Discovery of *Phytophthora cinnamomi RxLR* effector genes expressed in avocado during infection.

by

Melissa Joubert

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

Faculty of Natural and Agricultural Sciences

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Pretoria

December 2019

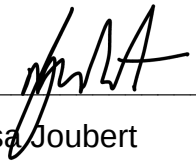
Supervisor: Prof Noëlani van den Berg

Co-supervisor: Dr S. Ashok Prabhu

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Declaration

I, Melissa Joubert, hereby declare that the thesis which I hereby submit for the degree Magister Scientiae to the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.



Melissa Joubert
December 2019

Acknowledgements

I would like to thank each of the following people and organisations that made the completion of this research project possible:

- To my primary supervisor, Prof Noëlani van den Berg, for your guidance and support. Thank you for keeping your door open to me, and for always having the time to nurture the aspiring scientist in me. Thank you for your understanding and encouragement, for your enthusiasm and care, and for providing the kind of research environment that is a joy to work in. Thank you for being an incredible leader, and for being an inspiration to all of the students who are lucky enough to learn under your supervision.
- To the Hans Merensky Foundation and the National Research Foundation for providing the funding that made this project possible.
- Dr Ashok Prabhu, Mr. Robert Backer, Ms. Juanita Engelbrecht and Dr Velushka Swart. Thank you for your willingness to answer my ceaseless questions and your effort to guide my understanding of the project. Thank you for your endless patience, and for being a sounding board for all my ideas and hypotheses. Your mentorship has been invaluable to me.
- Dr Tuan Duong, Dr Sarah Mwangi, Ms. Tsakani Miyambo and Ms. Wilma Nel. Thank you for all your assistance with the bioinformatics portions of this project, and for taking the time to explain the more complicated genetics aspects to me when the information was overwhelming.
- To my friends; Ms. Juanita Hanneman, Mr. Seamus Morgan, Mrs. Shannon Wilson, Ms. Julanie Stapelberg and Mr. Stephan Henning, thank you for going through all the ups and downs of this degree with me. Thank you for your encouragement and support, for providing a bright side to every downfall, and for bringing sunshine to every day.
- My parents; Johan, Liz, Michelle and Heinrich, and my sister Jessica, thank you for your motivation at all times, for having faith in me when I forgot how to believe in myself, and for always pushing me to be the best version of myself that I can be.
- To Mr. André Engelbrecht, thank you for supporting me through every part of this last phase of my project. Thank you for always listening to me talk about

scientific problems, even when I don’t make any sense. Thank you for your willingness to make a plan to assist me with my project and asking for nothing in return. Thank you for always offering the best advice and knowing exactly what to say to help me on my way to finishing this degree. I do not have words to express my gratitude.

- Finally, to all my friends and colleagues in the Avocado Research Programme and in the Forestry and Agricultural Biotechnology Institute, thank you for being a constant source of smiles, positive attitudes, commiserations and support. Thank you for creating a constructive work environment, and a home away from home. I am blessed to be surrounded by such wonderful people every day.

Preface

Phytophthora cinnamomi is a plant pathogenic oomycete that causes Phytophthora Root Rot in avocado trees, which has led to large crop losses in the past. Several control strategies are currently in place to limit the extent of disease caused by the pathogen, but efforts to develop new avenues for control are impeded by the lack of understanding regarding infection strategies used by this pathogen at the molecular level. Other *Phytophthora* species have been shown to use RxLR effector proteins as molecular weapons to interfere with host defenses in compatible plant-pathogen interactions. While recent studies have identified putative RxLRs from the *P. cinnamomi* genome, it remains unclear whether these candidate effectors contribute to virulence on avocado hosts.

Chapter 1 of this thesis presents a review of literature on RxLR effectors of *Phytophthora* species. First, the current knowledge of the characteristics of these effectors is assessed. Subsequently, methods of identification and functional characterization of RxLR effectors are reviewed, and detailed examples of characterized RxLRs are provided. Finally, possible implications of RxLR effector research are evaluated, with a focus on their application for disease control strategies.

Chapter 2 reports on the identification of *RxLR* effector genes expressed by *P. cinnamomi* during infection of avocado. Putative *RxLR*s were evaluated for their suitability as candidate effectors based on their fulfilment of selected criteria, and their expression profiles were assessed from RNA-seq data obtained from an infected avocado rootstock. Selected candidates were manually annotated, and their resultant protein sequences were subjected to phylogenetic analysis. Potential functions were then assigned to the candidate *P. cinnamomi* RxLRs based on their expression profiles and relatedness to previously characterized effectors.

Summary

Phytophthora cinnamomi is an oomycete that targets a broad range of plants, including several economically important forestry and agricultural crops. It is the causal agent of Phytophthora Root Rot, and causes significant economic losses within the agricultural and forestry industries. Recently, the use of effector molecules by pathogenic oomycetes during plant infection has become a subject of great interest to researchers. One class of these molecules, the RxLR effectors, has become a focus of *Phytophthora* research, and hundreds of putative *RxLR* effector genes have been predicted by bioinformatic analysis of *Phytophthora* genomic sequences. The characterization and validation of these effectors remains an ongoing process.

This study identified several *P. cinnamomi RxLR* genes upregulated during infection of a susceptible avocado rootstock. The genes were subjected to *in silico* analysis of expected RxLR characteristics and prediction of coding regions from the genomic sequence. Predictions were then validated by analysis of DNA sequences and the use of RNA-seq data, which were used to manually annotate these effector genes. The final prediction of RxLR proteins was then compared to the sequences of validated RxLRs in other *Phytophthora* species to enable inference of possible functions of the annotated genes.

In this study, a total of 25 *P. cinnamomi* candidate RxLRs were identified, which were proposed to play a role during avocado infection. While expression of this number of candidate *RxLR*s is relatively small, these candidates may represent effectors which are expressed specifically in this host-pathogen interaction, or may be a set of “elite” effectors which contribute to virulence in all hosts. The candidate genes were analysed for the presence of the desired motifs, and a subset of 16 *RxLR*s were chosen for further analysis. The expression profiles of these genes were investigated further, and it was found that four of the candidates were expressed most highly at 24 hpi, which correlates with expression profiles of RxLRs in other species. Twelve of the candidate RxLRs had expression profiles which were not similar to those which have been demonstrated for other RxLRs, while four were not significantly upregulated during specific timepoints of infection. These results warrant further investigation to determine the relevance of these unique expression profiles, which may present new insights into expression patterns of RxLR effectors.

Several of the 16 candidate effector genes were present in multiple copies in the *P. cinnamomi* genome, providing evidence for their roles in plant infection. Transcriptome data was used to manually annotate the genes, and the resulting protein predictions for most of the candidates were different from those originally predicted by gene prediction software. Not one of the prediction software used in this experiment accurately predicted the coding regions for all the genes – providing a substantial argument for the need for manual annotation of candidate effectors.

Phylogenetic analysis allowed functional inferences to be made for five of the candidate effectors, based on their shared evolutionary history with RxLRs characterized in other *Phytophthora* species. While no functional assays have been carried out for these candidate effectors yet, their identification as putative RxLRs presents a starting point for further investigation into their functions *in planta*. This study presents the first report *of P. cinnamomi RxLR*s with confirmed sequences and expression profiles, and as such offers the first insights into infection of avocado by this pathogen at the molecular level.

Table of contents **Page**

List of Abbreviations:

AA	Amino acid
ADP	Adenosine diphosphate
AFLP	Amplified fragment length polymorphism
ANOVA	Analysis of variance
ARP	Avocado Research Programme
Avh	Avirulence homolog
avr	Avirulence gene
BAC	Bacterial artificial chromosome
BAM	Binary alignment map
BAX	Blc-2-associated X protein
BiP	Binding immunoglobulin protein
BLAST	Basic local alignment search tool
bp	Base pairs
$CaCO_3$	Calcium carbonate
CBEL	Cellulose-binding elicitor lectin
cDNA	Complementary DNA
CDS	Coding sequence
CTAB	Cetrimonium bromide
C-terminal	Carboxyl terminal
DAMP	Damage-associated molecular pattern
DNA	Deoxyribonucleic acid
dpi	Days post-inoculation
ER	Endoplasmic reticulum
et al.	*Et alia* meaning 'and others'
ETI	Effector-triggered immunity
ETS	Effector-triggered susceptibility
E-value	Expect value
GFF	Generic feature format
GFP	Green fluorescent protein
HMM	Hidden Markov Model
hpi	Hours post-inoculation
HR	Hypersensitive response

ID	Identity
IGV	Integrated Genome Viewer
JGI	Joint Genome Institute
Kan50	50mg/ml kanamycin
Kb	Kilobases
LB	Luria-Bertani
LPV3	Large peripheral vesicle 3 gene
MAMP	Microbe-associated molecular pattern
MAP	Mitogen-activated protein
Mb	Megabases
MCMC	Markov chain Monte Carlo
min	Minute
miRNA	microRNA
mRNA	Messenger RNA
MUSCLE	Multiple Sequence Comparison by Log-Expectation
N_2	Nitrogen
NaAcO	Sodium acetate
NaCl	Sodium chloride
NADH	Nicotinamide adenine dinucleotide with hydrogen
NBS-LRR	Nucleotide-binding site leucine-rich repeat
NCBI	National Center for Biotechnology Information
NMR	Nuclear magnetic resonance
NN	Neural networks
NPR1	Non-expressor of pathogenesis related -1
N-terminal	Amino terminal
ORF	Open reading frame
PAMP	Pathogen-associated molecular pattern
PCD	Programmed cell death
PCR	Polymerase chain reaction
PI3P	Phosphatidylinositol-3-phosphate
PI4P	Phosphatidylinositol-4-phosphate
PINP	PSR1-interacting protein
PP1c	Protein phosphatase 1 catalytic

PRR	Pattern recognition receptor
PSE	Penetration-specific effector
PSR	Phytophthora suppressor of RNA silencing
PTI	Pattern-triggered immunity
R gene	Resistance gene
R protein	Resistance protein
RAPD	Random amplified polymorphic DNA
RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive oxygen species
rpm	Revolutions per minute
s	Seconds
SD	Standard deviation
SE	Standard error
siRNA	Small interfering RNA
SKRP	Serine/lysine/arginine-rich protein
SNE	Single nucleotide polymorphism
SNP	Suppressor of necrosis
TPM	Transcripts per million
UV	Ultraviolet

Chapter 1

Identification and functional analysis of RxLR effectors of the genus *Phytophthora*

Introduction

Oomycetes are a group of eukaryotic microorganisms belonging to the Kingdom Stramenopila. They resemble fungi in their filamentous growth and production of mycelia during vegetative growth stages, but hyphae are coenocytic, and cell walls contain cellulose and glucans rather than chitin. Moreover, oomycetes produce unique sexual and asexual spores known as oospores and zoospores, respectively (Latijnhouwers *et al.*, 2003, Agrios, 2005).

The majority of oomycetes are parasites of plants, with the most important plant pathogens belonging to the orders Peronosporales and Saprolegniales (Agrios, 2005, Thines & Kamoun, 2010). In particular, the order Peronosporales includes several genera of important plant pathogens such as *Albugo*, *Hyaloperonospora*, *Peronospora*, *Phytophthora* and *Pythium* (Agrios, 2005, Bozkurt *et al.*, 2012). The genus *Phytophthora* is of particular interest to plant pathologists, as it contains some of the most destructive plant pathogens worldwide (Zentmyer, 1976, Erwin & Ribeiro, 1996, Kroon *et al.*, 2012). Based on his studies on the potato late blight pathogen, *Phytophthora infestans*, De Bary (1876) named the genus the “Plant Destroyer”.

The genus *Phytophthora* includes over 140 species (Yang *et al.*, 2017a), including several economically important pathogens such as *P. infestans, P. ramorum, P. capsici, P. cinnamomi, P. cactorum, P. sojae* and *P. parasitica (*Agrios, 2005). Species in this genus vary greatly in their host ranges. Some have narrow host ranges, such as *P. infestans*, which causes diseases only on solanaceous hosts (Erwin & Ribeiro, 1996, Fry, 2008), and *P. sojae*, which causes root and stem rot in soybean (Erwin & Ribeiro, 1996, Tyler, 2007). Other species have broad host ranges and can cause diseases on hundreds to thousands of host plants. These generalists include species such as *P. ramorum*, which infects over 130 species of forestry and ornamental plants (Grünwald *et al.*, 2012a, Johnston *et al.*, 2015), and *P. cinnamomi*, with a host range of over 5000 agricultural, forestry and horticultural species (Hardham & Blackman, 2010, Allardyce *et al.*, 2012, Hardham & Blackman, 2017).

Species of *Phytophthora* also differ in their nutritional strategies, though most tend to be hemibiotrophs that infect the host biotrophically before switching to a necrotrophic growth stage (Tyler, 2009, Dou & Zhou, 2012). This means that these pathogens avoid detection *in planta* by suppressing plant immune responses, before exploiting plant defenses and causing harm to the plant host during their necrotrophic growth. To do this, *Phytophthora* species employ an arsenal of rapidly evolving pathogen effectors to manipulate host plants and facilitate disease progression (Glazebrook, 2005, Kamoun, 2007, Hein *et al.*, 2009, Fawke *et al.*, 2015).

Research into oomycete phytopathogens is only now beginning to elucidate infection at the molecular level, and recent advances in effector biology have played a major role in increasing our understanding of oomycete diseases (Hein *et al.*, 2009, Thines & Kamoun, 2010, Bozkurt *et al.*, 2012, Dou & Zhou, 2012, Fawke *et al.*, 2015, Kamoun *et al.*, 2015). The availability of sequenced oomycete genomes has allowed researchers to further progress in this field, by serving as valuable resources in basic and applied plant pathology (Pais *et al.*, 2013, Kamoun *et al.*, 2015), and effectoromics studies using these genomes as tools are instrumental in accelerating oomycete research (Vleeshouwers *et al.*, 2008, Schornack *et al.*, 2009, Vleeshouwers *et al.*, 2011, Grünwald, 2012b, Kamoun *et al.*, 2015).

This review will discuss how *Phytophthora* species cause plant disease through the use of pathogen effector molecules. The discovery and characterization of these effectors, and how genomics tools have accelerated their identification, will also be examined. Finally, we briefly review how knowledge of effectors and their interactive partners can be applied in improving disease management strategies.

Plant-pathogen molecular interactions

When a plant pathogen, such as *Phytophthora*, comes into contact with a susceptible host plant, disease symptoms appear as a result of the interaction between the pathogen and the host. Plants have evolved to possess specific immune responses to help combat such pathogens by limiting their spread within the host plant. Thus, the extent to which

disease is manifested is dependent on the extent to which the plant immune response can suppress the pathogen (Agrios, 2005). This interaction between plant and pathogen is mediated through the use of effectors by the pathogen and resistance (*R*) genes by the plant, as per the gene-for-gene hypothesis put forward by Harold Flor (1971).

The gene-for-gene hypothesis postulates that pathogens produce effector molecules encoded by avirulence genes (*avr*), which are then recognized by the products of corresponding resistance (*R*) genes in host plants, resulting in incompatible plant–pathogen interactions (Flor, 1955, Flor, 1971). These incompatible interactions are characterized by the initiation of defense responses upon effector recognition, which almost always includes the induction of localized cell death, characteristic of the hypersensitive response (HR) (Rouxel & Balesdent, 2010). The gene-for-gene hypothesis can also be extrapolated to indicate that the absence of R protein-Avr recognition, either due to lack of the *avr* gene in a virulent pathogen, or absence of the corresponding *R* gene in a susceptible host, would result in a compatible interaction and ensuing plant disease development (Glazebrook, 2005).

A more recent model for plant-pathogen interactions is the widely accepted zig-zag model of plant immunity (Jones & Dangl, 2006) (Figure 1). According to this model, when a pathogen first interacts with a plant, molecules that are usually exposed on the pathogen surface, known as microbe-associated molecular patterns (MAMPs) or pathogen-associated molecular patterns (PAMPs) are recognised in the plant apoplast by pattern recognition receptors (PRRs) localized on the membranes of plant cells. While this direct recognition of MAMPs/PAMPs by PRRs was put forward by Jones and Dangl in the original zig-zag model, it has since been shown that PRRs also play a role in indirect recognition upon infection, when damage-associated molecular patterns (DAMPs) are induced in plant cells (Cook *et al.*, 2015). These DAMPs, originally known as “endogenous elicitors” (Darvill & Albersheim, 1984), are host-derived molecules such as peptides or oligosaccharides that are induced in damaged plant cells upon wounding or infection (Lotze *et al.*, 2007, Boller & Felix, 2009).

Recognition of MAMPs/PAMPs or DAMPs by PRRs in plant cells results in pattern-triggered immunity (PTI) – the first line of defense against plant pathogens (Selin *et al.*, 2016). Typical PTI is a transient defense response involving activation of MAP (mitogen-activated protein) kinases upon pathogen recognition, as well as production of reactive oxygen species (ROS) and changes in hormone production (involving plant hormones such as salicylic acid, jasmonic acid and ethylene) (Tsuda & Katagiri, 2010). The zig-zag model proposes that pathogens subsequently release specific effector molecules to overcome PTI and increase virulence, resulting in what is known as effector-triggered susceptibility (ETS) in the plant host (Cook *et al.*, 2015).

In response to the release of these effectors by the pathogen, plant *R* genes are expressed to produce intracellular receptors known as nucleotide-binding site leucine-rich repeat (NBS-LRR) proteins, which recognize pathogen effectors inside the cells. This recognition causes effector-triggered immunity (ETI) which prevents disease development if the pathogen is sufficiently suppressed (Jones & Dangl, 2006). It should be noted, however, that ETI can also be triggered by indirect recognition by intracellular receptors rather than direct recognition of an effector by a plant R protein (Cook *et al.*, 2015). This model for recognition is known as the “guard hypothesis” (Van der Hoorn *et al.*, 2002, Jones & Dangl, 2006), and it proposes that some R proteins are responsible for monitoring specific host factors. Upon alteration of this host factor by pathogen effectors, the R protein “guarding” the host component will then activate host defense pathways (de Wit, 2007).

ETI often involves responses similar to PTI, with the most significant difference between these defense responses being that ETI tends to be a more rapid, vigorous immune response, which also lasts longer than PTI. It also more often includes HR which causes localized cell death (Jiang & Tyler, 2012). Pathogens do, however, evolve to produce novel effectors to suppress ETI and again establish ETS. This evolution of pathogen effectors drives the evolution of plant *R* genes to prevent ETS and again trigger ETI, resulting in an evolutionary “arms race” as plants and pathogens each try to establish resistance and disease, respectively (Jones & Dangl, 2006).

