AvrRpm1 (DEX:AvrRpm1/Col-0 and DEX:AvrRpm1/*rpm1-3*) were incubated in water with DEX. B and C, Wild type Col-0 Arabidopsis leaf discs were infiltrated with wild-type *P. syringae* pv *tomato* DC3000 (B) or the Δ*sax* mutant (C) expressing the effector AvrRpm1 (OD_{600} = 0.1) and incubated in water. At the indicated times, the discs were removed and the amount of sulforaphane in the bathing solution was analyzed. Average and range values for duplicate samples are shown. The experiments in A and B were performed three times with similar results, and the experiment in C was repeated twice with similar results. Fw, Fresh weight.

Mutants with a Reduced Capacity of Sulforaphane Production Display Decreased Cell Death Response

There are currently no specific biosynthesis mutants available that lack capacity to produce sulforaphane (Sønderby et al., 2010b). However, the two myeloblastosis (MYB) family transcription factors MYB28 and MYB29 have been found to control the production of aliphatic glucosinolate precursors in Arabidopsis (Beekwilder et al., 2008; Sønderby et al., 2010a). The *myb28 myb29* double mutant has severely reduced levels of the aliphatic

glucosinolates while maintaining wild-type levels of other glucosinolates. The myrosinase double mutant *thioglucoside glucohydrolase1* (*tgg1*) *tgg2* lacks the thioglycosidases cleaving the inactive glucosinolates and, therefore, is unable to accumulate sulforaphane and other glucosinolate breakdown products after wounding (Barth and Jander, 2006). The single mutants *myb28* and *myb29* and the double mutants *myb28 myb29* and *tgg1 tgg2* were tested for sulforaphane production after inoculation with *P. syringae* DC3000:AvrRpm1 (Fig. 5A; Supplemental

Figure 3. Infiltration of pure sulforaphane causes cell death in leaf. A to C, Sulforaphane was suspended in water to the indicated concentrations (A) or 1 mM (B and C) and syringe infiltrated into wild-type Col-0 Arabidopsis (A), broad bean (B), and sunflower (C) leaves. The left side of the leaf was infiltrated with sulforaphane, and the right side was mock infiltrated with deionized water. The leaves were detached and photographed after 24 h (A). D, Leaves were syringe infiltrated with the indicated isothiocyanates (ITC); leaf discs were punched out, washed, and incubated in water for the indicated times; and the conductance of the bathing solution was measured. Average and range values for duplicate samples are shown. Asterisks indicate statistical significance compared with mock treatment at the indicated times (one-way ANOVA, $P <$ 0.05). The experiments were performed twice with similar results.
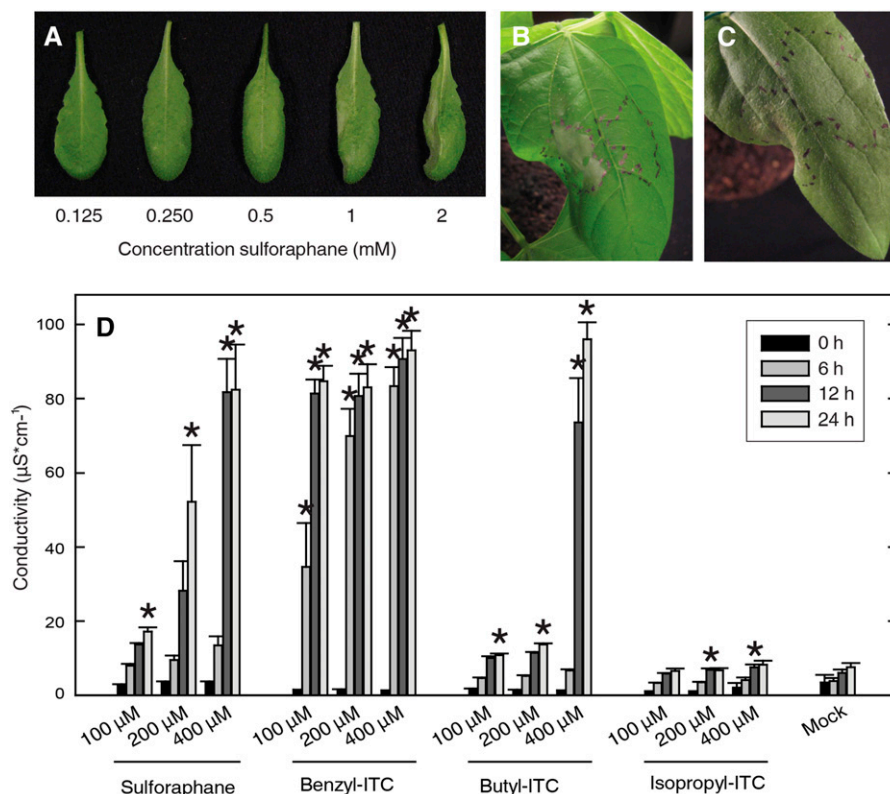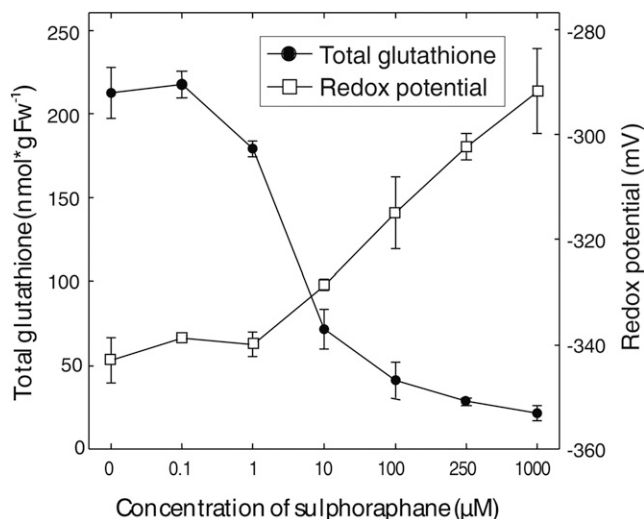
Figure 4. Infiltration of pure sulforaphane causes oxidation of the cellular glutathione pool. Leaf discs from wild-type Arabidopsis were vacuum infiltrated with sulforaphane suspended in water at the indicated concentrations. Reduced and oxidized glutathione contents were measured 30 min after infiltration, and the redox potential was calculated. Average and range values for triplicate samples are shown. The experiment was performed twice with similar results. Fw, Fresh weight.

Fig. S2). Sulforaphane release in both double mutants was reduced by more than 98% compared with wild-type Col-0.