In summary, the zigzag model, together with the gene-for-gene hypothesis proposes that the products of plant and pathogen genes interact in such a way as to either result in disease or resistance (Flor, 1971, Jones & Dangl, 2006). Effector molecules secreted by pathogens therefore play an important role in manipulating susceptible hosts and suppressing plant immunity (Stassen & Van den Ackerveken, 2011). These pathogen effectors have been classified as either virulence or avirulence proteins, depending on the response they elicit in a particular host due to the genotype of that host (Chen *et al.*, 2013).

Effectors in the broad sense can be defined as “molecules that alter host cell structure and function” (Selin *et al.*, 2016). These molecules may then facilitate infection as virulence factors or toxins to act in favor of the pathogen (Selin *et al.*, 2016), or they may be recognized by plant R proteins to trigger a defense response in favor of the plant, in which case they are known as avirulence effectors (Rouxel & Balesdent, 2010).

It is important to note that a specific effector may have an avirulence activity in one plant host and a virulence activity in another, depending on whether host plants have resistant or susceptible genotypes, respectively. This is known as the dual function of effectors, and it means that the activity of an effector in a given host depends on the genotype of that host (van't Slot & Knogge, 2002, Kamoun, 2007). Resistant plants, in this case, will contain *R* genes specific to the effectors of that pathogen, which can interact with those avirulence effectors to initiate a defense response (such as HR) in the host (Rouxel & Balesdent, 2010), while susceptible plants lacking the *R* genes will be incapable of mounting the required defense responses, resulting in compatible plant-pathogen interactions (Glazebrook, 2005). Pathogens must therefore maintain a continually evolving repertoire of virulent effector molecules in order to cause disease in susceptible host plants (Selin *et al.*, 2016).

Effectors of *Phytophthora* spp.

The effector molecules secreted by oomycetes such as *Phytophthora* spp. can be divided into two broad classes depending on their localization in the host plant once they have been secreted. These effectors are either apoplastic, with their site of action in the extracellular space, or cytoplasmic, where they are translocated into the plant cells to target various subcellular compartments (Morgan & Kamoun, 2007). These two broad classes can then be further classified according to the effector's mode of action.

Apoplastic effectors can be categorized as inhibitors of host enzymes, hydrolytic enzymes, disruptors of plasma membrane-cell wall adhesion, and extracellular toxins (Stassen & Van den Ackerveken, 2011, Jiang & Tyler, 2012). Enzyme inhibitors are effectors that act as a counter-defense against plant defense proteins by inhibiting plant apoplastic enzymes such as chitinases, glucanases and proteases, contributing to pathogen virulence (Kamoun, 2006). Hydrolytic enzymes secreted by pathogenic oomycetes mainly function in degrading carbohydrates such as cellulose and xylans, which can result in necrosis in host tissues due to pathogen virulence. These enzymes are normally lacking in obligate biotrophs since killing host tissues will limit the growth and survival of biotrophic pathogens (Jiang & Tyler, 2012). A third group of apoplastic effectors aims to disrupt the adhesion and signaling between the plasma membrane and cell wall of host cells, which is thought to interfere with defenses associated with the host cell wall, and thus promotes pathogen virulence (Stassen & Van den Ackerveken, 2011). Finally, some apoplastic effectors may be extracellular toxins produced by necrotrophic or hemibiotrophic oomycetes, which trigger cell death in host plants in order to promote disease progression (Stassen & Van den Ackerveken, 2011).

In the case of cytoplasmic effectors of oomycetes such as *Phytophthora*, two classes of host-translocated effectors are known; these are the RxLR protein family and the family of Crinkler proteins (Kamoun, 2006). Crinkler proteins are a family of small secreted effectors, named for their ability to produce crinkling and necrosis during transient expression assays. These Crinklers also have conserved motifs in their N-termini, including a signal peptide, an LXLFLAK-motif, a conserved DWL-domain, and a HVLVXXP-motif (Stassen & Van den Ackerveken, 2011), and are some of the most highly

expressed pathogen genes before and during infection of host plants (Jiang & Tyler, 2012). The RxLR effectors are so named due to a conserved amino acid motif (Arginine-any amino acid-Leucine-Arginine) common to all known avirulence proteins of pathogenic oomycetes. The RxLR motif is believed to play a role in translocation of these effectors to host cell cytoplasm (Jiang & Tyler, 2012). Hundreds of putative RxLRs have already been identified in *Phytophthora* genomes, and functional characterization of those effectors has become the major focus of studies on the classification and characterization of oomycete effectors (Morgan & Kamoun, 2007).

The RxLR effectors of oomycetes

Host-targeting motifs of RxLR effectors

Several putative avirulence proteins, all with a common RxLR motif, have been identified in oomycete pathogens such as *Phytophthora* spp. (Jiang & Tyler, 2012). RxLR proteins are characterized by a modular structure (Figure 2), with two main functional domains (Kamoun, 2007). The N-terminal domains typically contain a signal peptide for secretion from the pathogen, followed by the RxLR conserved amino acid motif (Birch *et al.*, 2008). Variants of this RxLR motif have been found, where amino acid residues may differ from the traditional "R-x-L-R" sequence, yet still maintain their ability to translocate into host cells (Kale & Tyler, 2011). In several oomycete species causing downy mildew on diverse host plants it was found that the RxLR motif of specific effectors was replaced by variant amino acid sequences such as QxLR, RxLQ, and GKLR (Fabro *et al.*, 2011, Tian *et al.*, 2011, Stassen *et al.*, 2013). Kale and colleagues (2010) also showed that certain mutations in the RxLR motif of *P. sojae* Avr1b would not change the translocation ability of this effector. Following functional mutagenesis experiments of the RxLR motif of Avr1b, researchers proposed that a more flexible RxLR-like motif [R,K,H]x[L,I,M,F,Y,W]x should be considered when identifying RxLR effectors in plant pathogenic species (Kale *et al.*, 2010, Kale & Tyler, 2011).

Many of these RxLR effectors also contain a second, less conserved motif in this region, known as the dEER (Aspartate-Glutamate-Glutamate-Arginine) motif, found at varying distances downstream of the RxLR motif (Jiang *et al.*, 2008). This dEER motif may not adhere strictly to the d-E-E-R amino acid sequence, but typically consists of a series of acidic residues ending with arginine, such as the dEER motif of *P. sojae* Avr1b, which has the sequence EEDAGER (Dou *et al.*, 2008a, Tyler, 2011, Tyler *et al.*, 2013). Finally, the proteins contain a highly variable C-terminal domain which is associated with the effector activities inside host cells (Bozkurt *et al.*, 2012). Additional conserved motifs, which make up the WY domain, have been found within this C-terminal region of RxLRs, and these are known as the K, W, Y and L motifs, based on the highly conserved amino acids within each motif (Jiang *et al.*, 2008, Dou *et al.*, 2008b). These motifs are found in some, but not all, predicted RxLRs within *Phytophthora* genomes (Jiang *et al.*, 2008, Goss *et al.*, 2013, Wirthmueller *et al.*, 2013), and their presence has been linked to the suppression of programmed cell death (PCD) by these effectors (Dou *et al.*, 2008b, Wang *et al.*, 2011).

The RxLR-dEER twin peptide is believed to play a significant role in translocation of these RxLR effectors into plant host cells, based on various scientific studies (Stassen & Van den Ackerveken, 2011). This conserved motif was found to be similar in sequence and relative location to the host-targeting signal required for translocation of the effectors of the malarial parasite, *Plasmodium falciparum*. The malarial host-targeting signal is known as the PEXEL/HT motif, and contains a conserved RxLx amino acid sequence, which allows the transfer of malarial effectors into the red blood cells of host organisms (Hiller *et al.*, 2004, Marti *et al.*, 2004). The presence of similar conserved motifs in the effectors of such divergent pathogens suggests that both animal and plant pathogens contain similar signals for host-targeting of effector molecules, leading to the hypothesis that these divergent microbes may also have conserved mechanisms for transport of pathogen effectors into host cells (Bhattacharjee *et al.*, 2006, Kamoun, 2006).

Based on the discovery of these similar host-targeting signals found in both animal and plant pathogens, Bhattacharjee *et al*. (2006) conducted an experiment where the RxLR-containing leader sequence from the *P. infestans* effector Avr3a was fused to green

fluorescent protein (GFP) and its effects on translocation from *P. falciparum* to host erythrocytes was analysed. Fusion of the intact RxLR domain of Avr3a to GFP resulted in unaltered translocation of the effector from the parasite vacuole to host red blood cells, while a mutation in the RxLR region of Avr3a abolished effector transport (Figure 3). These results supported the hypothesis that RxLR-dEER motifs are responsible for translocation of *Phytophthora* effectors from pathogens to host cells (Bhattacharjee *et al.*, 2006).

Whisson *et al*. (2007) further investigated the role of the RxLR-dEER motif in translocation of *Phytophthora* effectors by investigating the secretion and transport of the Avr3a effector from *P. infestans* when the RxLR-dEER motif was replaced with alternative amino acid residues. In this study, it was found that the RxLR-dEER motif was not required for secretion of Avr3a from the haustoria of *P. infestans*, and the motif was not needed for targeting of the effector to haustorial membrane, since Avr3a effectors could still be secreted from haustoria when the RxLR-dEER motif was replaced with other amino acids, such as alanine or KMIK-DDK residues. Further experiments showed that replacement of the RxLR-dEER motif with other amino acids resulted in the inability of *P. infestans* to deliver Avr3a into plant cells. The effector, in absence of the RxLR-dEER motif, can be secreted by haustoria, but then accumulates in the extra-haustorial matrix without being translocated into host cells (Whisson *et al.*, 2007).

The role of the RxLR-dEER motif in translocation was further confirmed by Dou *et al.* (2008a) using the Avr1b effector of *P. sojae* who found that the RxLR-dEER domain from Avr1b, fused to GFP, was sufficient to allow the GFP fusion construct to enter plant host cells directly from the *Escherichia coli* cells in which it was expressed. When soybean root tips were incubated with the transgenic *E. coli* cells, the expressed GFP fusion construct that contained the RxLR-dEER leader sequence from Avr1b was successfully transduced to the plant host cells, even in the absence of the pathogen. In contrast, this translocation was unsuccessful when amino acid substitutions were introduced into either the RxLR or dEER motif of the leader sequence (Figure 4) (Dou *et al.*, 2008a). In addition, Dou *et al.* (2008a) were able to support the findings by Bhattacharjee *et al.* (2006) in reciprocal experiments, in which the RxLR-dEER motif of the Avr1b effector was replaced

with the host-targeting signal of *P. falciparum*. It was found that replacement of the Avr1b N-terminal region with the *Plasmodium* targeting signals resulted in successful translocation of the Avr1b effector to soybean host cells (Dou *et al.*, 2008a), thereby documenting the role of RxLR-dEER motifs in translocation of oomycete effectors (Birch *et al.*, 2008).

Mechanism of host cell entry

Despite the above evidence for the role of the RxLR motif in translocation of RxLR effectors, the actual mechanism of cell entry was only elucidated in later experiments by Kale and colleagues (2010). Researchers performed experiments to see whether the RxLR-dEER domains of effectors could bind to phosphoinositides (phosphorylated phosphatidylinositol biomolecules) – a group of lipids that is common to all eukaryotic cells and are known to contribute to signaling within and between cells, regulate membrane trafficking, and play a role in the functioning of cellular cytoskeletons (Odorizzi *et al.*, 2000, Di Paolo & De Camilli, 2006, Shewan *et al.*, 2011). Kale *et al.* (2010) designed their experiments based on the fact that phosphatidylinositol-4-phosphate (PI4P) domains in rice and *Arabidopsis* PI4P kinases were shown to be made up of tandem repeats containing motifs reminiscent of RxLR-dEER motifs (Lee *et al.*, 2006). Kale *et al*. (2010) used fusions of GFP to the N-terminal regions of *P. sojae* RxLR effectors Avr1b, Avh331 and Avh5, or to full-length Avr1b, to probe membrane-spotted lipids. They found that full-length Avr1b- and N-terminal Avh331 fusions bound to both phosphatidylinositol-3-phosphate (PI3P) and PI4P. The N-terminal Avh5 fusion, meanwhile, bound predominantly to PI3P while the N-terminal Avr1b fusion bound predominantly to PI4P. Researchers hypothesised that differences in Avr1b fusions binding to these phospholipids was likely due to structural changes in the N-terminal region when the rest of the protein was missing (Kale *et al.*, 2010). Kale and colleagues (2010) suggested that the binding of these RxLR effectors to the membrane-associated phosphoinositides PI3P and PI4P may play a role during effector entry into host cells.

Subsequently, researchers substituted amino acids in the RxLR-dEER motif identical to mutations that were previously shown to abolish effector entry into soybean cells (Dou *et al.*, 2008a), and these mutations also abolished binding to phosphoinositides (Kale *et al.*, 2010). To further illuminate which phospholinositides were responsible for RxLR entry into host cells, researchers tested for the presence of PI3P and PI4P on the outer surface of both plant and animal cells by creating specific biosensors for both phospholipids. Results of these tests indicated that PI3P, but not PI4P, can be found on the outer plasma membrane of host cells, supporting the hypothesis that binding of RxLR-dEER motifs to PI3P is responsible for mediating effector entry into plant cells. Further experiments using excess of PI3P- and PI4P-binding proteins in incubations with soybean roots and effector-GFP fusions showed that the PI3P-binding proteins (which inhibit binding of the RxLRs by competing for binding sites) abolished effector entry into root cells, whereas PI4P-binding proteins did not affect cell entry of the effectors, again supporting the role of PI3P, but not PI4P, in cell entry of RxLR effectors (Kale *et al.*, 2010).

Finally, researchers used specific inhibitors of several types of endocytosis, as well as studies of effector localisation shortly after exposure to host cells, to determine the role of endocytosis in effector uptake. Results supported their hypothesis that RxLR effectors enter host cells by a type of endocytosis mediated by lipid rafts - lipid microdomains within the cell plasma membrane that contain increased concentrations of sphingolipids and cholesterol (Pike, 2003, Raven *et al.*, 2014). A visual representation of this proposed model for cell entry of RxLRs can be seen in Figure 5. PI3P mediated entry of RxLRs was further confirmed by Sun and colleagues (2013) for the *P. sojae* effector, Avh5.

Evidence has emerged to dispute the idea that the N-terminal RxLR motif is needed for PI3P-mediated cell entry, since Yaeno *et al.* (2011), Wawra *et al.* (2012) and Sun *et al*. (2013) showed that PI3P-binding was mediated by the C-terminus of *P. infestans* AVR3a and *P. sojae* Avh5 and Avr1b. Indeed, some researchers suggested that PI3P-binding of RxLR effectors is physiologically irrelevant, since *P. infestans* AVR3a only binds to Pi3P when it is denatured or not fully stabilized (Wawra *et al.*, 2012)**.**

A recent breakthrough study has provided a possible model to explain the puzzling results of these cell-entry experiments. In this study, researchers used NMR spectroscopy and native AVR3a proteins obtained from culture filtrates of *P. infestans* to show that this RxLR effector is, in fact, cleaved at the RxLR motif prior to secretion from the oomycete (Wawra *et al.*, 2017). This model, represented visually in Figure 6, proposes that the RxLR motif of this family of effectors is cleaved off by intracellular *Phytophthora* proteases, allowing the proteins to be acetylated within the cell before they are secreted from pathogen haustoria. This new hypothesis challenges the previously accepted beliefs about how RxLRs are translocated into host cells, such as the original hypothesis that the RxLR motif binding to Pi3P mediates cell entry (Kale *et al.*, 2010). Instead, the data of Wawra *et al*. (2017) indicate that the RxLR motif may have a role in processing of the effectors to allow secretion from the pathogen, rather than in uptake by plant cells. These findings also provide a possible explanation for the contrasting results of Pi3P-binding and cell entry experiments listed above, since the model explains why RxLR N-terminal deletion or substitution would abolish cell entry, while still taking into account the likelihood that C-terminal binding to PI3P may mediate effector entry into host cells (Wawra *et al.*, 2017). It is clear, however, that more studies are needed to test this hypothetical model for other RxLR effectors so that the mechanisms for secretion and cell uptake can be fully understood.

Furthermore, it is not currently understood how RxLRs are released from cell endosomes after they have been taken up into plant cells, although it was suggested by Tyler (2011) that the effector may directly cross the endosomal membrane following partial denaturation of the effector proteins. The conflicting results of translocation and PI3P binding studies, as well as the lack of understanding regarding endosome escape of effectors, show that more work is needed to investigate the specific mechanism of RxLR effector entry into host cells.

Identification of *RxLR* effector genes

Traditional identification of effector genes

The discovery of RxLR effectors and their predictable protein structures has enabled the use of bioinformatics analyses to predict large collections of candidate effector genes in *Phytophthora* species (Oh *et al.*, 2009). However, traditional identification of pathogenic effectors was based on biochemical purification of expressed proteins and subsequent genetic analysis.

One example of *RxLR* effector genes discovered by traditional methods is *Avr1a* and *Avr3a/5* from *P. sojae.* The genes encoding these effectors were first identified based on outcrossing experiments and subsequent linkage analysis using random amplified polymorphic DNA (RAPD) markers (Whisson *et al.*, 1994). In this analysis, segregation of virulence against soybean resistance genes *Rps1a*, *Rps3a* and *Rps5* was linked to segregation of RAPD markers in *P. sojae* F_1 hybrids and F_2 individuals, and it was found that pathogen avirulence was dominant to virulence against these resistance genes. Genes responsible for avirulence were named *Avr1a*, *Avr3a* and *Avr5,* though linkage data suggested that the latter two avirulence genes require the same dominant effector gene (with independent segregation) for avirulence on soybean (Whisson *et al.*, 1994).

Avr1a has since been genetically mapped based on how DNA markers and virulence on *Rps1a* segregate in two separate F_2 populations (MacGregor *et al.*, 2002), and it has been cloned and its copy number variations analysed (Qutob *et al.*, 2009). *Avr3a* and *Avr5* have been identified as alleles of the same effector gene, as they are encoded by the same genetic locus (Dong *et al.*, 2011a), which has since become known as *Avr3a/5*, since this single locus is responsible for conferring avirulence against both *Rps3a* and *Rps5* resistance genes. This effector has also been cloned and its functions in avirulence tested in soybean (Qutob *et al.*, 2009, Dong *et al.*, 2011a).