We hypothesized that, if sulforaphane acts as a mobile signal promoting PCD in cells adjacent to the infection site, then the double mutant would show diminished death of cells neighboring those in which type III effectors are activating an R protein. In accordance with this, electrolyte leakage was markedly decreased in the double mutants infiltrated with

DC3000:AvrRpm1 at low titer (Fig. 5B). We next tested if this was reflected in the ability to restrict reproduction of the pathogen in the plant tissue. Leaves of the mutants and wild-type plants were infiltrated with *P. syringae* DC3000:AvrRpm1, and the growth of the bacteria was measured 3 d later (Supplemental Fig. S3). The two double mutants demonstrated no change in resistance to *P. syringae* expressing AvrRpm1.

To further explore the role of sulforaphane in the plant's local resistance responses, we used the oomycete *Hyaloperonospora arabidopsidis* (*Hpa*). Arabidopsis wild type Col-0 carries the *Recognition of Peronospora Parasitica2* (*RPP2*) and *RPP4* genes that mediate recognition and trigger the HR in response to the isolates *Hpa* Cala2 and *Hpa* Emwa1, respectively (van der Biezen et al., 2002; Sinapidou et al., 2004). The wild type Col-0 and the *myb28 myb29* and *tgg1 tgg2* double mutants were inoculated with the isolate *Hpa* Cala2, and the extent of cell death at the infection sites was scored after trypan blue staining (Fig. 6A). In wild-type Col-0, about 90% of the interaction sites resulted in quick localized cell death and no visible growth of the pathogen (Fig. 6B), whereas about 10% of the interaction sites demonstrated some degree of trailing necrosis and growth of hyphae (Fig. 6C) within the given time period. In the two double mutants, the situation was reversed, with 93% and 73% of the interaction sites resulting in trailing necrosis for *myb28 myb29* and *tgg1 tgg2*, respectively. In a few instances, free hyphae, completely outgrowing the trailing necrosis, could be seen on the *myb28 myb29* double mutant. At an early time point of the infection, 36 hpi, the two mutants demonstrated less cell death than the wild type (Fig. 6D), and the number of dead cells per interaction site was decreased to less than half of the wild type in the two double mutants. This may suggest that slower cell death at early stages of infection results
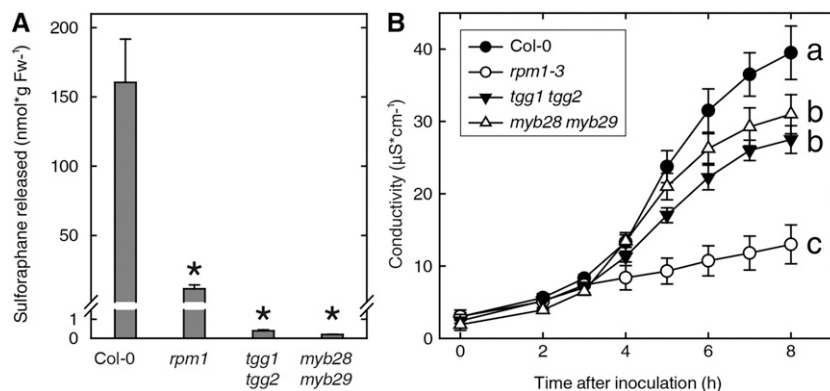

Figure 5. Compromised PCD in sulforaphane-deficient mutants. A, Leaf discs from Arabidopsis wild-type and mutant lines were infiltrated with *P. syringae* DC3000:AvrRpm1, and the amount of sulforaphane released into the bathing solution was determined at 6 hpi. Mean and range values for duplicate samples are shown. Asterisks indicate statistical significance compared with Col-0 (one-way ANOVA, $P < 0.05$). Fw, Fresh weight. B, Electrolyte leakage was measured at the indicated time points for the wild type (black circles), double mutant *myb28 myb29* (white triangles), double mutant *tgg1 tgg2* (black triangles), and *rpm1-3* (white circles). Mean ± SD of six replicate samples are shown. Letters a to c indicate statistically significant groups (one-way ANOVA with Tukey's posthoc test, $P < 0.05$). The experiments were performed twice with similar results.

Figure 6. Reduced resistance to a biotrophic oomycete in sulforaphane-deficient mutants. The indicated lines were inoculated at the cotyledon stage with *Hpa* conidia of the isolates Cala2 (A–D) or Emwa1 (E). The cotyledons were stained with Trypan Blue 7 (A–C) or 2 (D) dpi, and the extent of cell death was determined or sporophores were counted at 7 dpi (E). Rapid cell death in the wild type is shown in B, and trailing necrosis is shown in C. Average and range values for three replicate experiments, each including 200 interaction sites, are shown in A and D. Average and sd for 15 cotyledons are shown in E. Letters a to c indicate statistically significant groups (one-way ANOVA with Tukey's posthoc test, $P < 0.05$). The experiments were performed three times with similar results. Bars in B and C = 0.5 mm.
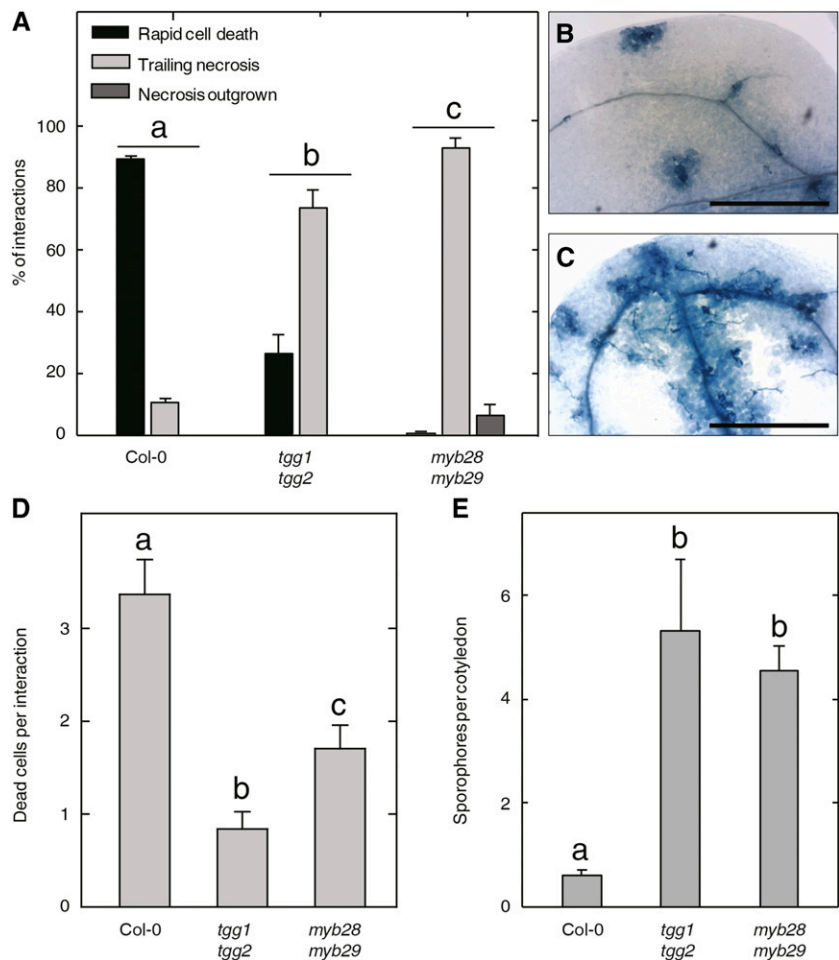
in less restriction of pathogen growth, and consequently, the growing hyphae are followed by delayed cell death. The isolate *Hpa* Emwa1 also showed increased virulence on the double mutants, as evidenced by an approximately 6-fold increase in the number of sporophores produced 7 d post induction (dpi) of cotyledons (Fig. 6E).