In *P. infestans*, avirulence genes were identified by similar methods. *Avr2* was identified following segregation ratio analysis of avirulence against *P. infestans* resistance genes (Spielman *et al.*, 1990). In original experiments, Spielman *et al.* concluded that avirulence against resistance gene *R2* was controlled by a single locus, but at this time no DNA markers had been developed to isolate the avirulence gene. After investigating

segregation of avirulence phenotypes in *P. infestans* linked to amplified fragment length polymorphism (AFLP) markers, van der Lee *et al.* (2001) named the responsible gene *Avr2*. The AFLP markers were subsequently used to isolate the gene from the bacterial artificial chromosome (BAC) library of *P. infestans* T30-4, enabling identification of *Avr2* by map-based cloning (Whisson *et al.*, 2001, Gilroy *et al.*, 2011a). *Avr4*, also discovered by van der Lee *et al.* (2001), was similarly identified, but gene sequence identity was later confirmed by a combination of map-based cloning and transcriptional profiling of cDNA-AFLP markers (van Poppel *et al.*, 2008).

These traditional methods of isolating *RxLR* effector genes were thus very effective as an identification strategy, but involved laborious genetic experiments and allowed isolation of only a few genes at a time. In the above-mentioned cases, it is important to note that identification of pathogen avirulence genes was only possible due to resistance genes already discovered in host plants, which limits the applicability of these studies to host-pathogen systems with known host *R* genes. With advances in functional genomics, identification of genes encoding putative effectors has become more straightforward, enabling high-throughput screening of candidate effectors and the identification of pathogen effectors in the absence of known host *R* genes.

Genome-wide identification of effector genes

In recent years, whole genomes of various *Phytophthora* species have become available (Tyler *et al.*, 2006, Haas *et al.*, 2009, Lamour *et al.*, 2012, Quinn *et al.*, 2013, Gao *et al.*, 2015, Ali *et al.*, 2017, Tabima *et al.*, 2017, Vetukuri *et al.*, 2018). With whole genome sequences, candidate effector genes, such as *RxLR*s, can be identified by using data mining tools (Grünwald, 2012b) to find sequences that have specific attributes, such as the presence of particular domains. Putative effector genes identified by bioinformatics can then be further analysed by functional studies to characterize effector proteins (Kamoun, 2006).

By using the genomes of certain *Phytophthora* species, several hundred putative RxLR effectors have been identified (Jiang & Tyler, 2012). The sequenced genomes of *P. infestans*, *P. sojae* and *P. ramorum* have, in fact, been found to each contain approximately 350 to 600 genes encoding putative RxLR effectors, although the number of candidate effectors identified depends largely on the bioinformatic criteria used in the analyses (Tyler *et al.*, 2006, Win *et al.*, 2007, Jiang *et al.*, 2008, Haas *et al.*, 2009).

Win *et al.* (2007), for example, identified 531 candidate *RxLR* effector genes in *P. ramorum*, and 672 in *P. sojae,* using an algorithm they developed based on features of 13 known RxLR proteins and 30 other sequences that were significantly homologous to these validated RxLRs. When researchers eliminated candidate RxLRs without the dEER motif and applied a Hidden-Markov Model (HMM) to the alignment of the sequences at the RxLR-dEER region, they ended up with 181 candidate RxLRs in the *P. ramorum* and 158 in the *P. sojae* genomes. Jiang *et al.* (2008), in contrast, identified 375 RxLRs in the genome of *P. ramorum*, and 396 in *P. sojae*, using a combination of recursive BLAST (using Avr1b protein sequence as the initial query) and subsequent HMM analysis. These results were similar to a previous study by Tyler *et al.* (2006), where 350 *RxLR* genes were predicted in each of the genomes of *P. sojae* and *P. ramorum,* although this study used sequences of the four *Avr1b-1* alleles, an *Avr1b-1* paralog (renamed *Avh1*) and *Avr3a* as initial queries in recursive BLAST prior to HMM analysis.

In *P. infestans*, a maximum of 425 *RxLR* genes were originally predicted from the draft genome by Whisson *et al.* (2007) using a combination of bioinformatics approaches. They identified 38 candidate *RxLR* genes from the draft genome using a heuristics approach, all of which had been shown to be expressed during different stages of infection. Whisson and colleagues (2007) then used an alignment of these 38 genes to develop an HMM, which was used to search the earliest genome assembly, resulting in the prediction of 284 *RxLR* genes. Researchers also used a method used earlier by Bhattacharjee *et al.* (2006) to predict 310 putative *RxLR*s. These approaches predicted a combined total of 425 *RxLR*s in the draft genome of *P. infestans*, 169 of which were common to gene sets predicted by both approaches and were therefore redundant (Whisson *et al.*, 2007).

Later, the assembly of the complete genome allowed Haas *et al.* (2009) to predict 563 *RxLRs* in *P. infestans* using a combination of an HMM profile for the RxLR motif, Regex programming to search for specific residue patterns, and the algorithm used by Win *et al*. (2007). The HMM and Regex methods each predicted a set of 395 candidate *RxLRs*, with an overlap of 326 genes between the methods. Using the Win (2007) algorithm, with selected additional criteria applied, Haas *et al.* (2009) identified 488 putative *RxLR* genes. After additional examination of the genome and manual alterations to the list of effector candidates, a final total of 563 was predicted.

The method used by Whisson (2007) to identify *RxLRs* in *P. infestans* was more recently also used by Reitmann *et al*. (2017) to predict candidate *RxLR* genes in *P. cinnamomi*, who found that 44 effector genes containing an RxLR motif could be detected in the transcriptome of pre-infection stages of the pathogen. In a separate analysis, Hardham & Blackman (2017) predicted 171 candidate *RxLR* genes in *P. cinnamomi* with pBLAST searches of the genome. They used 122 of the candidate *P. infestans RxLR*s predicted by Haas *et al.* (2009) to query the *P. cinnamomi* genome, which resulted in the detection of 340 effector genes with a confirmed signal peptide in the N-terminal region. However, manual inspection of the protein sequences showed that only 171 of these contained an RxLR and dEER motif (Hardham & Blackman, 2017). In recent analysis of other species, 336 candidate *RxLRs* were identified in *P. megakarya*, and 414 in *P. palmivora*, based on PSI-BLAST using candidate *RxLRs* from *P. sojae* and *P. ramorum*, combined with HMM analysis (Ali *et al.*, 2017).

It is important to note that these catalogues of candidate *RxLRs* are thus far only putative annotations, and further analysis and functional characterization of these genes is crucial before their functions as pathogen effector molecules can be validated. Validation and annotation of these genes will provide a better estimate of the number of functional *RxLR* effector genes that occur in different *Phytophthora* genomes.

Functional characterization of effector genes

Initial expression assays for characterisation

Hundreds of putative *RxLR* effector genes have already been predicted by bioinformatic analysis of *Phytophthora* genomic sequences (Tyler *et al.*, 2006, Win *et al.*, 2007, Jiang *et al.*, 2008, Haas *et al.*, 2009, Ali *et al.*, 2017), but the next challenge is to assign biological functions to these genes and validate their roles as pathogen effectors (Bhadauria *et al.*, 2009). Functional analyses are crucial in understanding the biochemical activities of these proteins and their roles in plant disease. Characterization of well-studied *RxLRs* to date, such as *Avr1b* and *Avr3b* from *P. sojae* and *Avr3a* from *P. infestans,* has relied extensively on *in planta* expression studies to obtain preliminary information regarding possible function (Kamoun, 2006).

For example, the RxLR effector Avr1b was found to result in increased virulence when it was overexpressed in *P. sojae* transformants used to infect soybean (*Glycine max*) seedlings. The *Avr1b-1* effector gene was also transiently expressed in infiltrated soybean and *Nicotiana benthamiana* together with the mouse BAX protein that elicits apoptosis. It was therefore deduced that the effector has a role in suppressing PCD during host infection (Dou *et al.*, 2008b).

Avr3b, another RxLR effector from *P. sojae*, was characterized by transient expression in *N. benthamiana*, where it was shown to increase host susceptibility to *P. capsici* and *P. parasitica*, and resulted in reduced accumulation of reactive oxygen species (ROS) at the sites of invasion (Dong *et al.*, 2011b). ROS accumulation is important in initiating defense responses in the host, since ROS can act as signaling molecules in host defense networks, but can also directly trigger cell death of the host or pathogen (Torres *et al.*, 2006, Tsuda & Katagiri, 2010). Further characterization of the *P. sojae* Avr3b effector was achieved by a biochemical analysis of the protein based on specific sequence motifs found to be present in the *Avr3b* region, and this analysis confirmed that Avr3b is an ADP-ribose/NADH pyrophosphorylase (Dong *et al.*, 2011b), a type of molecule implicated in the negative regulation of host immunity to certain pathogens (Bartsch *et al.*, 2006).

The RxLR effector Avr3a from *P. infestans* was initially characterized by its co-expression with the associated host *R*-gene from potato, *R3a*, in *N. benthamiana*. When co-expressed, interaction of the Avr3a effector protein with the R3a resistance protein caused a HR in *N. benthamiana* (Figure 8) (Armstrong *et al.*, 2005), characterized by localized cell death around the sites of agroinfiltration. In the context of plant defense, HR minimizes damage by the pathogen, since rapid cell death at the site of infection prevents an invading biotrophic pathogen from spreading through healthy tissues to the rest of the host plant (Raven *et al.*, 2014).

In a subsequent study by Bos *et al*. (2006), overexpression of the Avr3a effector in *N. benthamiana* infected with elicitin INF from *P. infestans* suppressed cell death typically caused by this toxin. Thus, expression of this RxLR in *N. benthamiana* helped elucidate its role in host-cell mediated cell death.

Further elucidation of effector functions

Although initial expression studies have helped understand the roles of some RxLR effectors in virulence for a generous set of these RxLR effectors, the activities of these effectors at the cellular level needs further study. For example, the researchers who discovered the INF-cell death suppression function of *P. infestans* Avr3a, subsequently elucidated the molecular mechanisms of this suppression activity (Bos *et al.*, 2010). They used yeast two-hybrid assays to identify potential host targets that interacted with the Avr3a effector protein. It was found that Avr3a associates with the potato protein CMPG1, a ubiquitin E3 ligase which functions in activation of host cell death upon recognition of various pathogen elicitors, including INF1 (González-Lamothe *et al.*, 2006). Bos and colleagues (2010) determined that Avr3a suppresses host cell death by stabilizing CMPG1 during this interaction, modifying its activity by preventing self-degradation. This stabilization thus prevents initiation of PCD as part of the host's defense responses during the biotrophic phase of infection by the pathogen (Bos *et al.*, 2010). Further investigations revealed that Avr3a could also suppress CMPG1-dependent cell death triggered by the interactions of *Cf-9/Avr9*, *Cf-4/Avr4* and *Pto/AvrPto*, as well as by the *Phytophthora* cellulose-binding elicitor lectin (CBEL). However, cell death triggered by the host NBS-

LRRs R3a, Rx and R2 did not require CMPG1, and was therefore not suppressed by Avr3a (Gilroy *et al.*, 2011b).

Recently, Chaparro-Garcia *et al.* (2015) found that Avr3a can interfere with plant immunity independent of CMPG1-targetting. They found that Avr3a has an additional role in virulence, where the effector interferes with internalization of immune receptors by associating with the host Dynamin-Related Protein DRP2. Although they did not determine how the association took place, they did determine that it disrupted signaling cascades necessary for PTI responses in host plants (Chaparro-Garcia *et al.*, 2015).

Another *P. infestans* RxLR effector, Avrblb2, was also shown to interfere with host defense through its interaction with a plant host protein. Bozkurt *et al.* (2011) used a combination of co-immunoprecipitation and liquid chromatography-tandem mass spectrometry to identify *N. benthamiana* proteins that interacted with the Avrblb2 effector *in planta*. They chose the C14 cysteine protease for further study. Cysteine proteases, such as C14, had been previously implicated in plant defense (Shindo & Van Der Hoorn, 2008, Kaschani *et al.*, 2010), and as such serve as common host targets for filamentous pathogen effectors (Shabab *et al.*, 2008, Song *et al.*, 2009, Kaschani *et al.*, 2010). Bozkurt and colleagues (2011) confirmed the association of Avrblb2 with C14 *in planta* and found that this interaction leads to modification of the subcellular localization of C14. Further experiments showed that Avrblb2, in the absence of its cognate R protein Rpi-blb2, specifically prevents the secretion of C14 cysteine protease to its site of action, the cell apoplast, thereby inhibiting the defense responses of solanaceous plants to *P. infestans* (Bozkurt *et al.*, 2011).

The molecular functions have been elucidated for several other RxLRs in *P. infestans*. For example, PexRD2 was shown to interact with a MAP kinase involved in host defense, thereby obstructing signaling pathways to weaken plant immunity (King *et al.*, 2014). The RxLR effector Pi03192 was found to promote disease by modifying the subcellular localization of two NAC transcription factors by interacting with these host targets (McLellan *et al.*, 2013). Pi04314 was shown to play a role in infection through its interaction with protein phosphatase 1 catalytic (PP1c) isoforms, which resulted in their re-localisation from the nucleolus to the nucleoplasm and subsequently interfered with

induction of immune signaling pathways (Boevink *et al.*, 2016). Thus, different *P. infestans* RxLRs were shown to modulate host defenses by interacting with various host targets.

Qiao *et al.* (2013) were the first to identify RxLR effectors that inhibited RNA-silencing. They screened 59 of the RxLR effectors originally predicted from the *P. sojae* genome by Tyler *et al.* (2006) for suppression activity when co-expressed with *GFP* in *N. benthamiana* line 16c, which constitutively expresses *GFP*. When exogenous *GFP* was infiltrated into leaves of 16c, small interfering RNAs (siRNAs) were produced to silence expression of both endogenous and exogenous *GFP* genes. Effectors with RNA-silencing suppression activity that are co-expressed in this system would then interfere with silencing of *GFP*, allowing their functions to be pinpointed. With this system, two of the 59 RxLRs screened by Qiao *et al.* (2013) were positively identified as having RNA-silencing suppression activity. These two genes, designated *PsAvh18* and *PsAvh146* upon initial identification from the genome (Tyler *et al.*, 2006) were then renamed *PSR1* (*Phytophthora* suppressors of RNA silencing 1) and *PSR2*, respectively. Their expression in transgenic *Arabidopsis* was shown to reduce levels of both small interfering RNAs (siRNAs) and microRNAs (miRNAs).

Upon further investigation, Qiao *et al.* (2013) determined that the inhibition of silencing by PSR1 was not due to binding of the effectors to small RNAs, but rather due to interference with the DICERLIKE1-mediated processing of primary miRNAs and precursor miRNAs to mature miRNA transcripts. PSR1 was also shown to affect biogenesis of siRNAs. Qiao *et al.* (2015) later found that PSR1 carried out these functions by directly targeting a host protein designated PSR1-Interacting Protein 1 (PINP1) which, in *Arabidopsis*, regulates miRNA and endogenous siRNA accumulation. PSR2, however, did not display the same broad inhibition of small RNAs as PSR1, and instead specifically controlled levels of two trans-acting siRNAs by a different mechanism, which were thought to be involved in suppression of host NBS-LRR defense genes. Both *PSR1* and *PSR2,* when expressed *in planta* were shown to enhance susceptibility of plants to *Phytophthora* and potato virus X infection (Qiao *et al.,* 2013).

Another *P. sojae* RxLR effector that contributes to plant infection is Avr3b, identified by Dong *et al.* (2011) using a combination of genetic and bioinformatic analyses. Examination of the coding sequence for this effector revealed the presence of a Nudix hydrolase motif. This motif is characteristic of a family of enzymes known as pyrophosphatases, which are produced by organisms in all domains of life (Mildvan *et al.*, 2005). Researchers used biochemical assays to prove that Avr3b had pyrophosphatase activity, using ADP-ribose and NADH as substrates (Dong *et al.*, 2011b). It had previously been proposed that such ADP-ribose/NADH pyrophosphorylases play a role in the negative regulation of plant defense against bacteria and oomycetes (Ge *et al.*, 2007). Dong *et al.* (2011b) speculated that the Avr3b effector of *P. sojae* may mimic the action of these host enzymes to modulate plant defense through the suppression of ROS accumulation during infection.

Other *P. sojae* RxLR effectors have been assigned functions over the years. As with *P. infestans* PexRD2 (King *et al.*, 2014), *P. sojae* Avh331 was found to impair host defenses by interrupting MAP kinase signaling pathways, although it was proposed to act downstream of the MAP kinase cascade since it did not interact directly with MAP kinases in biochemical assays (Cheng *et al.*, 2012). Avr3c was shown to disrupt plant immunity by binding to the host SKRP proteins, allowing the manipulation of alternative splicing processes involved in plant defense (Huang *et al.*, 2017) and Avh262 stabilized host endoplasmic reticulum-luminal binding immunoglobulin proteins (BiPs) that are negative regulators of plant immunity. This interaction results in suppression of PCD during infection to subvert host defense responses (Jing *et al.*, 2016).

Another strategy used by RxLRs to subvert host immunity was revealed in *P. capsici* by Li *et al.* (2019a). In this study, researchers used a yeast two-hybrid assay to screen *P. capsici* RxLR effectors using the host protein, non-expressor of pathogenesis related-1 (NPR1), as bait. They found that NPR1 interacts with *P. capsici* RxLR48, serving as its virulence target in host plants. In subsequent experiments, Li *et al.* (2019a) proved that in binding to NPR1, RxLR48 also promotes the nuclear localization and accumulation of this host protein and prevents its degradation by a 26S proteasome. This finding is significant, since the degradation of NPR1 is required for the host to mount an effective

salicylic acid-mediated defense response (Spoel *et al.*, 2009). Increased levels of cellular NPR1 have also been found to decrease HR in host plants, and so by altering the distribution ratio and degradation patterns of NPR1, RxLR48 would ultimately undermine the host's ability to combat hemibiotrophic pathogens such as *P. capsici* (Spoel *et al.*, 2009, Backer *et al.*, 2019, Li *et al.*, 2019a).