Sulforaphane Pretreatment of Arabidopsis Induces Defense against Virulent Pathogens

To test whether sulforaphane can activate defense and provide increased immunity to naive plant tissue, Arabidopsis was treated with sulforaphane before inoculation with virulent *Hpa* Cala2 (Fig. 7). In this case, the accession Landsberg *erecta* and the *enhanced disease resistance1* (*eds1*) mutant in the Wassilewskija-0 background were used as controls, as these lines are highly susceptible to the Cala2 isolate (Parker et al., 1996; Bailey et al., 2011). Sporulation at 4 dpi was markedly decreased when plants had been pretreated with sulforaphane before inoculation with *Hpa* Cala2. Treatment with sulforaphane 24 hpi with Cala2 also led to decreased sporulation at 4 dpi, but the effect was

significantly smaller. However, the effect of the sulforaphane was indistinguishable from that of the untreated controls at 7 dpi (Supplemental Fig. S4).

DISCUSSION

The HR was described over 100 years ago (Stakman, 1915; Mur et al., 2008). Eighty years later, the basic genetics of effector-triggered immunity started to be unraveled, and the past decade has seen an explosive increase in the understanding of the molecular mechanisms behind plant defense responses to pathogens. Many pathogenic effectors and their cognate plant R proteins have been identified, and several intracellular signaling components have been discovered (Desveaux et al., 2006; Torres et al., 2006; Bent and Mackey, 2007). Much effort has gone into the elucidation of phloem-mobile compounds that mediate the induction of systemic acquired resistance (Dempsey and Klessig, 2012). In contrast, comparatively little effort has been spent on compounds active at the local tissue level in plant defenses. The dead and dying cells of HR lesions might appear uniform. However, some cells directly see the bacteria and recognize the effector, whereas adjacent
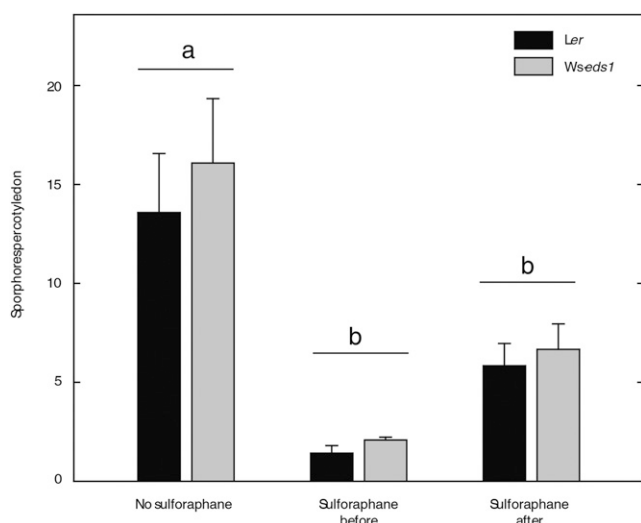
Figure 7. Sulforaphane treatment of plants provides increased resistance against pathogens. Seedlings of wild-type Landsberg *erecta* (L*er*) or *eds1* in the Wassilewskija (Ws) background were sprayed with 200 μM sulforaphane 24 h before or after, as indicated, inoculation with *Hpa* isolate Cala2, and the resulting sporulation was counted at 4 dpi. Average numbers of spores and range values for three replicate samples are shown. Letters a to c indicate statistically significant groups (one-way ANOVA with Tukey's posthoc test, *P* < 0.05). The experiment was repeated twice with identical results.

cells only sense the response of their immediate neighbors. Given that the HR lesion spreads to a size considerably larger than that reasonably reached by the pathogens in terms of single bacterial cells, germinated oomycetes, or fungal spores, it seems likely that signals produced by the initially responding cells contribute to propagating cell death. Importantly, cell death must also stop from spreading, and it has been described that cells adjacent to an HR site become more resistant to further pathogen attacks (Ross, 1961).

In this study, we set out to identify cell death-promoting compounds released by plant cells undergoing effector-triggered immunity. To this end, we used bioassay-guided fractionation of a diffusate obtained from transgenic Arabidopsis leaf tissue where the bacterial effector AvrRpm1 was expressed in planta. A small-molecule extract that caused necrotic lesions when infiltrated into naive wild-type tissue was obtained from a bathing solution in which the transgenic tissue had been incubated. The extract was further fractionated, and the single most active compound could be unambiguously identified as the isothiocyanate sulforaphane. We could also demonstrate that Arabidopsis wild-type tissues infiltrated with *P. syringae* expressing AvrRpm1 released quantities of sulforaphane sufficient to induce cell death in naive leaf tissue (Figs. 2B and 3D).

Sulforaphane belongs to a large and structurally diverse family of plant defensive compounds derived from glucosinolate precursors (Fahey et al., 2001). When these are cleaved by thioglucosidases known as myrosinases, the unstable intermediates are converted

to nitriles, isothiocyanates, and thiocyanates. These compounds are found in many plant families, especially in the Brassicale order, but are far from universal in the plant kingdom (Fahey et al., 2001). In, fact, the presence of glucosinolates in plant species outside the Brassicale is a rare occurrence. The plant species that do contain glucosinolates display a large degree of variation in the composition of the side chain of their glucosinolates (Fahey et al., 2001; Agerbirk and Olsen, 2012). Sulforaphane is thus restricted to a few plant species, and the amount of both the precursor glucoraphanine and sulforaphane varies between Arabidopsis accessions (Kliebenstein et al., 2001). Thus, sulforaphane production in response to HR must be a very limited occurrence in the plant kingdom, and any conservation should be sought in function rather than in structure.

The glucosinolates were previously considered as primarily antifeeding compounds active against insect herbivores (Hopkins et al., 2009). However, recent studies on glucosinolates in Arabidopsis have firmly established roles for them in the defense against microbial pathogens as well (Bednarek et al., 2009; Clay et al., 2009; Fan et al., 2011; Bednarek, 2012a; Johansson et al., 2014). Indole glucosinolates are linked to broad-spectrum nonhost resistance against powdery mildew fungi (Bednarek et al., 2009), signaling triggered by the flagellin recognition in Arabidopsis (Clay et al., 2009), and initiation of the HR following the recognition of *P. syringae* effectors (Johansson et al., 2014). Sulforaphane, specifically, was previously linked to resistance against nonadapted strains of *P. syringae* in Arabidopsis (Fan et al., 2011), and it was reported that nonadapted bacterial strains were highly sensitive to sulforaphane, whereas strains adapted for the infection of Arabidopsis and other cruciferous plants were shown to harbor *SAX* genes mediating the detoxification of sulforaphane produced by the plant host. We found that when wild-type *P. syringae* expressing AvrRpm1 was infiltrated into Arabidopsis leaf tissue, the released sulforaphane was degraded at 24 hpi (Fig. 2B). This was not observed with the expression of transgenic AvrRpm1 (Fig. 2A), leading us to believe that the SAX genes in *P. syringae* are initially unable to keep up with the rapid production of sulforaphane. This was supported by the observation that the sulforaphane released was not removed by a Δ*sax P. syringae* mutant expressing AvrRpm1 (Fig. 2C). The exact mechanism for the apparent removal of sulforaphane by *P. syringae* remains unclear.

Sulforaphane is but one of a family of similar compounds in Arabidopsis. The composition of the glucosinolate precursors and the cleavage products varies largely between different Arabidopsis accessions (Kliebenstein et al., 2001). The sulforaphane precursor glucoraphanine is the dominant glucosinolate in the accession Col-0 leaf tissue (Kliebenstein et al., 2001; Brown et al., 2003). We infiltrated a number of structurally dissimilar isothiocyanates into Arabidopsis Col-0 wild-type leaf tissue and measured electrolyte leakage (Fig. 3D). The results presented show that

several different isothiocyanates could induce cell death. Thus, it seems likely that other glucosinolate breakdown products play a similar role in other accessions. Moreover, given the ability of sulforaphane to induce cell death in a variety of plants (Fig. 3, B and C), the activity of these compounds might be conserved across the plant kingdom. Furthermore, indole-3-acetonitrile also can induce cell death upon infiltration into leaf tissue (Johansson et al., 2014). Our admittedly very small screen of compounds did reveal very different potency depending on the nature of the side chains. This is in agreement with the dependence on the side chain for the effects of isothiocyanates in various mammalian systems as well (Wu et al., 2009). However, any conclusions about the importance of chain length, structure, or substituents would require a more thorough screen of different compounds.

To assess the in vivo function of the sulforaphane released during the HR, two different double mutant lines were used: *myb28 myb29* and *tgg1 tgg2*. Both double mutants lack the ability to produce sulforaphane (Barth and Jander, 2006; Beekwilder et al., 2008; Sønderby et al., 2010a). Cell death induced by *P. syringae* expressing AvrRpm1 decreased only moderately in each of these double mutants, and there was no detectable change in the ability of the mutant plants to restrict bacterial growth (Fig. 5B; Supplemental Fig. S3). Thus, the HR pathway triggered by AvrRpm1 is able to proceed relatively unhindered despite an inability to produce sulforaphane. This is perhaps not surprising, given that the defense response induced by AvrRpm1 is particularly robust (Tsuda et al., 2009). In contrast, there was a clear cell death and resistance phenotype of the double mutants toward two different isolates of the oomycete *Hpa*, with clearly delayed cell death and enhanced pathogen growth and sporulation (Fig. 6, A, D, and E). The *tgg* and *myb* double mutants are arguably not ideal in their specificity toward lowered sulforaphane production, but given the very different nature of the mutated genes in the lines, it seems reasonable that the shared phenotypes are indeed caused by the lack of aliphatic glucosoinolate breakdown products rather than other indirect effects of the mutations. Since the biosynthesis pathway for the aliphatic glucosinolates is not fully understood, no specific biosynthetic mutants are available, and we deem that these two double mutants represent the best available genetic test for the involvement of sulforaphane in effector-triggered immunity. It is difficult to entirely separate the possible effects of sulforaphane on the plant cells and the possible direct effect on the pathogen in the mutant experiment. However, since sulforaphane by itself clearly affects plant tissue, at the very least both effects are likely present during the defense reaction. Future transcriptional profiling studies of the effects of sulforaphane treatment of plant tissue are likely to shed more light on the direct effects on the plant.

As sulforaphane infiltration caused a quick decrease in the cellular glutathione pool (Fig. 4), we propose the reaction of sulforaphane with glutathione and the subsequent increase in redox potential as the most straightforward explanation for the action of sulforaphane. Thus, if sulforaphane acts by simply decreasing the glutathione pool of the cell and thus making this pool more oxidative, there is no need for specific receptors. It should be noted that the redox potential and the size of the glutathione pool are well known to have profound influences on many different cellular processes, including cell death (Noctor et al., 2011). It has even been suggested that the half-cell reduction potential of cellular glutathione is a universal marker for stress and cell death (Kranner et al., 2006). The notion that lower concentrations and thus higher oxidation of the cellular glutathione pool could induce defense is supported by our observation of protection induced by sulforaphane. Treatment with sulforaphane at a concentration that did not induce visible lesions offered increased protection against a virulent *Hpa* isolate (Fig. 7).

It is well established that hydrogen peroxide and other reactive oxygen species (Overmyer et al., 2003; Torres, 2010) as well as other reactive electrophiles are produced during plant defense responses (Farmer and Davoine, 2007; Durand et al., 2009; Mueller and Berger, 2009; Farmer and Mueller, 2013). Furthermore, Arabidopsis tissue undergoing the HR triggered by AvrRpt2 releases salicylic acid, jasmonic acid, and 12-oxo-phytodienoic acid (Kourtchenko et al., 2007), and other studies indicate large-scale production of other electrophilic oxidation products of fatty acids (Farmer and Davoine, 2007; Tsuda et al., 2009; Farmer and Mueller, 2013). Thus, it could be envisioned that sulforaphane acts in combination with other reactive electrophilic compounds produced during plant defense responses against pathogens to execute the HR. This would provide a high degree of adaptive flexibility to plants. Several different reactive electrophiles could be utilized alone or in combination, and this would hamper the pathogen's ability to counteract the signaling pathway. It seems clear that a cocktail of compounds, which might have both direct antimicrobial effects and signaling properties in the plant, are secreted from the tissue undergoing the HR. Some of these electrophilic compounds can also activate the expression of defense genes in neighboring tissue (Bailey et al., 2011). This specific role might be played by structurally very different compounds in different plant species. Some support for this notion comes from research on benzoxazinone glucosides, which are secondary metabolites stored in both monocot and dicot plant species (Frey et al., 2009). These compounds are also stored as inactive precursors and cleaved to unstable aglycones upon tissue disintegration or possibly other stresses. Maize (*Zea mays*) mutants lacking the capacity to produce these show striking similarities in pathogen defense to Arabidopsis mutants impaired in glucosinolate production (Ahmad et al., 2011; Bednarek, 2012a). Furthermore, maize benzoxazinone 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one directly reacts with glutathione and other biologically relevant thiol groups forming adducts (Dixon et al., 2012).