Clearly, *Phytophthora* species have evolved diverse mechanisms to interfere with host defenses through the use of RxLR effectors. Despite their different modes of action, they all play a role in combating plant defense responses. Better understanding of the functions of RxLR effectors are needed, as they could ultimately help improve strategies for managing diseases caused by these plant pathogens.

Application of effector knowledge

Identifying and understanding effector genes is an important step in developing disease management strategies for plant pathogens. Effector genes can be specifically targeted through gene silencing. Sanju *et al.* (2015) demonstrated that transgenic potato lines displayed partial resistance to late blight when an RNA interference (RNAi) construct was introduced that allowed moderate host-induced silencing of *P. infestans* effector gene *Avr3a* (Sanju *et al.*, 2015). Known effectors can therefore provide targets for control using host-induced gene silencing, but further research is needed to determine whether such control is sufficiently effective and durable.

Knowledge of effector genes mainly contributes to disease management strategies by allowing the identification of interacting R proteins, which can then be exploited in resistance breeding or marker-assisted selection of crop cultivars (Vleeshouwers *et al.*, 2008, van Damme *et al.*, 2011). In one study, Vleeshouwers *et al.* (2008) used computationally predicted RxLR effectors to screen for activation of avirulence activity on wild *Solanum* species, which allowed the isolation of *Rpi-sto1* from *Solanum stoloniferum* and *Rpi-pta1* from *Solanum papita*. These resistance genes were found to be functionally equivalent to a previously cloned *R* gene, *Rpi-blb1*. Studies such as this facilitate the accelerated cloning of *R* genes needed for resistance breeding (Schornack *et al.*, 2009),

and also help identify *R* gene homologs in sexually compatible material that can be used in selective breeding programs (Vleeshouwers *et al.*, 2008, Lokossou *et al.*, 2010, van Damme *et al.*, 2011).

The use of effectors to find their cognate *R* genes can also help to avoid redundant cloning efforts, so that only *R* genes with different activities are used in resistance breeding efforts (Schornack *et al.*, 2009). Four *R* genes from *Solanum* species were cloned and characterized by Lokossou *et al*. (2009). *Rpi-blb3*, *Rpi-abpt*, *R2* and *R2-like,* were isolated separately, but each recognized the *P. infestans* Avr2 effector (Lokossou *et al.*, 2009). In cases such as this, cloning and exploitation of only one of the cognate R proteins would be sufficient to confer resistance in the presence of Avr2, and efforts to clone separate *R* genes with the same functionality are not needed.

Finally, identified *R* genes can be altered to expand their recognition specificity, conferring resistance in the presence of additional effectors without characterized interacting R proteins (Vleeshouwers *et al.*, 2011). Modification of existing *R* genes by random mutagenesis can enable recognition of novel virulence alleles of pathogen effector genes, even before pathogens evolve new virulence effectors (van Damme *et al.*, 2011, Vleeshouwers *et al.*, 2011). For example, Segretin *et al.* (2014) modified the potato *R3a* resistance gene with single amino acid substitutions, which enabled recognition of new virulence variants of *P. infestans Avr3a* and subsequent host immune responses.

Identification and characterization of pathogen effectors is therefore especially valuable in providing improved strategies for disease management. This knowledge can be used to provide targets for silencing of effector genes, or aid in the characterization of cognate *R* genes for improved resistance breeding (Vleeshouwers *et al.*, 2011).

Conclusion

Phytophthora species are destructive plant pathogens which have become the focus of research on oomycete phytopathogens. Of particular interest is how these pathogens use effector molecules, such as the RxLRs, to undermine plant defense responses, as they are thought to play a crucial role during the infection process. Recent availability of whole

genomes of *Phytophthora* has allowed many putative RxLRs to be identified, which can then be further analysed for functionality during host infection. Thus, genome sequencing and subsequent data mining provides opportunities to streamline research in effector biology, enabling accelerated identification, isolation and characterization of pathogen effectors. This knowledge can then be used to identify interacting host proteins, which can be applied to improve disease management strategies for *Phytophthora* plant pathogens.

Tables

Table 1. Examples of RxLRs with characterized functions in other *Phytophthora* species.

Species	RxLR Effector	Cognate R Protein (Host isolated from)	Direct Host Target	Proposed Function	References
P. infestans	PiSNE1 (*Phytophthora* suppressor of necrosis 1)	Not identified	Unknown	Suppresses plant cell necrosis during the biotrophic growth phase of the pathogen	(Kelley *et al.*, 2010)
	IPI-O1/Avrblb1	Rpi-Blb1(RB) (*Solanum bulbocastinum*) Rpi-sto1 (*Solanum stoloniferum*) Rpi-pta1 (*Solanum papita*)	LecRK-I.9 (lectin receptor kinase I.9)	Disrupts adhesions between cell wall and plasma membrane through binding of an RGD tripeptide motif, which interferes with host defense responses	(Senchou *et al.*, 2004, Bouwmeester *et al.*, 2011, Chen *et al.*, 2012)
	Avrblb2	Rpi- Blb2 (*S. bulbocastinum*)	C14 cysteine protease	Associates with and prevents the secretion of C14 to interfere with its role in host defense	(Bozkurt *et al.*, 2011)
	PiAvr3a	R3a (*Solanum tuberosum*)	CMPG1 ubiquitin E3 ligase	Stabilises CMPG1 to prevent activation of cell death during biotrophic phase of infection	(Bos *et al.*, 2010, Gilroy *et al.*, 2011b)
	PexRD2	Not identified	MAPKKKε (MAPK kinase kinase ε)	Associates with the kinase domain MAPKKKε to interfere with host immune signaling pathways and undermine host defense	(King *et al.*, 2014)
	Pi03192	Not identified	NTP (NAC Targeted by *Phytophthora*) 1 & NTP2	Interacts with the NAC transcription factors NTP1 and NTP2 to prevent their re-localisation from the ER to the	(McLellan *et al.*, 2013)

				nucleus, suppressing their role in plant defense	
	Pi04314	Not identified	PP1c-1 (protein phosphatase type 1 catalytic-1), PP1c-2 & PP1c-3	Interacts with PP1c isoforms in host to modify their localization and interfere with downstream immune signaling pathways	(Boevink *et al.*, 2016)
	PiPSR2 (*Phytophthora* suppressor of RNA-silencing 2)	Not identified	Unknown	Suppresses RNA-silencing and promotes *Phytophthora* infection of host plants	(Xiong *et al.*, 2014, de Vries *et al.*, 2017)
	PexRD54	Not identified	ATG8CL (autophagy-related cargo receptor)	Interferes with autophagy-related host defense by outcompeting Joka2, a host cargo receptor, for binding to autophagy-related protein ATG8CL	(Dagdas *et al.*, 2016)
	Pi22926	Not identified	MAPKKKβ2 (MAPK kinase kinase β2)	Interacts with MAPKKKβ2 to disrupt defense signaling pathways in host	(Ren *et al.*, 2019)
	Pi14054	Not identified	Unknown	Suppresses RNA-silencing in host plants to undermine plant defenses	(Vetukuri *et al.*, 2017)
	Pi17316	Not identified	StVIK (*S. tuberosum* VASCULAR HIGHWAY1-interacting kinase)	Binds to the host MAPKKK protein StVIK to interfere with immune signaling and suppress plant defense	(Murphy *et al.*, 2018)
P.sojae	PsPSR1	Not identified	PINP1 (PSR1-interacting protein 1)	Interacts with nuclear protein PINP1 to interfere with its role in accumulation of small RNAs, thereby inhibiting RNA-silencing by host plants to attenuate defense responses	(Qiao *et al.*, 2013, Qiao *et al.*, 2015)
	PsPSR2	Not identified	Unknown	Suppresses RNA-silencing and promotes infection of host plants, likely interferes with SA-dependent defense responses	(Qiao *et al.*, 2013, Xiong *et al.*, 2014)

	PsAvr1b-1	Rps1b (*Glycine max*) Rps1k (*G. max*)	Unknown	Suppresses programmed cell death in host cells during *in planta* expression studies	(Song *et al.*, 2013)
	PsAvr3b	Rps3b (*G. max*)	Unknown	Negatively regulates plant immunity by mimicking host enzymes through its ADP-ribose/NADH pyrophosphatase activity	(Dong *et al.*, 2011b)
	Avh331 (PsAvr1k)	Rps1k (*G. max*)	Unknown	Interferes with MAPK signaling pathways to suppress host basal defense responses	(Cheng *et al.*, 2012)
	PsAvr3c	Rps3c (*G. max*)	GmSKRPs (*G. max* serine/lysine/arginine-rich proteins)	Binds to SKRPS to modify alternative splicing processes involved in host defense responses	(Huang *et al.*, 2017)
	Avh262	Not identified	BiPs (binding immunoglobulin proteins)	Stabilises BiPs involved in ER-stress related cell death to suppress defense responses	(Jing *et al.*, 2016)
	Avh23	Not identified	ADA2 (alteration/deficiency in activation 2)	Competitively binds to the ADA2 subcomplex of a histone acetylation complex, reducing histone acetyltransferase activity and suppressing host defense gene expression	(Kong *et al.*, 2017)
	Avh238	Not identified	Type2 GmACSs (*G. max* 1-aminocyclopropane-1-carboxylate synthases)	Disturbs phytohormone defense pathways by destabilizing soybean Type2 ACSs - enzymes involved in ethylene biosynthesis - resulting in suppressed host immune responses	(Yang *et al.*, 2017b, Yang *et al.*, 2019)
	Avh240	Not identified	GmAP1 (*G. max* aspartic protease 1)	Associates with host GmAP1 to inhibit its secretion to cell apoplast, disrupting its function in host immune response	(Guo *et al.*, 2019)
P. parasitica	PSE1 (penetration-	Not identified	Unknown	Manipulates plant hormone physiology by decreasing cellular	(Evangelisti *et al.*, 2013)

	specific effector 1)			concentrations of auxins, and suppresses localized cell death in hosts	
	PpRxLR2	Not identified	Unknown	Enhances infection by pathogen by overcoming programmed cell death immune responses	(Dalio *et al.*, 2018)
P. capsici	RxLR48	Not identified	NPR1 (non-expressor of pathogenesis related-1)	Interacts with NPR1 to interfere with the protein's localization and degradation, resulting in compromised plant immunity, and interferes with PTI signaling for first layer of plant defense	(Li *et al.*, 2019a)
	PcAvr3a12	Not identified	FKBP15-2 (FK506 binding protein 15-2)	Targets FKBP15-2, a host peptidyl-prolyl isomerase, to suppress its function in ER-stress related host defense responses	(Fan *et al.*, 2018)
	PcAvh1	Not identified	PP2Aa (protein phosphatase 2A)	Binds to the scaffolding subunit of host PP2Aa to interfere with plant immunity	(Chen *et al.*, 2019)
	RxLR207	Not identified	BPA1 (Binding partner of ACD11)	Interacts with BPA1 to destabilise its binding partner ACD11, promoting ROS-mediated cell death to progress necrotrophic phase of infection	(Li *et al.*, 2019b)

Abbreviations: ACD11, Arabidopsis accelerated cell death 11; ADP, adenosine diphosphate; ER, endoplasmic reticulum; MAPK, mitogen-activated protein kinase; NAC transcription factors, NAM/ATAF/CUC transcription factors; NADH, nicotinamide adenine dinucleotide + hydrogen; PTI, pattern triggered immunity; R protein, resistance protein; RGD motif, Arginine-glycine-aspartate motif; ROS, reactive oxygen species; RNA, ribonucleic acid.

Figures

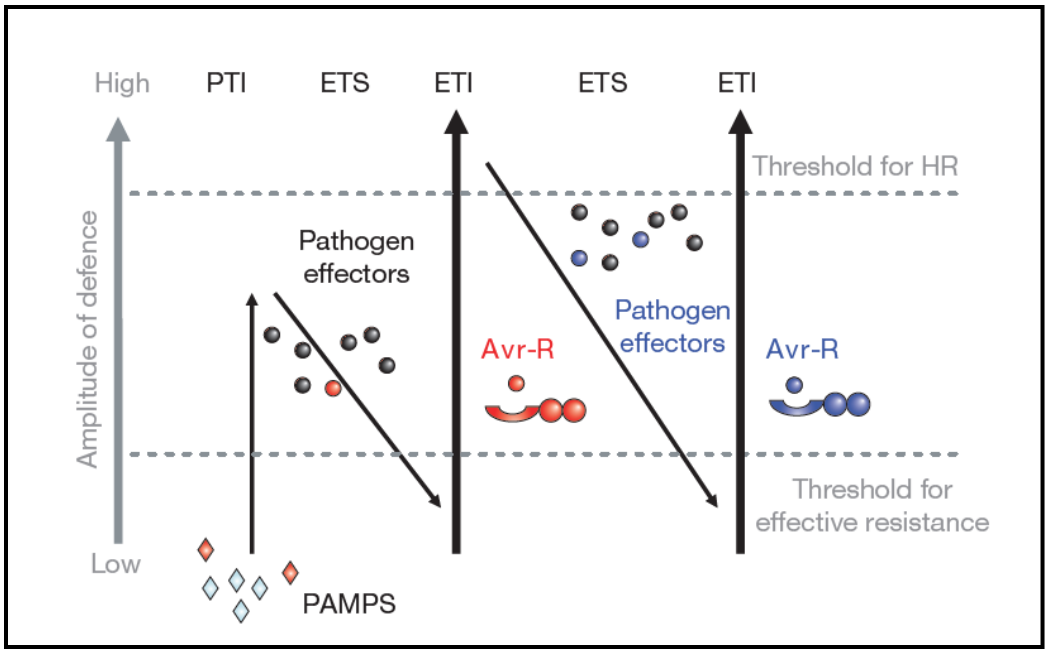


Figure 1. The zigzag model described by Jones & Dangl (2006). Initially, pathogen-associated molecular patterns (PAMPS) on the pathogen surface are recognized by pattern recognition receptors (PRRs) of the host, triggering pathogen triggered immunity (PTI) – the first phase of defense against the pathogen. Pathogens then secrete effectors to suppress PTI and result in effector-triggered susceptibility (ETS) in the host. Pathogen effectors can then be recognised by host resistance (R) proteins, known as nucleotide-binding site leucine-rich repeat proteins (NBS-LRRs), and effector-triggered immunity (ETI) is initiated to eliminate the pathogen and prevent disease. The pathogen can then evolve to produce novel effectors which are able to suppress ETI and again produce disease by ETS. In turn, this evolution of pathogen effectors drives the evolution of plant *R* genes to combat ETS and again establish ETI.

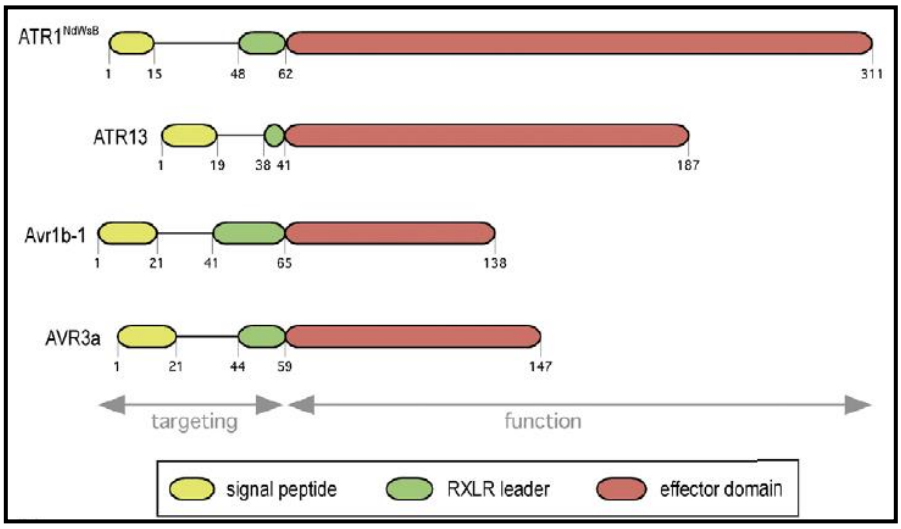


Figure 2. Domain organization of RxLR effectors, based on the structure of known RxLR effectors ATR1 and ATR13 from *Hyaloperonospora parasitica*, Avr1b-1 from *Phytophthora sojae* and AVR3a from *Phytophthora infestans*, as demonstrated by Morgan and Kamoun (2007). Here, numbers indicate the amino acid positions, and the RxLR leader includes both the RxLR conserved motif and the downstream dEER motif. The grey arrows indicate separation of the N-terminal domain for targeting to host cells from the C-terminal domain for effector activity.

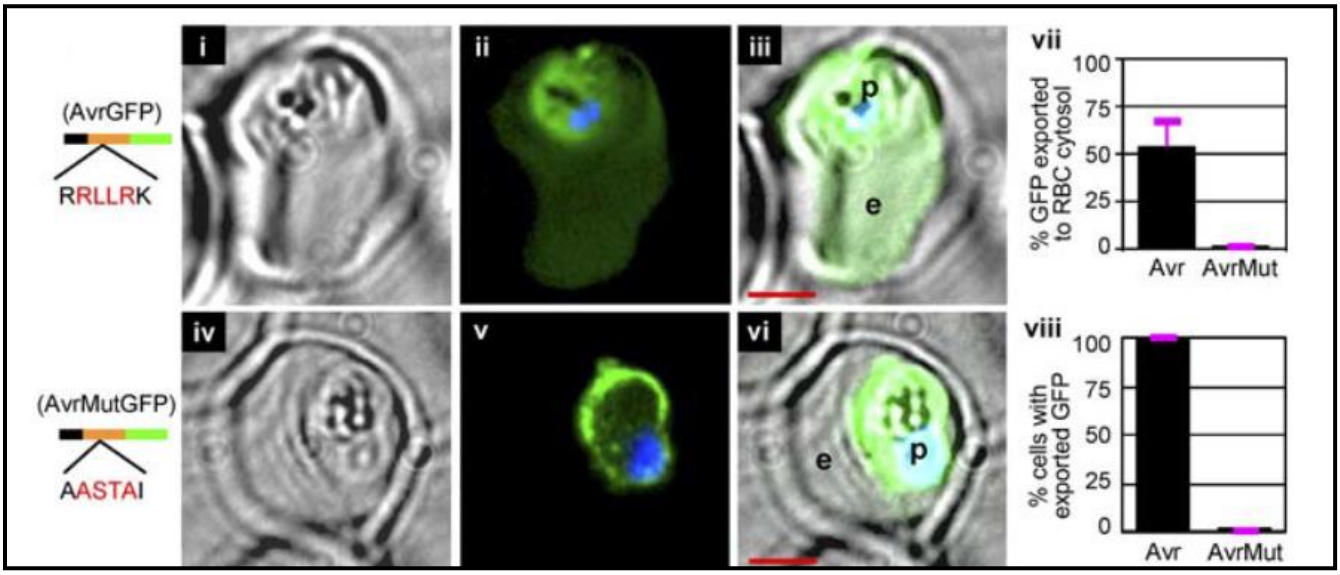


Figure 3. Effects of fusion of AVR3a leader sequence to green fluorescent protein (GFP), as demonstrated by Bhattacharjee *et al.* (2006). Images (i-iii) show that GFP containing the AVR3a leader sequence was successfully translocated from the *Plasmodium falciparum* vacuole to host erythrocytes. When the RxLR domain of the AVR3a leader sequence was replaced with an alternative, mutated amino acid sequence, GFP was no longer translocated from the parasite vacuole, as shown in images (iv-vi). Image (vii) shows that there was a significant reduction in the percentage of GFP exported from *P. falciparum* when the RxLR domain of AVR3a was mutated, and graph (viii) shows that all cells with the AVR3a RxLR domain contained exported GFP, while none of the cells containing the mutated leader sequence contained exported GFP.