Our results support that sulforaphane and possibly other glucosinolate breakdown products might also function in the regulation of the expression of defense genes in neighboring tissue. It is apparent that natural products such as sulforaphane occupy a gray zone between antimicrobials and signaling compounds (Bednarek, 2012b, 2012a). However, it is important to note that, in this case, signal does not refer to a classical ligand-receptor pair. Rather, the signal is composed of multiple reactive compounds that together provide an activity that drives the cell in the direction of certain outcomes: PCD or defense priming.

To conclude, we here provide evidence that the isothiocyanate sulforaphane is released from Arabidopsis cells recognizing a bacterial effector protein. The released sulforaphane stimulates cell death and possibly other types of plant defense responses in the complex multicellular defense reaction, the HR.

MATERIALS AND METHODS

Plant Cultivation and Bacterial Strains

Arabidopsis (*Arabidopsis thaliana*) was cultivated under short-day conditions (8-h day and 16-h night) at 22°C daytime and 18°C nighttime and 60% relative humidity for 4 to 6 weeks. The *myb28 myb29* mutant lines were confirmed by PCR using flanking and insert-specific primers (Supplemental Fig. S2). *Pseudomonas syringae* pv *tomato* DC3000 transformed with a vector carrying the avirulence gene AvrRpm1 was maintained on solid King's B medium with appropriate antibiotics at room temperature. The vector pVSP61 containing the AvrRpm1 gene was isolated from DC3000 (Mackey et al., 2002) and transformed into the Δsax *P. syringae* mutant (Fan et al., 2011).

Purification of Cell Death-Inducing Substances

Ten grams of leaf tissue from 5- to 6-week-old DEX:AvrRpm1/Col-0 or DEX:AvrRpm1/*rpm1-3* plants was harvested into a beaker containing 50 mL of deionized water with 20 μM DEX (Sigma-Aldrich) and incubated for 6 h with gentle agitation. The solution was filtered through filter paper and stored at -20°C under N_2 as a precaution. Small molecules were captured on a preconditioned 500-mg C18 solid-phase extraction column (Discovery SPE; Supelco). The column was washed with water, and bound compounds were eluted with methanol. The eluate was dried under nitrogen, suspended in water, and used for infiltration or further purified by partitioning against chloroform:methanol:water (1:1:0.9). The chloroform phase was collected, dried, dissolved in methanol, and subjected to HPLC on a 125- × 4.5-mm C18 LiChrocart column (Merck). Flow was maintained at 1 mL min^{-1}, and isocratic elution with methanol:water (70:30) was followed by a linear gradient to 80% (v/v) 2-propanol for 30 min. Fractions were collected manually, dried under nitrogen, dissolved in 0.5 mL of water, and used for leaf infiltration. The purified active fraction was dried in vacuo and dissolved in 100 μL of distilled ethyl acetate, and aliquots of 0.5 μL were injected into a gas chromatograph (Hewlett-Packard model 5890) equipped with a mass selective detector (Hewlett-Packard model 5970). The detector was operated in the scan mode (m/z 50–600), and a capillary column of 5% phenyl-methylsiloxane (12 m, 0.33-μm film thickness) with helium as the carrier gas was used. The oven temperature was raised from 80°C to 300°C at a rate of 10°C min^{-1}. The injector (temperature of 200°C) was operated in the split mode.

Pathogen Growth Assays, Electrolyte Leakage, and Trypan Blue Staining

Exponentially growing *P. syringae* pv *tomato* was resuspended from plates in 10 mM $MgCl_2$ and diluted to the proper optical density at 600 nm (OD_{600}).

For bacterial growth experiments, whole leaves were pressure inoculated with the suspension ($OD_{600} = 0.00002$, corresponding to 7.1×10^3 colony-forming units [cfu]) using a needleless syringe. Plants were returned to growth chambers, and samples were collected at the indicated time points (four 7-mm leaf discs from separate plants in four replicates). The discs were homogenized in 10 mM $MgCl_2$, serial diluted, and plated on King's B medium agar containing rifampicin and kanamycin. The number of cfu was determined 3 d after extraction and represented as described (Morel and Dangl, 1999). Electrolyte leakage assays were performed as described (Mackey et al., 2002) with minor modifications. Bacteria suspended in 10 mM $MgCl_2$ were diluted to $OD_{600} = 0.01$, corresponding to 3.55×10^6 cfu mL^{-1}, and vacuum infiltrated into leaf discs (diameter, 7 mm). The discs were rinsed in deionized water and transferred to six-well microtiter plates containing 10 mL of deionized water (four discs per well in six replicates). Conductance was measured at the indicated time points. Maintenance of *Hyaloperonospora arabidopsidis* isolates, preparation of inoculum for experiments, and assessment of sporulation were as described (Tör et al., 2002). Arabidopsis seedlings were sprayed with 400 μM sulforaphane suspended in deionized water 24 h before or after inoculation with *Hpa* spores as indicated. Cell death was visualized by Trypan Blue staining as described (Koch and Slusarenko, 1990).

Synthesis of [1,1,2,2,3,3,4,4-^{2}H$_8$]D,L-Sulforaphane

Octadeuterio-sulforaphane was synthesized starting with [1,1,2,2,3,3,4,4-^{2}H$_8$]4-bromo-1-butanol (QMX Laboratories) by modification of a previously described synthetic route to unlabeled sulforaphane (Ding et al., 2006). The product was purified by reverse-phase HPLC (solvent system, acetonitrile:water, 1:3 [v/v]), affording the title compound (15 mg) as a colorless oil. The electron-impact mass spectrum showed prominent ions at m/z 185 (M$^+$), 167 (M$^+$-O^{2}H), 122, 118, 90, 74 ([S=C=N-C^{2}H$_2$]$^+$), and 62 ([C^{2}H$_2$-C^{2}H$_2$-C^{2}H=C^{2}H$_2$]$^+$). The identity of the material was further established by gas-liquid chromatography, reverse-phase HPLC, and UV spectrometry using a sample of authentic sulforaphane (Sigma-Aldrich) as a reference.