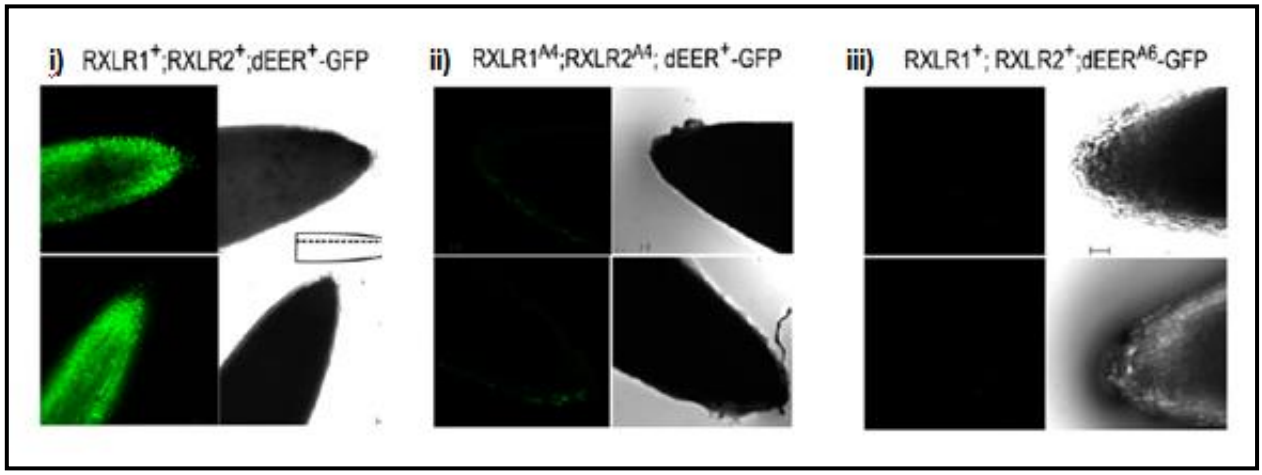


Figure 4. Translocation of GFP fusion constructs when soybean root tips were incubated with *Esherichia coli* cells expressing GFP fused to Avr1b leader sequences, as demonstrated by Dou *et al*. (2008a). i) GFP fused to the unaltered RxLR-dEER motif from Avr1b was successfully transported into soybean host cells. ii) When the RxLR motifs were replaced with alanine residues, transport of GFP into host cells was abolished, even when the dEER motif was unchanged. iii) When the dEER motif was replaced with alanine residues, GFP could not be successfully transported into host cells, even when the RxLR motifs were kept intact.

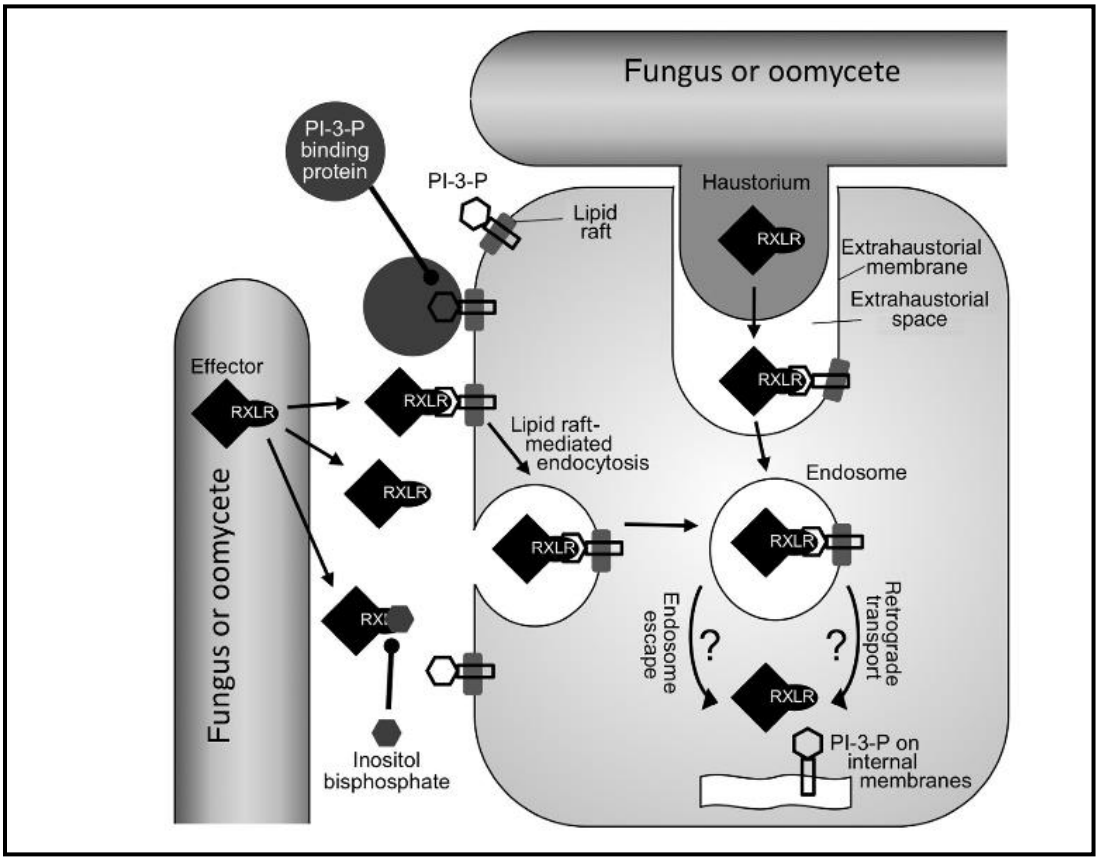


Figure 5. Tyler (2011) depicts the proposed method of entry for RxLR effectors into host cells. The image illustrates the entry of secreted RxLR effectors into a host cell via lipid raft-mediated endocytosis after binding of the RxLR amino acid motif to phosphatidylinositol-3-phosphate (PI3P). It is still unknown how these effectors subsequently escape from endosomes to function within the cell. The diagram also shows that the use of inositol bisphosphate or PI3P-binding proteins blocks effector entry into host cells.

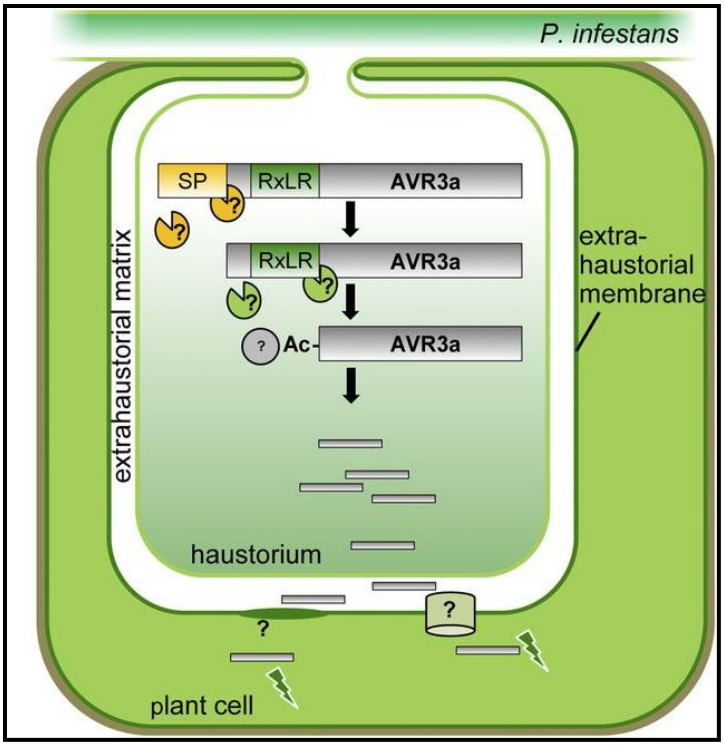


Figure 6. Wawra and colleagues (2017) suggest a new model for RxLR translocation and secretion. Researchers suggest that RxLR effectors are cleaved at the RxLR motif and then acetylated before they can be secreted from *Phytophthora* haustoria, as is the case for *Phytophthora infestans* effector AVR3a. The presence of the RxLR motif is thus proposed here to play a role in secretion of the effectors from pathogen haustoria, rather than in translocation of the effectors directly into host cells.

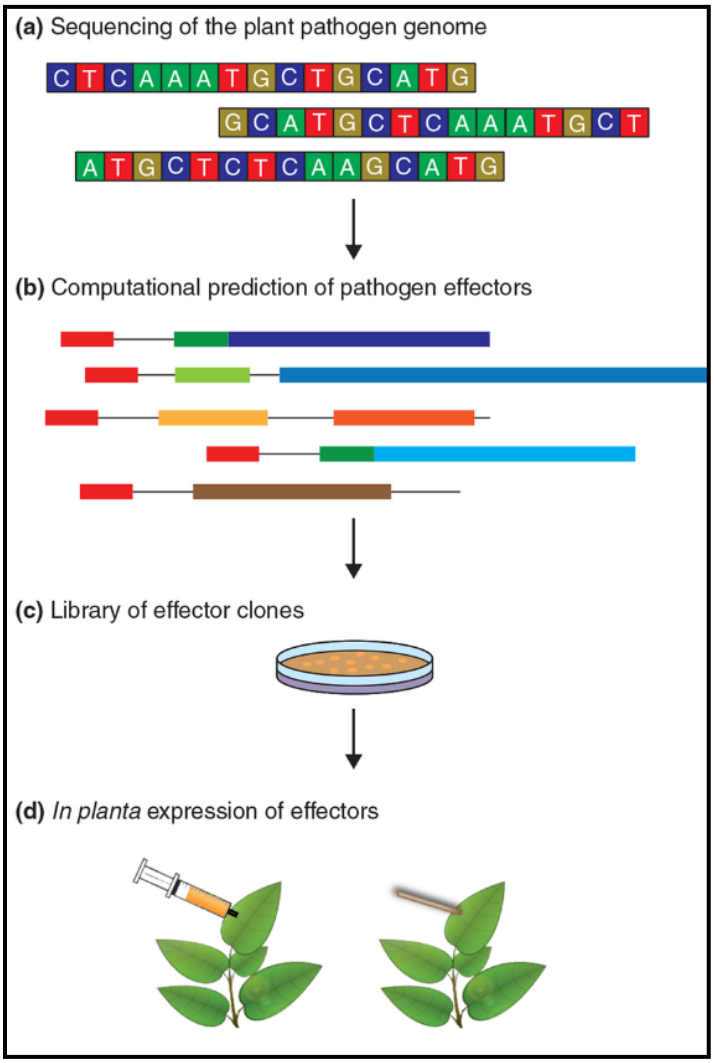


Figure 7. Overview of effector identification as described by Pais and colleagues (2013). Availability of pathogen genome sequences (a) allows the prediction of candidate effectors by computational methods (b). Sets of predicted effectors can then be cloned (c), allowing for their expression in host plants (d) and, ultimately, characterization of their functions during infection.

Figure 8. Agro-infiltration of *Nicotiana benthamiana* by Armstrong *et al.* (2005). Transient expression of the unchanged Avr3a effector from *Phytophthora infestans* alone did not result in cell death around infiltration sites. Similarly, expression of only the mutant avr3a effector, which lacked the N-terminal signal peptide, did not cause localized cell death. Co-expression of the mutant avr3a effector together with the R3a resistance protein from potato plants did not result in lesions on *N. benthamiana* leaves. When the unaltered Avr3a effector was co-expressed with the R3a resistance protein, however, lesions were formed on *N. benthamiana* leaves. These lesions were caused by localized cell death, characteristic of hypersensitive response in plant cells.

References

Agrios G (2005) *Plant Pathology*. Elsevier Academic Press, Burlington.

Ali SS, Shao J, Lary DJ, Kronmiller BA, Shen D, Strem MD, Amoako-Attah I, Akrofi AY, Begoude BD & ten Hoopen GM (2017) *Phytophthora megakarya* and *Phytophthora palmivora*, closely related causal agents of Cacao Black Pod Rot, underwent increases in genome sizes and gene numbers by different mechanisms. *Genome Biology and Evolution* **9**: 536-557.

Allardyce JA, Rookes JE & Cahill DM (2012) Defining plant resistance to *Phytophthora cinnamomi*: a standardized approach to assessment. *Journal of Phytopathology* **160**: 269-276.

Armstrong MR, Whisson SC, Pritchard L*, et al.* (2005) An ancestral oomycete locus contains late blight avirulence gene *Avr3a,* encoding a protein that is recognized in the host cytoplasm. *Proceedings of the National Academy of Sciences* **102**: 7766-7771.

Backer R, Naidoo S & van den Berg N (2019) The NON EXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) and Related Family: Mechanistic Insights in Plant Disease Resistance. *Frontiers in Plant Science* **10**: doi: 10.3389/fpls.2019.00102.

Bartsch M, Gobbato E, Bednarek P, Debey S, Schultze JL, Bautor J & Parker JE (2006) Salicylic acid–independent ENHANCED DISEASE SUSCEPTIBILITY1 signaling in *Arabidopsis* immunity and cell death is regulated by the monooxygenase *FMO1* and the nudix hydrolase *NUDT7*. *The Plant Cell* **18**: 1038-1051.

Bhadauria V, Banniza S, Wei Y & Peng Y-L (2009) Reverse genetics for functional genomics of phytopathogenic fungi and oomycetes. *Comparative and Functional Genomics* **2009**: doi:10.1155/2009/380719.

Bhattacharjee S, Hiller NL, Liolios K, Win J, Kanneganti T-D, Young C, Kamoun S & Haldar K (2006) The malarial host-targeting signal is conserved in the Irish potato famine pathogen. *PLoS Pathogens* **2**: e50.

Birch PR, Boevink PC, Gilroy EM, Hein I, Pritchard L & Whisson SC (2008) Oomycete RXLR effectors: delivery, functional redundancy and durable disease resistance. *Current Opinion in Plant Biology* **11**: 373-379.

Boevink PC, Wang X, McLellan H*, et al.* (2016) A *Phytophthora infestans* RXLR effector targets plant PP1c isoforms that promote late blight disease. *Nature Communications* **7**: 10311.

Boller T & Felix G (2009) A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology* **60**: 379-406.

Bos JI, Armstrong MR, Gilroy EM, Boevink PC, Hein I, Taylor RM, Zhendong T, Engelhardt S, Vetukuri RR & Harrower B (2010) *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proceedings of the National Academy of Sciences* **107**: 9909-9914.

Bouwmeester K, De Sain M, Weide R, Gouget A, Klamer S, Canut H & Govers F (2011) The lectin receptor kinase LecRK-I. 9 is a novel *Phytophthora* resistance component and a potential host target for a RXLR effector. *PLoS Pathogens* **7**: e1001327.

Bozkurt TO, Schornack S, Banfield MJ & Kamoun S (2012) Oomycetes, effectors, and all that jazz. *Current Opinion in Plant Biology* **15**: 483-492.

Bozkurt TO, Schornack S, Win J, Shindo T, Ilyas M, Oliva R, Cano LM, Jones AM, Huitema E & van der Hoorn RA (2011) *Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. *Proceedings of the National Academy of Sciences* **108**: 20832-20837.

Chaparro-Garcia A, Schwizer S, Sklenar J, Yoshida K, Petre B, Bos JI, Schornack S, Jones AM, Bozkurt TO & Kamoun S (2015) *Phytophthora infestans* RXLR-WY effector AVR3a associates with dynamin-related protein 2 required for endocytosis of the plant pattern recognition receptor FLS2. *PLoS One* **10**: DOI:10.1371/journal.pone.0137071.

Chen X-R, Zhang Y, Li H, Zhang Z-H, Sheng G-L, Li Y-P, Xing Y-P, Huang S-X, Tao H & Kuan T (2019) The RXLR effector PcAvh1 is required for full virulence of *Phytophthora capsici*. *Molecular Plant-Microbe Interactions* **32**: 986-1000.

Chen XR, Xing YP, Li YP, Tong YH & Xu JY (2013) RNA-Seq reveals infection-related gene expression changes in *Phytophthora capsici*. *PLoS One* **8**: e74588.

Chen Y, Liu Z & Halterman DA (2012) Molecular determinants of resistance activation and suppression by *Phytophthora infestans* effector IPI-O. *PLoS Pathogens* **8**: e1002595.

Cheng B, Yu X, Ma Z, Dong S, Dou D, Wang Y & Zheng X (2012) *Phytophthora sojae* effector Avh331 suppresses the plant defence response by disturbing the MAPK signalling pathway. *Physiological and Molecular Plant Pathology* **77**: 1-9.

Cook DE, Mesarich CH & Thomma BP (2015) Understanding plant immunity as a surveillance system to detect invasion. *Annual Reviews in Phytopathology* **53**: 541-563.

Dagdas YF, Belhaj K, Maqbool A*, et al.* (2016) An effector of the Irish potato famine pathogen antagonizes a host autophagy cargo receptor. *eLife* **5**.