Quantification of the Released Sulforaphane

Two leaf discs were incubated in 1 mL of water after infiltration with *P. syringae* or DEX solution. At the times indicated, the discs were removed and deuterated sulforaphane was added as an internal standard to the aqueous bathing solution. The sulforaphane was extracted by partitioning against 1 mL of dichloromethane. The organic phase was dried under N_2 and analyzed by GC-MS as described above or using an Agilent 7820 gas chromatograph coupled to an Agilent 5975 mass selective detector. The ions 160 and 167 m/z were used for quantification of the endogenous and deuterated internal standards, respectively.

Quantification of Total and Oxidized Glutathione

Leaf discs were vacuum infiltrated with sulforaphane suspended in water and left at room temperature for 30 min, after which 20 discs (130 mg of tissue) were ground into a fine powder in liquid nitrogen. Total and oxidized glutathione content in the tissue extracts was determined with an enzymatic method (Tietze, 1969; Queval and Noctor, 2007). Samples were prepared in triplicate and measured three times. The glutathione redox potential was calculated as described (Queval and Noctor, 2007).

Statistical Analysis

Statistical analysis was performed as described (Johansson et al., 2014) using GraphPad Prism 6 (GraphPad Software). The conductivity at the final time point (6 h) of ion leakage assays using bacterial infiltration (Fig. 5B), the electrolyte leakage at each time point for chemical infiltration compared with mock treatment (Fig. 3D), the bacterial growth at 3 dpi (Supplemental Fig. S3), the cell death progression and sporulation after *Hpa* inoculation (Figs. 6, A, D, and E, and 7), and the sulforaphane release following bacterial infiltration (Fig. 5; Supplemental Fig. S2) were subjected to one-way ANOVA with Tukey's posthoc analysis, with $P < 0.05$ considered significant. Samples were grouped based on significant differences between them.

Andersson et al.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Experimental setup for identification of cell death-inducing compounds.

Supplemental Figure S2. Confirmation and characterization of *myb* mutants.

Supplemental Figure S3. In planta growth of *P. syringae* expressing AvrRpm1 in sulforaphane-deficient mutants.

Supplemental Figure S4. Sulforaphane pretreatment effects on *Hpa* infection at 7 dpi.

LITERATURE CITED

Agerbirk N, Olsen CE (2012) Glucosinolate structures in evolution. Phytochemistry 77: 16–45

Ahmad S, Veyrat N, Gordon-Weeks R, Zhang Y, Martin J, Smart L, Glauser G, Erb M, Flors V, Frey M, et al (2011) Benzoxazinoid metabolites regulate innate immunity against aphids and fungi in maize. Plant Physiol 157: 317–327

Andersson MX, Kourtchenko O, Dangl JL, Mackey D, Ellerström M (2006) Phospholipase-dependent signalling during the AvrRpm1- and AvrRpt2-induced disease resistance responses in *Arabidopsis thaliana*. Plant J 47: 947–959

Bailey K, Cevik V, Holton N, Byrne-Richardson J, Sohn KH, Coates M, Woods-Tör A, Aksoy HM, Hughes L, Baxter L, et al (2011) Molecular cloning of ATR5(Emoy2) from Hyaloperonospora arabidopsidis, an avirulence determinant that triggers RPP5-mediated defense in Arabidopsis. Mol Plant Microbe Interact 24: 827–838

Barth C, Jander G (2006) Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense. Plant J 46: 549–562

Bednarek P (2012a) Chemical warfare or modulators of defence responses: the function of secondary metabolites in plant immunity. Curr Opin Plant Biol 15: 407–414

Bednarek P (2012b) Sulfur-containing secondary metabolites from Arabidopsis thaliana and other Brassicaceae with function in plant immunity. ChemBioChem 13: 1846–1859

Bednarek P, Pislewska-Bednarek M, Svatos A, Schneider B, Doubsky J, Mansurova M, Humphry M, Consonni C, Panstruga R, Sanchez-Vallet A, et al (2009) A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. Science 323: 101–106

Beekwilder J, van Leeuwen W, van Dam NM, Bertossi M, Grandi V, Mizzi L, Soloviev M, Szabados L, Molthoff JW, Schipper B, et al (2008) The impact of the absence of aliphatic glucosinolates on insect herbivory in Arabidopsis. PLoS ONE 3: e2068

Bent AF, Mackey D (2007) Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. Annu Rev Phytopathol 45: 399–436

Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J (2003) Variation of glucosinolate accumulation among different organs and developmental stages of Arabidopsis thaliana. Phytochemistry 62: 471–481

Clay NK, Adio AM, Denoux C, Jander G, Ausubel FM (2009) Glucosinolate metabolites required for an Arabidopsis innate immune response. Science 323: 95–101

Coates ME, Beynon JL (2010) Hyaloperonospora arabidopsidis as a pathogen model. *In* NK VanAlfen, G Bruening, JE Leach, eds, Annual Review of Phytopathology, Vol 48. Annual Reviews, Palo Alto, CA, pp 329–345

Dempsey DA, Klessig DF (2012) SOS: too many signals for systemic acquired resistance? Trends Plant Sci 17: 538–545

Desveaux D, Singer AU, Dangl JL (2006) Type III effector proteins: doppelgangers of bacterial virulence. Curr Opin Plant Biol 9: 376–382

Ding TJ, Zhou L, Cao XP (2006) A facile and green synthesis of sulforaphane. Chin Chem Lett 17: 1152–1154

Dixon DP, Sellars JD, Kenwright AM, Steel PG (2012) The maize benzoxazinone DIMBOA reacts with glutathione and other thiols to form spirocyclic adducts. Phytochemistry 77: 171–178

Durand T, Bultel-Poncé V, Guy A, Berger S, Mueller MJ, Galano JM (2009) New bioactive oxylipins formed by non-enzymatic free-radical-catalyzed pathways: the phytoprostanes. Lipids 44: 875–888

Fahey JW, Zalcmann AT, Talalay P (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56: 5–51

Fan J, Crooks C, Creissen G, Hill L, Fairhurst S, Doerner P, Lamb C (2011) Pseudomonas sax genes overcome aliphatic isothiocyanate-mediated non-host resistance in Arabidopsis. Science 331: 1185–1188

Farmer EE, Davoine C (2007) Reactive electrophile species. Curr Opin Plant Biol 10: 380–386

Farmer EE, Mueller MJ (2013) ROS-mediated lipid peroxidation and RES-activated signaling. Annu Rev Plant Biol 64: 429–450

Feng F, Zhou JM (2012) Plant-bacterial pathogen interactions mediated by type III effectors. Curr Opin Plant Biol 15: 469–476

Frey M, Schullehner K, Dick R, Fiesselmann A, Gierl A (2009) Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. Phytochemistry 70: 1645–1651

Hopkins RJ, van Dam NM, van Loon JJA (2009) Role of glucosinolates in insect-plant relationships and multitrophic interactions. Annu Rev Entomol 54: 57–83