Dalio R, Maximo H, Oliveira T, Dias R, Breton M, Felizatti H & Machado M (2018) *Phytophthora parasitica* effector PpRxLR2 suppresses *Nicotiana benthamiana* immunity. *Molecular Plant-Microbe Interactions* **31**: 481-493.

Darvill AG & Albersheim P (1984) Phytoalexins and their elicitors - a defense against microbial infection in plants. *Annual Review of Plant Physiology* **35**: 243-275.

De Bary A (1876) Researches into the nature of the potato fungus, *Phytophthora infestans*. *Journal of the Royal Agricultural Society of England* **2**: 239-269.

de Vries S, von Dahlen JK, Uhlmann C, Schnake A, Kloesges T & Rose LE (2017) Signatures of selection and host-adapted gene expression of the *Phytophthora infestans* RNA silencing suppressor PSR2. *Molecular Plant Pathology* **18**: 110-124.

de Wit PJ (2007) How plants recognize pathogens and defend themselves. *Cellular and Molecular Life Sciences* **64**: 2726-2732.

Di Paolo G & De Camilli P (2006) Phosphoinositides in cell regulation and membrane dynamics. *Nature* **443**: 651-657.

Dong S, Yu D, Cui L, Qutob D, Tedman-Jones J, Kale SD, Tyler BM, Wang Y & Gijzen M (2011a) Sequence variants of the *Phytophthora sojae* RXLR effector Avr3a/5 are differentially recognized by Rps3a and Rps5 in soybean. *PloS One* **6**: 1-8.

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

Dou D & Zhou JM (2012) Phytopathogen effectors subverting host immunity: different foes, similar battleground. *Cell Host & Microbe* **12**: 484-495.

Dou D, Kale SD, Wang X, Jiang RH, Bruce NA, Arredondo FD, Zhang X & Tyler BM (2008a) RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. *Plant Cell* **20**: 1930-1947.

Dou D, Kale SD, Wang X*, et al.* (2008b) Conserved C-terminal motifs required for avirulence and suppression of cell death by *Phytophthora sojae* effector Avr1b. *Plant Cell* **20**: 1118-1133.

Erwin DC & Ribeiro OK (1996) *Phytophthora diseases worldwide*. American Phytopathological Society (APS Press).

Evangelisti E, Govetto B, Minet-Kebdani N, Kuhn ML, Attard A, Ponchet M, Panabières F & Gourgues M (2013) The *Phytophthora parasitica* RXLR effector Penetration-Specific Effector 1 favours *Arabidopsis thaliana* infection by interfering with auxin physiology. *New Phytologist* **199**: 476-489.

Fabro G, Steinbrenner J, Coates M, Ishaque N, Baxter L, Studholme DJ, Körner E, Allen RL, Piquerez SJ & Rougon-Cardoso A (2011) Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host plant immunity. *PLoS Pathogens* **7**: e1002348.

Fan G, Yang Y, Li T, Lu W, Du Y, Qiang X, Wen Q & Shan W (2018) A *Phytophthora capsici* RXLR effector targets and inhibits a plant PPIase to suppress endoplasmic reticulum-mediated immunity. *Molecular Plant* **11**: 1067-1083.

Fawke S, Doumane M & Schornack S (2015) Oomycete interactions with plants: infection strategies and resistance principles. *Microbiology and Molecular Biology Reviews* **79**: 263-280.
Flor H (1955) Host-parasite interaction in flax rust - its genetics and other implications. *Phytopathology* **45**: 680-685.

Flor HH (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9**: 275-296.

Fry W (2008) *Phytophthora infestans*: the plant (and R gene) destroyer. *Molecular Plant Pathology* **9**: 385-402.

Gao R, Cheng Y, Wang Y, Guo L & Zhang G (2015) Genome sequence of *Phytophthora fragariae* var. *fragariae*, a quarantine plant-pathogenic fungus. *Genome Announcements* **3**: e00034-00015.

Ge X, Li G-J, Wang S-B, Zhu H, Zhu T, Wang X & Xia Y (2007) AtNUDT7, a negative regulator of basal immunity in Arabidopsis, modulates two distinct defense response pathways and is involved in maintaining redox homeostasis. *Plant Physiology* **145**: 204-215.

Gilroy EM, Taylor RM, Hein I, Boevink P, Sadanandom A & Birch PR (2011b) CMPG1-dependent cell death follows perception of diverse pathogen elicitors at the host plasma membrane and is suppressed by *Phytophthora infestans* RXLR effector AVR3a. *New Phytologist* **190**: 653-666.

Gilroy EM, Breen S, Whisson SC*, et al.* (2011a) Presence/absence, differential expression and sequence polymorphisms between *PiAVR2* and *PiAVR2-like* in *Phytophthora infestans* determine virulence on *R2* plants. *New Phytologist* **191**: 763-776.

Glazebrook J (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology* **43**: 205-227.

González-Lamothe R, Tsitsigiannis DI, Ludwig AA, Panicot M, Shirasu K & Jones JD (2006) The U-box protein CMPG1 is required for efficient activation of defense mechanisms triggered by multiple resistance genes in tobacco and tomato. *The Plant Cell* **18**: 1067-1083.

Goss EM, Press CM & Grünwald NJ (2013) Evolution of RXLR-class effectors in the oomycete plant pathogen *Phytophthora ramorum*. *PLoS One* **8**: e79347.

Grünwald NJ (2012b) Genome sequences of *Phytophthora* enable translational plant disease management and accelerate research. *Canadian Journal of Plant Pathology* **34**: 13-19.

Grünwald NJ, Garbelotto M, Goss EM, Heungens K & Prospero S (2012a) Emergence of the sudden oak death pathogen *Phytophthora ramorum*. *Trends in Microbiology* **20**: 131-138.

Guo B, Wang H, Yang B, Jiang W, Jing M, Li H, Xia Y, Xu Y, Hu Q & Wang F (2019) *Phytophthora sojae* effector PsAvh240 inhibits host aspartic protease secretion to promote infection. *Molecular Plant* **12**: 552-564.

Haas BJ, Kamoun S, Zody MC*, et al.* (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* **461**: 393-398.

Hardham AR & Blackman LM (2010) Molecular cytology of *Phytophthora*–plant interactions. *Australasian Plant Pathology* **39**: 29-35.

Hardham AR & Blackman LM (2017) *Phytophthora cinnamomi*. *Molecular Plant Pathology* **19**: 260-285.

Hein I, Gilroy EM, Armstrong MR & Birch PR (2009) The zig-zag-zig in oomycete-plant interactions. *Molecular Plant Pathology* **10**: 547-562.

Hiller NL, 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.

Huang J, Gu L, Zhang Y, Yan T, Kong G, Kong L, Guo B, Qiu M, Wang Y & Jing M (2017) An oomycete plant pathogen reprograms host pre-mRNA splicing to subvert immunity. *Nature Communications* **8**: 2051.

Jiang RH & Tyler BM (2012) Mechanisms and evolution of virulence in oomycetes. *Annual Reviews in Phytopathology* **50**: 295-318.

Jiang RH, Tripathy S, Govers F & Tyler BM (2008) RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the Natural Academy of Sciences* **105**: 4874-4879.

Jing M, Guo B, Li H*, et al.* (2016) A *Phytophthora sojae* effector suppresses endoplasmic reticulum stress-mediated immunity by stabilizing plant Binding immunoglobulin Proteins. *Nature Communications* **7**: 11685.

Johnston SF, Cohen MF, Torok T, Meentemeyer RK & Rank NE (2015) Host phenology and leaf effects on susceptibility of California bay laurel to *Phytophthora ramorum*. *Phytopathology* **106**: 47-55.

Jones JDG & Dangl JL (2006) The plant immune system. *Nature* **444**: 323-329.

Kale SD & Tyler BM (2011) Entry of oomycete and fungal effectors into plant and animal host cells. *Cellular Microbiology* **13**: 1839-1848.

Kale SD, Gu B, Capelluto DG, Dou D, Feldman E, Rumore A, Arredondo FD, Hanlon R, Fudal I & Rouxel T (2010) External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell* **142**: 284-295.

Kamoun S (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Phytopathology* **44**: 41.

Kamoun S (2007) Groovy times: filamentous pathogen effectors revealed. *Current Opinion in Plant Biology* **10**: 358-365.

Kamoun S, Furzer O, Jones JD, Judelson HS, Ali GS, Dalio RJ, Roy SG, Schena L, Zambounis A & Panabières F (2015) The Top 10 oomycete pathogens in molecular plant pathology. *Molecular Plant Pathology* **16**: 413-434.

Kaschani F, Shabab M, Bozkurt T, Shindo T, Schornack S, Gu C, Ilyas M, Win J, Kamoun S & van der Hoorn RA (2010) An effector-targeted protease contributes to defense against *Phytophthora infestans* and is under diversifying selection in natural hosts. *Plant Physiology* **154**: 1794-1804.

Kelley BS, Lee SJ, Damasceno CM, Chakravarthy S, Kim BD, Martin GB & Rose JK (2010) A secreted effector protein (SNE1) from *Phytophthora infestans* is a broadly acting suppressor of programmed cell death. *The Plant Journal* **62**: 357-366.

King SR, McLellan H, Boevink PC, Armstrong MR, Bukharova T, Sukarta O, Win J, Kamoun S, Birch PR & Banfield MJ (2014) *Phytophthora infestans* RXLR effector PexRD2 interacts with host MAPKKKε to suppress plant immune signaling. *The Plant Cell* **26**: 1345-1359.

Kong L, Qiu X, Kang J*, et al.* (2017) A *Phytophthora* Effector Manipulates Host Histone Acetylation and Reprograms Defense Gene Expression to Promote Infection. *Current Biology* **27**: 981-991.

Kroon LP, Brouwer H, de Cock AW & Govers F (2012) The genus *Phytophthora* anno 2012. *Phytopathology* **102**: 348-364.

Lamour KH, Mudge J, Gobena D, Hurtado-Gonzales OP, Schmutz J, Kuo A, Miller NA, Rice BJ, Raffaele S & Cano LM (2012) Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. *Molecular Plant-Microbe Interactions* **25**: 1350-1360.

Latijnhouwers M, de Wit PJ & Govers F (2003) Oomycetes and fungi: similar weaponry to attack plants. *Trends in Microbiology* **11**: 462-469.

Lee S-J, Kelley BS, Damasceno CM, John BS, Kim B-S, Kim B-D & Rose JK (2006) A functional screen to characterize the secretomes of eukaryotic pathogens and their hosts in planta. *Molecular Plant-Microbe Interactions* **19**: 1368-1377.

Li Q, Chen Y, Wang J, Zou F, Jia Y, Shen D, Zhang Q, Jing M, Dou D & Zhang M (2019a) A *Phytophthora capsici* virulence effector associates with NPR1 and suppresses plant immune responses. *Phytopathology Research* **1**.

Li Q, Ai G, Shen D, Zou F, Wang J, Bai T, Chen Y, Li S, Zhang M & Jing M (2019b) A *Phytophthora capsici* effector targets ACD11 binding partners that regulate ROS-mediated defense response in *Arabidopsis*. *Molecular Plant* **12**: 565-581.

Lokossou AA, Park T-h, van Arkel G, Arens M, Ruyter-Spira C, Morales J, Whisson SC, Birch PR, Visser RG & Jacobsen E (2009) Exploiting knowledge of R/Avr genes to rapidly clone a new LZ-NBS-LRR family of late blight resistance genes from potato linkage group IV. *Molecular Plant-Microbe Interactions* **22**: 630-641.

Lokossou AA, Rietman H, Wang M, Krenek P, van der Schoot H, Henken B, Hoekstra R, Vleeshouwers VG, van der Vossen EA & Visser RG (2010) Diversity, distribution, and evolution of *Solanum bulbocastanum* late blight resistance genes. *Molecular Plant-Microbe Interactions* **23**: 1206-1216.

Lotze MT, Zeh HJ, Rubartelli A, Sparvero LJ, Amoscato AA, Washburn NR, DeVera ME, Liang X, Tör M & Billiar T (2007) The grateful dead: damage-associated molecular pattern molecules and reduction/oxidation regulate immunity. *Immunological Reviews* **220**: 60-81.

MacGregor T, Bhattacharyya M, Tyler B, Bhat R, Schmitthenner AF & Gijzen M (2002) Genetic and physical mapping of *Avr1a* in *Phytophthora sojae*. *Genetics* **160**: 949-959.

Marti M, Good RT, Rug M, Knuepfer E & Cowman AF (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**: 1930-1933.

McLellan H, Boevink PC, Armstrong MR, Pritchard L, Gomez S, Morales J, Whisson SC, Beynon JL & Birch PR (2013) An RxLR effector from *Phytophthora infestans* prevents re-localisation of two plant NAC transcription factors from the endoplasmic reticulum to the nucleus. *PLoS Pathogens* **9**: e1003670.

Mildvan A, Xia Z, Azurmendi H, Saraswat V, Legler P, Massiah M, Gabelli S, Bianchet M, Kang L-W & Amzel L (2005) Structures and mechanisms of Nudix hydrolases. *Archives of Biochemistry and Biophysics* **433**: 129-143.

Morgan W & Kamoun S (2007) RXLR effectors of plant pathogenic oomycetes. *Current Opinion in Microbiology* **10**: 332-338.

Murphy F, He Q, Armstrong M, Giuliani LM, Boevink PC, Zhang W, Tian Z, Birch PR & Gilroy EM (2018) The potato MAP3K StVIK is required for the *Phytophthora infestans* RXLR effector Pi17316 to promote disease. *Plant Physiology* **177**: 398-410.

Odorizzi G, Babst M & Emr SD (2000) Phosphoinositide signaling and the regulation of membrane trafficking in yeast. *Trends in Biochemical Sciences* **25**: 229-235.

Oh SK, Young C, Lee M*, et al.* (2009) In planta expression screens of *Phytophthora infestans* RXLR effectors reveal diverse phenotypes, including activation of the *Solanum bulbocastanum* disease resistance protein Rpi-blb2. *Plant Cell* **21**: 2928-2947.

Pais M, Win J, Yoshida K, Etherington GJ, Cano LM, Raffaele S, Banfield MJ, Jones A, Kamoun S & Saunders DG (2013) From pathogen genomes to host plant processes: the power of plant parasitic oomycetes. *Genome Biology* **14**: 211.

Pike LJ (2003) Lipid rafts bringing order to chaos. *Journal of Lipid Research* **44**: 655-667.

Qiao Y, Shi J, Zhai Y, Hou Y & Ma W (2015) *Phytophthora* effector targets a novel component of small RNA pathway in plants to promote infection. *Proceedings of the National Academy of Sciences* **112**: 5850-5855.

Qiao Y, Liu L, Xiong Q, Flores C, Wong J, Shi J, Wang X, Liu X, Xiang Q & Jiang S (2013) Oomycete pathogens encode RNA silencing suppressors. *Nature Genetics* **45**: 330-333.

Quinn L, O'Neill PA, Harrison J, Paskiewicz KH, McCracken AR, Cooke LR, Grant MR & Studholme DJ (2013) Genome-wide sequencing of *Phytophthora lateralis* reveals genetic variation among isolates from Lawson cypress (*Chamaecyparis lawsoniana*) in Northern Ireland. *FEMS Microbiology Letters* **344**: 179-185.

Qutob D, Tedman-Jones J, Dong S, Kuflu K, Pham H, Wang Y, Dou D, Kale SD, Arredondo FD & Tyler BM (2009) Copy number variation and transcriptional polymorphisms of *Phytophthora sojae* RXLR effector genes *Avr1a* and *Avr3a*. *PloS One* **4**: e5066.

Raven P, Johnson G, Mason K, Losos J & Singer S (2014) *Biology*. McGraw-Hill, New York.

Ren Y, Armstrong M, Qi Y, McLellan H, Zhong C, Du B, Birch PR & Tian Z (2019) *Phytophthora infestans* RXLR effectors target parallel steps in an immune signal transduction pathway. *Plant Physiology* **180**: 2227-2239.

Rouxel T & Balesdent M-H (2010) Avirulence Genes. *Encyclopaedia of Life Sciences,* doi: 10.1002/9780470015902.a0021267. John Wiley & Sons, Ltd, Chichesters.

Sanju S, Siddappa S, Thakur A, Shukla PK, Srivastava N, Pattanayak D, Sharma S & Singh B (2015) Host-mediated gene silencing of a single effector gene from the potato pathogen *Phytophthora infestans* imparts partial resistance to late blight disease. *Functional & Integrative Genomics* **15**: 697-706.

Schornack S, Huitema E, Cano LM*, et al.* (2009) Ten things to know about oomycete effectors. *Molecular Plant Pathology* **10**: 795-803.

Segretin ME, Pais M, Franceschetti M, Chaparro-Garcia A, Bos JI, Banfield MJ & Kamoun S (2014) Single amino acid mutations in the potato immune receptor R3a expand response to *Phytophthora* effectors. *Molecular Plant-Microbe Interactions* **27**: 624-637.

Selin C, de Kievit TR, Belmonte MF & Fernando WGD (2016) Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. *Frontiers in Microbiology* **7**: doi: 10.3389/fmicb.2016.00600.

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**: 502-509.

Shabab M, Shindo T, Gu C, Kaschani F, Pansuriya T, Chintha R, Harzen A, Colby T, Kamoun S & van der Hoorn RA (2008) Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato. *The Plant Cell* **20**: 1169-1183.

Shewan A, Eastburn DJ & Mostov K (2011) Phosphoinositides in cell architecture. *Cold Spring Harbor Perspectives in Biology* **3**: a004796.

Shindo T & Van Der Hoorn RA (2008) Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. *Molecular Plant Pathology* **9**: 119-125.

Song J, Win J, Tian M, Schornack S, Kaschani F, Ilyas M, van der Hoorn RA & Kamoun S (2009) Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3. *Proceedings of the National Academy of Sciences* **106**: 1654-1659.

Song T, Kale SD, Arredondo FD, Shen D, Su L, Liu L, Wu Y, Wang Y, Dou D & Tyler BM (2013) Two RxLR avirulence genes in *Phytophthora sojae* determine soybean *Rps 1k*-mediated disease resistance. *Molecular Plant-Microbe Interactions* **26**: 711-720.

Spielman LJ, Sweigard JA, Shattock RC & Fry WE (1990) The genetics of *Phytophthora infestans*: Segregation of allozyme markers in F2 and backcross progeny and the inheritance of virulence against potato resistance genes *R2* and *R4* in F1 progeny. *Experimental Mycology* **14**: 57-69.

Spoel SH, Mou Z, Tada Y, Spivey NW, Genschik P & Dong X (2009) Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. *Cell* **137**: 860-872.

Stassen JH & Van den Ackerveken G (2011) How do oomycete effectors interfere with plant life? *Current Opinion in Plant Biology* **14**: 407-414.

Stassen JH, Boer Ed, Vergeer PW, Andel A, Ellendorff U, Pelgrom K, Pel M, Schut J, Zonneveld O & Jeuken MJ (2013) Specific *in planta* recognition of two GKLR proteins of the downy mildew *Bremia lactucae* revealed in a large effector screen in lettuce. *Molecular Plant-Microbe Interactions* **26**: 1259-1270.

Sun F, Kale SD, Azurmendi HF, Li D, Tyler BM & Capelluto DG (2013) Structural basis for interactions of the *Phytophthora sojae* RxLR effector Avh5 with phosphatidylinositol 3-phosphate and for host cell entry. *Molecular Plant-Microbe Interactions* **26**: 330-344.

Tabima J, Kronmiller B, Press C, Tyler B, Zasada I & Grunwald N (2017) Whole genome sequences of the raspberry and strawberry pathogens *Phytophthora rubi* and *P. fragariae*. *Molecular Plant-Microbe Interactions* 115824.

Thines M & Kamoun S (2010) Oomycete–plant coevolution: recent advances and future prospects. *Current Opinion in Plant Biology* **13**: 427-433.

Tian M, Win J, Savory E, Burkhardt A, Held M, Brandizzi F & Day B (2011) 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Molecular Plant-Microbe Interactions* **24**: 543-553.

Torres MA, Jones JD & Dangl JL (2006) Reactive oxygen species signaling in response to pathogens. *Plant Physiology* **141**: 373-378.

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

Tyler B (2011) Entry of oomycete and fungal effectors into host cells. *Effectors in Plant-Microbe Interactions,*(Martin F & Kamoun S, eds.), p.^pp. 243-278. Wiley-Blackwell, Oxford.

Tyler BM (2007) *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. *Molecular Plant Pathology* **8**: 1-8.

Tyler BM (2009) Entering and breaking: virulence effector proteins of oomycete plant pathogens. *Cellular Microbiology* **11**: 13-20.

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D & Beynon JL (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* **313**: 1261-1266.

Tyler BM, Kale SD, Wang Q, Tao K, Clark HR, Drews K, Antignani V, Rumore A, Hayes T & Plett JM (2013) Microbe-independent entry of oomycete RxLR effectors and fungal RxLR-like effectors into plant and animal cells is specific and reproducible. *Molecular Plant-Microbe Interactions* **26**: 611-616.

van't Slot KA & 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.

van Damme M, Cano LM, Oliva R, Schornack S, Segretin ME, Kamoun S & Raffaele S (2011) Evolutionary and functional dynamics of oomycete effector genes. *Effectors in Plant–Microbe Interactions,*(Martin F & Kamoun S, eds.), p.^pp. 101-102. Wiley-Blackwell, Oxford, UK.

Van der Hoorn RA, De Wit PJ & Joosten MH (2002) Balancing selection favors guarding resistance proteins. *Trends in Plant Science* **7**: 67-71.

van der Lee T, Robold A, Testa A, van't Klooster JW & Govers F (2001) Mapping of avirulence genes in *Phytophthora infestans* with amplified fragment length polymorphism markers selected by bulked segregant analysis. *Genetics* **157**: 949-956.

van Poppel PM, Guo J, van de Vondervoort PJ, Jung MW, Birch PR, Whisson SC & Govers F (2008) The *Phytophthora infestan*s avirulence gene *Avr4* encodes an RXLR-dEER effector. *Molecular Plant-Microbe Interactions* **21**: 1460-1470.

Vetukuri RR, Whisson SC & Grenville-Briggs LJ (2017) *Phytophthora infestans* effector Pi14054 is a novel candidate suppressor of host silencing mechanisms. *European Journal of Plant Pathology* **149**: 771-777.

Vetukuri RR, Kushwaha S, Sen D, Whisson SC, Lamour K & Grenville-Briggs L (2018) Genome sequence resource for the oomycete taro pathogen *Phytophthora colocasiae*. *Molecular Plant-Microbe Interactions*.

Vleeshouwers VG, Rietman H, Krenek P, Champouret N, Young C, Oh S-K, Wang M, Bouwmeester K, Vosman B & Visser RG (2008) Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes. *PLoS one* **3**: e2875.

Vleeshouwers VG, Raffaele S, Vossen JH, Champouret N, Oliva R, Segretin ME, Rietman H, Cano LM, Lokossou A & Kessel G (2011) Understanding and exploiting late blight resistance in the age of effectors. *Annual Review of Phytopathology* **49**: 507-531.

Wang Q, Han C, Ferreira AO*, et al.* (2011) Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. *Plant Cell* **23**: 2064-2086.

Wawra S, Trusch F, Matena A, Apostolakis K, Linne U, Zhukov I, Stanek J, Koźmiński W, Davidson I & Secombes CJ (2017) The RxLR motif of the host targeting effector AVR3a of *Phytophthora infestans* is cleaved before secretion. *The Plant Cell* **29**: 1184-1195.

Wawra S, Agacan M, Boddey JA*, et al.* (2012) Avirulence protein 3a (AVR3a) from the potato pathogen *Phytophthora infestans* forms homodimers through its predicted translocation region and does not specifically bind phospholipids. *Journal of Biological Chemistry* **287**: 38101-38109.

Whisson S, Drenth A, Maclean D & Irwin J (1994) Evidence for outcrossing in *Phytophthora sojae* and linkage of a DNA marker to two avirulence genes. *Current Genetics* **27**: 77-82.

Whisson S, Lee T, Bryan G, Waugh R, Govers F & Birch P (2001) Physical mapping across an avirulence locus of *Phytophthora infestans* using a highly representative, large-insert bacterial artificial chromosome library. *Molecular Genetics and Genomics* **266**: 289-295.

Whisson SC, Boevink PC, Moleleki L*, et al.* (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* **450**: 115-118.

Win J, Morgan W, Bos J, Krasileva KV, Cano LM, Chaparro-Garcia A, Ammar R, Staskawicz BJ & Kamoun S (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* **19**: 2349-2369.

Wirthmueller L, Maqbool A & Banfield MJ (2013) On the front line: structural insights into plant–pathogen interactions. *Nature Reviews Microbiology* **11**: 761-776.

Xiong Q, Ye W, Choi D, Wong J, Qiao Y, Tao K, Wang Y & Ma W (2014) *Phytophthora* suppressor of RNA silencing 2 is a conserved RxLR effector that promotes infection in soybean and *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions* **27**: 1379-1389.

Yaeno T, Li H, Chaparro-Garcia A, Schornack S, Koshiba S, Watanabe S, Kigawa T, Kamoun S & Shirasu K (2011) Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector AVR3a is required for its stability in host cells to modulate plant immunity. *Proceedings of the National Academy of Sciences* **108**: 14682-14687.

Yang B, Wang Q, Jing M*, et al.* (2017b) Distinct regions of the *Phytophthora* essential effector Avh238 determine its function in cell death activation and plant immunity suppression. *New Phytologist* **214**: 361-375.

Yang B, Wang Y, Guo B*, et al.* (2019) The *Phytophthora sojae* RXLR effector Avh238 destabilizes soybean Type2 GmACSs to suppress ethylene biosynthesis and promote infection. *New Phytologist* **222**: 425-437.

Yang X, Tyler BM & Hong C (2017a) An expanded phylogeny for the genus *Phytophthora*. *IMA fungus* **8**: 355-384.

Zentmyer GA (1976) *Phytophthora*—plant destroyer. *BioScience* **26**: 686-689.

Chapter 2

Identification and manual annotation of RxLR effectors in *Phytophthora cinnamomi*

Abstract

Phytophthora cinnamomi is a plant pathogenic oomycete that causes Phytophthora Root Rot of avocado (PRR). Currently, there is a limited understanding of the molecular interactions underlying this disease. Other *Phytophthora* species employ an arsenal of effector proteins to manipulate host physiology, of which the RxLR effectors contribute to virulence by interfering with host immune responses. The aim of this study was to identify candidate RxLRs in *P. cinnamomi* that play a role in PRR development, and to infer possible functions for these effectors. We identified 16 candidate *RxLR*s which were expressed during infection of a susceptible avocado rootstock, and these candidates contained all the characteristics of RxLR effector proteins. Several of these genes were present in multiple copies in the *P. cinnamomi* genome, suggesting that they may contribute to pathogen fitness. Transcriptome data were used to manually annotate and predict protein sequences for these effectors. Phylogenetic analysis was combined with evaluation of the expression profiles of the genes over the time-course of infection to assign putative functions to the candidate genes. This study represents the first investigation of the expression of *P. cinnamomi* RxLR effectors during avocado infection, and provides a foundation for the future functional characterization of RxLRs that contribute to *P. cinnamomi* virulence in avocado.

Introduction

Phytophthora cinnamomi Rands is a soil-borne, hemibiotrophic, plant-pathogenic oomycete with a broad range of host plants. It is globally distributed and affects at least 5000 plant species, including economically important crop plants such as avocado, macadamia, peach and chestnut. *P. cinnamomi* typically infects the roots of host plants causing root rot, as well as stem cankers and dieback of shoots (Hardham, 2005, Hardham & Blackman, 2017). The pathogen causes Phytophthora root rot (PRR) of avocado, which has had a devastating impact on this crop (Zentmyer, 1984, Coffey, 1987, Reeksting *et al.*, 2016, Belisle *et al.*, 2019). Control of PRR is primarily through the application of chemicals such as phosphites, the use of good agricultural practices, and the planting of tolerant or resistant rootstocks. Research into mechanisms of PRR resistance against the pathogen is essential, since eradication of *P. cinnamomi* is unlikely once it has established in soil (Hardy *et al.*, 2001).

Once a plant has been infected with *P. cinnamomi*, disease symptoms appear as a result of the release of effector proteins by the pathogen. These effector molecules are used by *Phytophthora* spp. to manipulate host plants (Stassen & Van den Ackerveken, 2011). According to the broadly accepted models of plant-pathogen interactions, effector proteins of most plant pathogens are important in eliciting specific responses in the host plant, by acting either as virulence or avirulence effectors (Cook *et al.*, 2015). Oomycetes use two different classes of effectors to contribute to pathogen virulence. Apoplastic effectors are secreted to the cell apoplast and function outside of host cells, whereas cytoplasmic effectors are secreted directly into the host cells in which they function (Kamoun, 2006).

One group of the cytoplasmic class of effectors in *Phytophthora* spp. have a common RxLR amino acid motif in their N-terminal regions, which has been hypothesized to play a role in their translocation and localization inside host cells (Whisson *et al.*, 2007, Dou *et al.*, 2008a, Kale *et al.*, 2010) The conserved RxLR motif has enabled the identification of hundreds of putative RxLR effectors in various *Phytophthora* genomes (Win *et al.*, 2007, Jiang *et al.*, 2008, Haas *et al.*, 2009, Ali *et al.*, 2017), though the functional characterization of the identified effectors remains an ongoing process.

In *P. cinnamomi*, several candidate *RxLR*s have been predicted, none of which have been functionally characterized. Reitmann *et al.* (2017) described the expression of 44 putative *RxLR* genes in cysts and germinating cysts of *P. cinnamomi*. Hardham & Blackman (2017) predicted 171 candidate RxLRs in the *P. cinnamomi* genome, based on similarity to candidate RxLRs in *Phytophthora infestans*. Most recently, McGowan & Fitzpatrick (2017) identified a total of 68 candidate *P. cinnamomi* RxLR effectors, using a combination of several prediction methods. To date, the only available expression data for *P. cinnamomi* RxLR genes is the transcriptome produced for pre-infection structures of the pathogen (Reitmann *et al.*, 2017).

The aim of this study was to identify candidate RxLR effectors in *P. cinnamomi* that play a role during infection of avocado, by focusing on the discovery of putative *RxLR*s which were upregulated during infection of a susceptible avocado rootstock. Potential RxLR genes were mined from the *P. cinnamomi* genome using a bioinformatics pipeline, and their expression profiles were investigated using available RNA-sequencing data. The candidate genes were manually annotated using transcriptome data, and their resultant protein sequences were predicted. Finally, these protein sequences were investigated for evolutionary relatedness to other *Phytophthora* effectors to assign putative functions to candidate *P. cinnamomi* RxLRs.

Methods and Materials

Identification of candidate *RxLR*s in the *P. cinnamomi* genome

Data mining of the P. cinnamomi *genome*

Although a draft genome has been produced for *P. cinnamomi* by the Joint Genome Institute (JGI) (USA Department of Energy, California), there is been no conclusive information on how many pathogenic effector genes occur in it. To that end, a bioinformatics pipeline adapted from Win *et al*. (2007) was used in this study to identify candidate *RxLR* genes in *P. cinnamomi*. Peptides predicted by the JGI annotated genes were submitted to SignalP Version 3.0 (Department of Bio and Health Informatics,

Technical University of Denmark) to determine which of the predicted peptides contained a signal peptide in their N-terminal region. Default parameters were used, the sequences were submitted as part of the eukaryotic organism group, and both Hidden Markov model (HMM) and Neural networks (NN) methods were used in the predictions. SignalP 3.0 predictions were filtered by only selecting peptides with a HMM score above 0.9 and a NN cleavage site between residues 10 and 30. Peptides predicted to contain a signal peptide were then further searched for the presence of an RxLR (arginine-any amino acid-leucine-arginine) motif between residues 30 and 60 of the protein sequence using perl regular expression. All proteins that passed the above steps were considered as candidate *P. cinnamomi* RxLRs.

Phylogenetic analysis of predicted RxLR *effector genes*

Protein sequences of the candidate *RxLR* effector genes were subjected to a multiple sequence alignment and subsequent phylogenetic analysis to determine whether there was any notable similarity between candidates within the set of putative genes. Only the N-terminal of the peptide sequences was used for the alignment, similar to the method used by Goss *et al.* (2013). Sequences were manually edited to include only the residues up to and including the dEER motif, and when there was no dEER motif the first 80 residues were used as the N-terminal sequence for alignment.

Alignment was performed using the MUSCLE alignment method (Edgar, 2004) in Geneious Prime 2020.0.3 with a maximum of 10 iterations and 5 trees. Parameters were automatically selected for the algorithm using the “optimise profile-dependant parameters” option. Aligned sequences were then manually edited to remove uninformative and ambiguously aligned regions before they were exported for phylogenetic analysis.

The edited alignment of the N-terminal peptide sequences of the candidate RxLRs was subjected to Bayesian inference analysis using MrBayes 3.2.7a (Ronquist *et al.*, 2012). The Poisson substitution model was used, and 10 parallel runs were conducted with four random chains each. The analysis used five million generations of the Markov chain

Monte Carlo (MCMC) analysis, with trees being sampled at every 100th generation. Subsequent to the MCMC analysis, 25% of the trees were discarded as a burn-in phase, and the remaining trees were used to calculate posterior probabilities.

A second phylogenetic analysis was performed to compare the set of candidate effectors to characterized RxLRs in other species. RxLRs with well-defined functions in other oomycetes were found based on a literature search for RxLR effectors and their protein sequences were obtained from Genbank (National Center for Biotechnology Information, US National Institutes of Health, USA) (Table 1). Full protein sequences that were not from Genbank sequences were obtained from either the UniprotKB database (Uniprot Consortium, 2014) or supplementary information from relevant research articles. Sequences were aligned and edited as before, and a phylogenetic tree was generated by Bayesian inference analysis in MrBayes 3.2.7a using the previously described method.

Identifying candidate *P. cinnamomi RxLRs* likely to play a role in avocado infection

Dual RNA-sequencing of avocado rootstocks infected by P. cinnamomi *GKB4*

The produced catalogue of candidate *RxLR* effector genes was used to determine which effectors were most likely to play a role in the infection of a susceptible avocado host. To that end, data from an RNA-sequencing experiment previously conducted by the Avocado Research Programme (ARP) was used to obtain expression information for the candidate effector genes.

Expression data were obtained by dual RNA-sequencing of avocado infected with *P. cinnamomi*. Roots of a susceptible avocado rootstock, R0.12, were infected with 1.4 x 10^5 zoospores/ml of *P. cinnamomi* isolate GKB4, and were harvested after 12 hours, 24 hours and 5 day. A partially resistant rootstock, R0.09, was inoculated as above, but roots were only harvested 5 days post inoculation. A culture of *P. cinnamomi* GKB4 was grown separately for 2 weeks on 20% V8 medium (200 ml clarified V8 juice, 2 g $CaCO_3$, 15 g agar and distilled water to a volume of 1L) before mycelium was harvested to determine

which pathogen genes were constitutively expressed and which were upregulated during infection.

Harvested samples were flash frozen in liquid N_2 and stored at -70°C until they were ground to a fine powder using an IKA® MT 40 tube mill. RNA was extracted from powdered samples using a modified CTAB extraction method (Chang *et al.*, 1993). Extracted RNA was purified using a Qiagen RNeasy clean up kit (Qiagen, Valencia, California, USA) subsequent to treatment with DNase I (Fermentas Life Sciences, Hanover, USA). The quality and purity of extracted RNA was then measured using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and prepared samples were stored at -70°C before 2 µg of each sample was sent to Novogene (Novogene Corporation Inc., Chula Vista, California, USA) for paired-end sequencing using Illumina Hiseq PE150.

Identification of expressed candidate RxLR*s from RNA-seq data*

RNA-seq reads were subjected to quality control using FASTQC and Trimmomatic (Bolger *et al.*, 2014) was used to trim random hexamers. Transcript abundance was quantified by performing a pseudo-alignment using Kallisto (Bray *et al.*, 2016). Abundance levels of transcripts within the RNA-seq libraries were calculated across three time-points (12hpi, 24hpi, 5dpi) and the mycelia library, with 100 bootstrap replicates for each library, using the JGI *P. cinnamomi* transcriptome as a reference.

After transcript abundance estimation, the Sleuth package (Pimentel *et al.*, 2017) was implemented in R to conduct the time-series differential expression using the Kallisto output. In Sleuth, transcript counts were filtered and normalized, and differentially expressed genes were identified using bootstrapping to account for variation within biological replicates and between different timepoints (Pimentel *et al.*, 2017). Statistically significant expression was determined using a Wald test, with the threshold for false discovery set as 0.05.