Hückelhoven R, Panstruga R (2011) Cell biology of the plant-powdery mildew interaction. Curr Opin Plant Biol 14: 738–746

Johansson ON, Fantozzi E, Fahlberg P, Nilsson AK, Buhot N, Tör M, Andersson MX (2014) Role of the penetration-resistance genes PEN1, PEN2 and PEN3 in the hypersensitive response and race-specific resistance in Arabidopsis thaliana. Plant J 79: 466–476

Kliebenstein DJ, Kroymann J, Brown P, Figuth A, Pedersen D, Gershenzon J, Mitchell-Olds T (2001) Genetic control of natural variation in Arabidopsis glucosinolate accumulation. Plant Physiol 126: 811–825

Koch E, Slusarenko A (1990) *Arabidopsis* is susceptible to infection by a downy mildew fungus. Plant Cell 2: 437–445

Kourtchenko O, Andersson MX, Hamberg M, Brunnström A, Göbel C, McPhail KL, Gerwick WH, Feussner I, Ellerström M (2007) Oxophytodienoic acid-containing galactolipids in Arabidopsis: jasmonate signaling dependence. Plant Physiol 145: 1658–1669

Kranner I, Birtić S, Anderson KM, Pritchard HW (2006) Glutathione half-cell reduction potential: a universal stress marker and modulator of programmed cell death? Free Radic Biol Med 40: 2155–2165

Mackey D, Belkhadir Y, Alonso JM, Ecker JR, Dangl JL (2003) Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. Cell 112: 379–389

Mackey D, Holt BF III, Wiig A, Dangl JL (2002) RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. Cell 108: 743–754

Monaghan J, Zipfel C (2012) Plant pattern recognition receptor complexes at the plasma membrane. Curr Opin Plant Biol 15: 349–357

Morel JB, Dangl JL (1999) Suppressors of the Arabidopsis lsd5 cell death mutation identify genes involved in regulating disease resistance responses. Genetics 151: 305–319

Mueller MJ, Berger S (2009) Reactive electrophilic oxylipins: pattern recognition and signalling. Phytochemistry 70: 1511–1521

Mur LAJ, Kenton P, Lloyd AJ, Ougham H, Prats E (2008) The hypersensitive response: the centenary is upon us but how much do we know? J Exp Bot 59: 501–520

Noctor G, Queval G, Mhamdi A, Chaouch S, Foyer CH (2011) Glutathione. The Arabidopsis Book 9: e0142, doi/10.1199/tab.0142

Overmyer K, Brosché M, Kangasjärvi J (2003) Reactive oxygen species and hormonal control of cell death. Trends Plant Sci 8: 335–342

Parker JE, Holub EB, Frost LN, Falk A, Gunn ND, Daniels MJ (1996) Characterization of eds1, a mutation in *Arabidopsis* suppressing resistance to Peronospora parasitica specified by several different RPP genes. Plant Cell 8: 2033–2046

Queval G, Noctor G (2007) A plate reader method for the measurement of NAD, NADP, glutathione, and ascorbate in tissue extracts: application to redox profiling during Arabidopsis rosette development. Anal Biochem **363:** 58–69

Ross AF (1961) Localized acquired resistance to plant virus infection in hypersensitive hosts. Virology **14:** 329–339

Shah J (2009) Plants under attack: systemic signals in defence. Curr Opin Plant Biol **12:** 459–464

Sinapidou E, Williams K, Nott L, Bahkt S, Tör M, Crute I, Bittner-Eddy P, Beynon J (2004) Two TIR:NB:LRR genes are required to specify resistance to Peronospora parasitica isolate Cala2 in Arabidopsis. Plant J **38:** 898–909

Sønderby IE, Burow M, Rowe HC, Kliebenstein DJ, Halkier BA (2010a) A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic glucosinolates in Arabidopsis. Plant Physiol **153:** 348–363

Sønderby IE, Geu-Flores F, Halkier BA (2010b) Biosynthesis of glucosinolates: gene discovery and beyond. Trends Plant Sci **15:** 283–290

Stakman EC (1915) Relation between Puccinia graminis and plants highly resistant to its attack. J Agric Res **4:** 193–200

Tierens KF, Thomma BPH, Brouwer M, Schmidt J, Kistner K, Porzel A, Mauch-Mani B, Cammue BPA, Broekaert WF (2001) Study of the role of antimicrobial glucosinolate-derived isothiocyanates in resistance of Arabidopsis to microbial pathogens. Plant Physiol **125:** 1688–1699

Tietze F (1969) Enzymic method for quantitative determination of nanogram amounts of total and oxidized glutathione: applications to mammalian blood and other tissues. Anal Biochem **27:** 502–522

Tör M, Gordon P, Cuzick A, Eulgem T, Sinapidou E, Mert-Türk F, Can C, Dangl JL, Holub EB (2002) *Arabidopsis* SGT1b is required for defense signaling conferred by several downy mildew resistance genes. Plant Cell **14:** 993–1003

Torres MA (2010) ROS in biotic interactions. Physiol Plant **138:** 414–429

Torres MA, Jones JDG, Dangl JL (2006) Reactive oxygen species signaling in response to pathogens. Plant Physiol **141:** 373–378

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

Valgimigli L, Iori R (2009) Antioxidant and pro-oxidant capacities of ITCs. Environ Mol Mutagen **50:** 222–237

van der Biezen EA, Freddie CT, Kahn K, Parker JE, Jones JDG (2002) Arabidopsis RPP4 is a member of the RPP5 multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components. Plant J **29:** 439–451

Wu X, Zhou QH, Xu K (2009) Are isothiocyanates potential anti-cancer drugs? Acta Pharmacol Sin **30:** 501–512
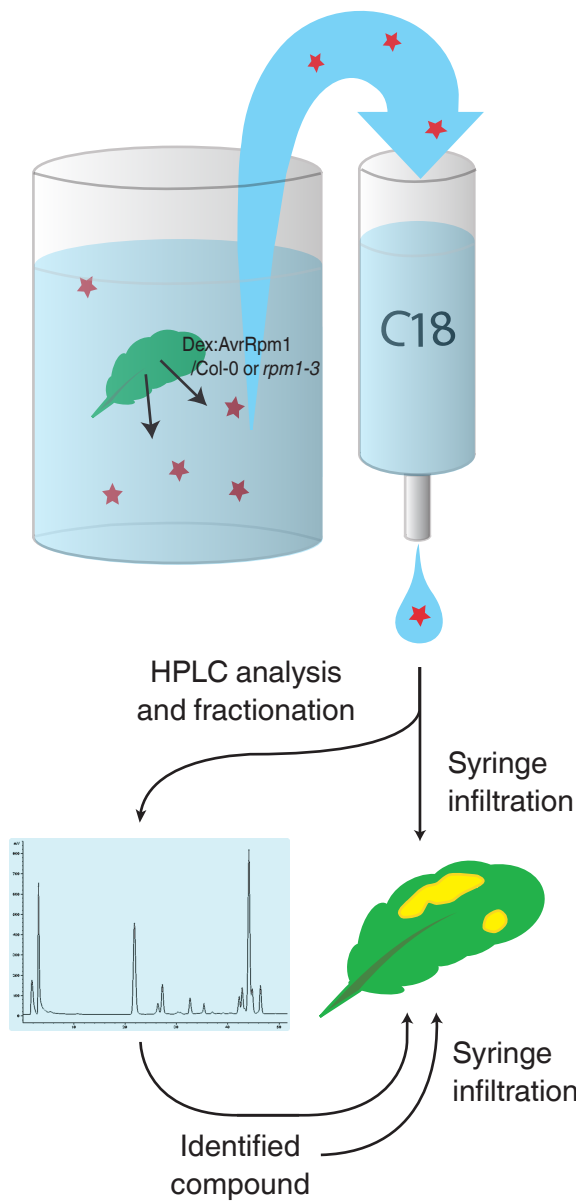
Fig. S1. *Experimental setup for isolating compounds inducing cell death released from tissue undergoing the HR*

Transgenic Arabidopsis leaf tissue expressing the bacterial effector AvrRpm1 with or without the cognate R protein RPM1 is incubated in water. The bathing solution is run through a C18 column to capture small molecules and the captured compounds are infiltrated into leaf tissues or subjected to semi preparative HPLC. The fractions obtained from HPLC are infiltrated into wild type tissues.
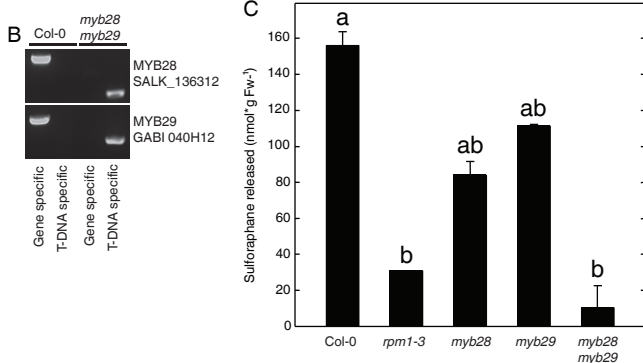

Fig. S2. *Confirmation of myb-mutants and release of sulforaphane during HR induced by P. syringae expressing AvrRpm1*

The indicated insertion mutant lines were confirmed using flanking and insert specific primers (A and B). Leaf discs from wild type Arabidopsis and the indicated mutants were infiltrated with P. syringae DC3000:AvrRpm1 at OD 0.01 and incubated in water. After 8 hours, the discs were removed and the amount of sulforaphane in the bathing solution determined (C). Average and range of duplicate samples are shown. Letters a-c indicates statistically significant groups (One way ANOVA with Tukeys post hoc test p < 0.05). The experiment was repeated twice with similar results.
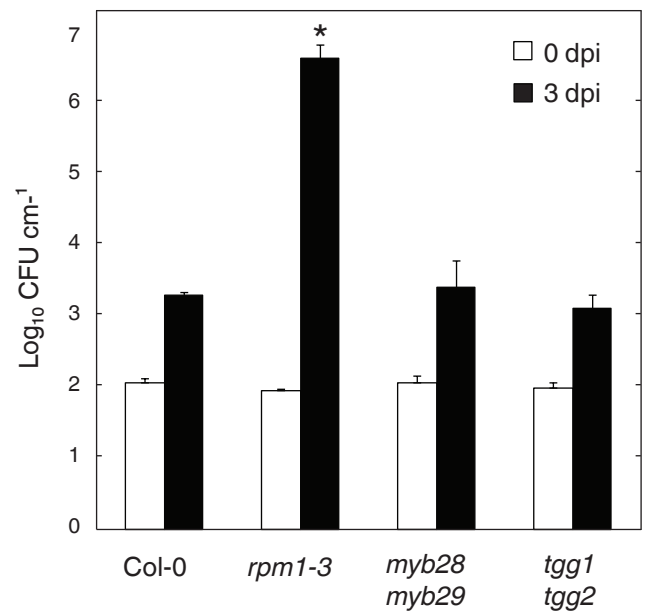

Fig. S3. *In planta growth of P. syringae expressing AvrRpm1 in sulforaphane deficient mutants*

Wild type and the indicated double mutant leaves were syringe inoculated with a solution containing $7.1*10^3$ cfu ml-1 of P. syringae expressing AvrRpm1. The bacteria were extracted immediately or after three days, serially diluted and plated on selective medium. Shown are average bacterial density and standard deviation from 4 replicates. Asterisks indicate statistical significance compared to Col-0 (One way ANOVA p <0.05). The experiment was repeated twice with similar results.
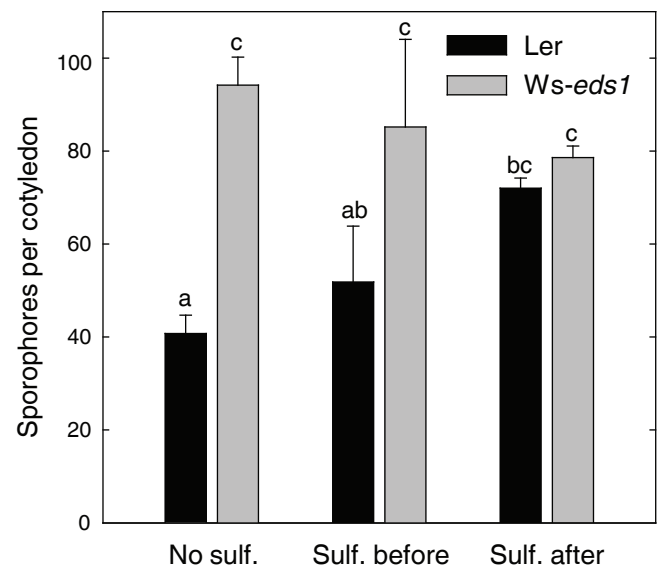

Fig. S4. *Sulforpahane pretreatment is ineffective at 7 dpi*

Seedlings of wild type Ler or eds1 in the Ws-0 background were sprayed with 200 μM of sulforaphane 24 hours before or 24 hours after inoculation with Hpa isolate Cala2 and the resulting sporulation counted at 7 dpi. Average number of spores and range of three replicate samples are shown. Letters a-c indicates statistically significant groups (One way ANOVA with Tukeys post hoc test p < 0.05). The experiment was repeated twice with identical results.