To identify and visualize transcriptionally active *RxLR*s, the gene identities of candidate *RxLR*s (gene IDs) were used as input to search for their expression in Shiny, which allows visualization of results generated by Sleuth. Additionally, the gene IDs of six *P. cinnamomi* candidate RxLR effectors predicted by Hardham & Blackman (2017) were included in the search. These six candidate effectors lacked the exact R-x-L-R motif, but they were classified as likely to encode RxLR effector proteins, due to similarity with characterized RxLRs in other *Phytophthora* species. The expression patterns of the candidate genes that were expressed during infection were then visualized as a heatmap in Shiny (in R).

For any *RxLR*s found to be expressed during infection, supposed genomic sequences, and predicted protein sequences, were obtained from JGI (USA Department of Energy, California).

Selection of expressed *RxLR*s for sequencing

Confirming the presence of required motifs within RxLR *genes*

To select expressed RxLRs for in-depth sequence analysis, gene sequences obtained from JGI were subjected to a second round of screening using an *in-silico* approach. Peptide sequences were manually screened in CLC Main Workbench 8.0.1 for the presence of an R-x-L-R amino acid motif within the first 60 amino acids in the N-terminal region. Selected genes were aligned according to previously mentioned parameters, but with alignment fixpoints set at the RxLR motif in each sequence. Sequences were also visually screened for the presence of a dEER motif downstream of the RxLR motif. The presence of a predicted signal peptide was confirmed by analyzing amino acid sequences using SignalP Version 4.1 (Department of Bio and Health Informatics, Technical University of Denmark). Protein sequences were submitted to SignalP 4.1 using default search parameters, searching for signal peptides specific to eukaryotes, with default cut-off values, that predicted signal peptides that could contain transmembrane segments. Proteins that listed “YES” in the SignalP 4.1 output for the signal peptide prediction were taken to have confirmed signal peptides present. Sequences with an output listing signal peptide presence as “NO” were discarded as potential RxLR effector proteins.

Determining similarity of putative RxLRs to proteins identified in other species

Putative *RxLR* genes that were found to be expressed during infection of susceptible avocado roots were screened for similarity to known effector proteins by using Basic Local Alignment Search Tool (BLAST) (National Center for Biotechnology Information, US National Library of Medicine, Maryland, USA). The JGI predicted protein sequences for chosen genes were investigated for similarity to known peptide sequences by searching a non-redundant protein database (nr) using the protein-protein BLAST (BLASTp) algorithm, without specifying organisms or exclusions.

Annotation of possible coding regions for putative RxLR*s*

The genomic sequences for expressed *RxLR*s were analysed by various gene prediction software to give an indication of which genes probably encoded proteins in reality. For the prediction of putative protein sequences, the resources used were: FGENESH (Softberry, Inc., New York, USA), EumicrobeDB (CSIR-Indian Institute of Chemical Biology, Kolkata, India), and the protein predictions previously obtained from the Joint Genome Institute.

Sequences that were previously shown not to contain a predicted signal peptide were excluded from further analysis subsequent to prediction software analysis. Where any of the software used predicted a protein sequence where the N-terminal region differed in sequence from that predicted by JGI, the alternate predicted protein sequences were also analysed in SignalP for presence of a signal peptide in the alternative protein sequence prediction.

Statistical analysis of expression of candidate RxLR*s*

To confirm the expression profiles for the chosen set of candidate RxLR effectors, the statistical analysis was repeated for the time-course infection of the susceptible avocado rootstock, using the RNA-seq read counts quantified as TPM in Kallisto. The Kallisto

outputs for each of the candidate genes were re-analysed in Microsoft® Office Excel v16.0.1230 (Microsoft® Corporation Inc., Redmond, Washington, USA), using the data generated across three timepoints (12hpi, 24hpi and 5dpi) after infection of the susceptible avocado rootstock R0.12, as well as data generated from RNA sequencing of *P. cinnamomi* mycelia. The statistical analysis of transcript abundances for the time-course was performed using a one-way ANOVA in XLSTAT v2019.3.2 (Addinsoft LLC., Montmartre, Paris, FRA). The same package was then used to perform a post hoc Fisher's Least Significant Difference (LSD) test, and statistical significance was evaluated using a confidence interval of 95%.

Confirming the genome sequence of the selected *RxLR*s

Extraction of P. cinnamomi *DNA*

P. cinnamomi GKB4 was grown on 20% V8 Agar for 4 days and the mycelia were harvested and crushed by manual grinding before 500 µl TES buffer (Tris pH 7.5, 10 mM; EDTA pH 8, 10 mM; sodium dodecyl sulfate 0.5% w/v) and 5 µl Proteinase K was added before vortexing for 3 s. Reactions were incubated at 60°C for 1 hour before 140 µl of 5 M NaCl and 65 µl of 10% CTAB was added. The tubes were vortexed and incubated at 60°C for another 10 min, before adding 905 µl chloroform isoamyl alcohol (24:1). Reactions were mixed and incubated on ice for 10 min. Samples were then centrifuged for 10 min at 12,000 rpm and the aqueous phase transferred to a sterile 1.5 ml tube. A volume of 440 µl isopropanol was added, reactions were mixed and incubated on ice for 30 min. Samples were again centrifuged for 5 minutes at 12,000 rpm, the aqueous phase was discarded, and the pellet washed twice with 1 ml of ice cold 70% ethanol. Pellets were allowed to dry by incubating the open tubes at 60°C for 5 min and then resuspended in 50 µl sterile water with 2 µl of RNase A (Thermo Fisher Scientific). DNA concentrations were determined using a NanoDrop 2000 (Thermo Fisher Scientific) and isolated DNA was diluted to working solutions of 100 ng/µl each.

To confirm successful isolation of *P. cinnamomi* DNA, the *LPV3* gene region was amplified and sequenced. Amplification reactions contained 1 U FastStart *Taq* DNA Polymerase (Roche Applied Science), 2.5 µl 10x PCR reaction buffer, 2.5 mM $MgCl_2$, 200 µM of each dNTP, 0.2 µM of forward and reverse primer, respectively, and sterile water to a total volume of 25 µl. The reaction was placed in a 2720 Thermal Cycler (Applied Biosystems, Thermo Fisher Scientific) which was set for an initial denaturation at 95°C for 6 min, 30 cycles of 45 s denaturation at 95°C, 30 s annealing at 55°C and 45 s extension at 72°C, followed by a final extension for 7 min at 72°C. Products were visualized on 1% agarose gel and sent to Inqaba Biotec™ (Inqaba Biotechnical Industries (Pty) Ltd, Pretoria, South Africa) for Sanger sequencing. The nucleotide sequence obtained from this PCR product was submitted as a BLASTn query (National Center for Biotechnology Information, US National Library of Medicine, Maryland, USA) to confirm isolate identity.

Amplification of RxLR *genes from* P. cinnamomi *DNA*

For amplification of gene sequences for the chosen set of RxLRs from the genome, primers were designed using Clone Manager (Clone Manager Professional Version 9, Sci-Ed Software). Primers were designed to anneal upstream and downstream of the coding sequences based on the combined results of all prediction software used. Where additional exons were predicted in the C-terminus of the peptide, two reverse primers were designed. Primer sequences for amplification from DNA are shown in Table 2.

Chosen genes (shown in Table 2) were amplified from *P. cinnamomi* GKB4 DNA by PCR using Phusion Green Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific). To optimize amplification conditions of the chosen genes, PCR of each gene was performed multiple times to account for differing primer annealing temperatures. For each gene, five 20µl reactions were set up, where each reaction tube contained the following final reagent concentrations: 1X Phusion HF buffer, 200 µM of each dNTP, 0.2 U Phusion Hot Start II High-Fidelity DNA Polymerase, 100 ng GKB4 DNA and 0.5 µM of each primer. A gradient PCR was used to determine optimal annealing temperature for

each primer set, and so the Veriti 96 Well Thermal Cycler (Applied Biosystems, Thermo Fisher Scientific) was set for an initial denaturation at 98°C for 1 minute, 25 cycles of 10 s denaturation at 98°C, 30 s annealing at a range of temperatures (61-65°C for *RxLR 1-4*, *9-14*, and *Hardham 5* and *6*; 66-70°C for *RxLR 5-8*) and 30 s extension at 72°C, followed by a final extension for 10 minutes at 72°C. For each gene, reaction tubes containing negative PCR controls were run at the same reaction conditions as the optimisation reactions, except that negative controls contained no *P. cinnamomi* DNA, and annealing temperatures for primers were carried out at the recommended temperatures calculated using the Thermo Fisher Scientific online Phusion T_m Calculator. The PCR products were then processed by electrophoresis on a 1% agarose gel and visualized under UV light. Successfully amplified genes were selected for subsequent cloning and sequencing; genes that failed to amplify after repeated optimisations were discarded from further experiments.

Candidate genes that were successfully amplified in optimization steps were then selected for large scale PCR to allow for gel extraction of these products later on. The reactions were thus carried out as before using Phusion Green Hot Start II High-Fidelity DNA Polymerase, with the same final concentrations of reagents, but made up to a total volume of 50 µl in each reaction tube, and each reaction was carried out in triplicate. The thermocycler was set as before, except that the final extension at 72°C was carried out for 7 minutes, and the annealing temperatures used for each gene were the highest temperatures previously optimized for each – *RxLR 1-4*, *9-14* and *Hardham 5* and *6* had an annealing temperature of 64°C; *RxLR 5-8* had an annealing temperature of 69°C. The products of the large-scale PCR were then processed by electrophoresis on a 1% agarose gel, visualized under UV light and segments of the gel containing the relevant products were cut out for extraction. Gel products were then purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research) according to manufacturer's instructions. The concentrations of the extracted DNA products were determined using a NanoDrop™ 2000 Spectrophotometer.

Sequencing of putative *RxLR* genes

Cloning of candidate RxLR*s and transformation of competent cells*

Once pure DNA products were obtained for each of the candidate RxLRs, these were cloned into a blunt vector to allow efficient sequencing. Cloning was performed using the Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen). The TOPO® cloning reaction was prepared according to manufacturer's guidelines, with 15-30 ng of PCR product added to each respective mixture (one reaction was prepared for each candidate *RxLR*), along with 1 µl salt solution, 1 µl pCR™II-Blunt-TOPO®, and sterile water to make up a final volume of 6 µl. Individual reactions were mixed gently, incubated at room temperature for 10 min and then placed on ice prior to transformation.

Transformation procedure was modified from the guidelines provided by the TOPO® cloning kit. A total of 6 µl of each prepared cloning reaction was added to respective tubes of 50 µl *Escherichia coli* DH5α competent cells, tapped lightly to mix the reactions and then placed on ice for 40 min before being heat shocked at 42°C for 90 s. Tubes were again placed on ice and incubated for 15 min, before 300 µl of Luria-Bertani (LB) broth was added under sterile conditions. The reactions were incubated at 37°C for 1 h with horizontal shaking (180 rpm). Volumes of 40 µl of each transformation reaction were then spread onto respective LB/Kan50 agar plates (2.5% w/v Luria-Bertani medium, 1.5% w/v agar bacteriological, 0.1% v/v 50 mg/ml Kanamycin), and plates were incubated in the dark at 37°C overnight.

Screening of putative RxLR *clones*

After the first overnight incubation, 16 putative *RxLR* clones were transferred from each plate to a new LB/Kan50 plate to ensure that growth of clones on selective media was not due to degraded antibiotics on the first plates. Clones were transferred by picking up single colonies from the original plates and streaking them into small sections on the new plates. In cases where there were less than 16 putative clones on the first LB/Kan50

plates, all of the available colonies were transferred. Plates were again incubated overnight at 37°C in the dark.

Following the second incubation, five clones of each *RxLR* were selected for PCR screening to confirm that putative clones were successful recombinants. Clones were transferred to 10 µl sterile water and then boiled at 99°C for 10 min to break open cells and release DNA. Tubes were centrifuged for 1 min at 12,000 rpm to remove cell debris. A volume of 3 µl of each supernatant was added as the template in respective PCR reactions. *RxLR*s were then amplified from each clone using insert-specific primers (as shown in Table 2) at PCR conditions optimized for Phusion Green Hot Start II High-Fidelity DNA Polymerase according to the protocol previously described, and products were analysed on a 1.5% agarose gel.

Extraction of plasmids for sequencing

Once recombination had been confirmed, three or four recombinant clones of each *RxLR* was selected for plasmid extraction prior to sequencing. Selected clones were transferred from LB/Kan50 streak plates to 5 ml LB/Kan50 broth and incubated at 37°C overnight with shaking (150 rpm). Plasmids were extracted using QIAprep® Spin Miniprep Kit (Qiagen) according to manufacturer's instructions. The concentration of plasmid DNA in the eluted solution was measured using a NanoDrop™ 2000 Spectrophotometer.

Successful extraction of recombinant plasmids was confirmed by PCR amplification using the vector-specific M13 primers (Fwd M13F-5'GTAAAACGACGGCCAGT3'; Rev M13R-5'CAGGAAACAGCTATGAC3'). The PCR was performed as before, at PCR conditions optimized for Phusion Green Hot Start II High-Fidelity DNA Polymerase, although the annealing temperature was amended to 55°C and the extension times for each reaction were increased by 15 s. Results of this PCR were analysed by gel electrophoresis.

Sanger sequencing of candidate RxLRs

Once isolation of recombinant *RxLR* plasmids was confirmed to be successful, extracted plasmids were sequenced by Sanger sequencing using BigDye® Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) and vector-specific M13 primers. Sequencing PCR reactions were set up in a volume of 12 µl for each extracted plasmid, where each reaction tube contained the following final reagent concentrations: 0.85X Sequencing buffer, 4.17% v/v BigDye 3.1, 0.83 µM primer, 40-200 ng plasmid DNA. The sequencing PCR was run in a Veriti 96 Well Thermal Cycler (Applied Biosystems, Thermo Fisher Scientific) set for an initial denaturation at 96°C for 5 s, followed by 25 cycles of 10 s denaturation at 96°C, 5 s annealing at a 55°C and 4 min extension at 60°C. Each plasmid was sequenced in both the forward and reverse direction.

For genes that could not be successfully cloned into sequencing vectors, three replicates of the PCR product of each *RxLR* were sequenced twice in each direction for three replicates of each gene. Gel extracted PCR product was cleaned up using ExoSAP-IT™ Express PCR product clean-up (Thermo Fisher Scientific) according to manufacturer guidelines. The purified PCR product was then sequenced by the same sequencing reaction used for plasmids, except that 40-100 ng of PCR product was added to each sequencing reaction, and the primers used in sequencing reactions were the *RxLR*-specific primers designed for original amplification from *P. cinnamomi* DNA.

Products of sequencing reactions were then precipitated using a sodium acetate clean-up. The full volume of each sequencing PCR product was added to a precipitation mix containing 2 µl NaAcO 3M, pH 5.2, and 50 µl 100% ethanol, made up to a final volume of 72 µl for each precipitation reaction. Tubes were incubated on ice for 10 min, then centrifuged for 30 min at 12,000 g. Each reaction was then subjected to two ethanol wash steps, where 150 µl 70% ethanol was added to each tube and tubes were centrifuged for 10 min at 12,000 g for each wash step. The supernatant in each tube was removed and DNA pellets were allowed to air dry at room temperature for 15 min. Samples were stored at -20°C until they could be sequenced. All samples were submitted to the DNA Sanger sequencing facility at the University of Pretoria for sequencing using an ABI 3500xl genetic analyser (Applied Biosystems, Thermo Fisher Scientific).

Obtained sequences were then analysed using CLC Main Workbench 8.0.1, where sequencing contigs were created using the forward and reverse sequences for each recombinant RxLR. Sequencing contigs were trimmed to only include *P. cinnamomi RxLR* sequences without additional vector sequences attached. The same software was then used to align the obtained sequences for all three plasmids for each of the cloned *RxLR*s with the original sequence obtained from JGI, and a genomic consensus sequence was generated for each candidate *P. cinnamomi RxLR*.

Manual annotation of candidate *RxLR*s

Once genomic sequences had been determined for the chosen set of candidate RxLR effectors, they could be manually annotated using Integrated Genome Viewer 2.7.2. (IGV), which allows transcriptome reads to be mapped to a selected genome.

First, genomic coordinates for the candidate genes were found in Geneious v7.06, using the newly assembled *P. cinnamomi* genome obtained by the ARP (unpublished data). A new sequence database was added using the Custom BLAST service in Geneious – this database consisted of all the scaffolds of the new genome in FASTA format and was named “PcGenome”. A second sequence database was created using the putative protein sequence predictions based on annotation of the *P. cinnamomi* genome. These protein sequences were inferred from gene prediction using BRAKER2 (Hoff *et al.*, 2019), and imported into Geneious as an amino acid (AA) file. The database created using the AA file was named “PcGenomeCDS”. To obtain the genomic coordinates using the Custom BLAST service in Geneious, the “Sequence Search” functionality was used with the confirmed *RxLR* DNA sequences as input, first against the PcGenome Custom BLAST database and then against the PcGenomeCDS database. For candidate *RxLR*s that did not have their DNA sequences confirmed experimentally in this study, the original DNA sequences obtained from JGI were used in the sequence search against the same databases described above.

To view the candidate *RxLR*s in IGV, the following files were imported into IGV: the FASTA file containing the *P. cinnamomi* genome scaffolds, the GFF3 file containing the BRAKER2 gene predictions, and the BAM file containing all the *P. cinnamomi* reads from the transcriptome generated by the RNA-seq experiment. The *P. cinnamomi* genome was then searched using the genomic coordinates according to contig matches obtained for each gene in the Geneious custom BLAST. Where transcripts from the BAM file mapped to the genomic coordinates provided, a consensus sequence was generated based on the regions of the gene that were expressed. If the coverage values of the genomic region in question was interrupted by sections without transcripts, these regions were considered introns and no consensus would be generated for that region of the gene. Alternatively spliced transcripts were manually annotated according to the coverage of the BAM file with the genomic region – where the number of reads mapping to the genome were decreased within the genomic region, but not reduced to zero, these were considered differentially processed transcripts, and were taken into account for manual annotations.

Manual prediction of RxLR protein sequences from manually annotated genes

The consensus sequences obtained for each of the candidate genes based on the transcriptomic reads mapping in IGV were imported into CLC Main Workbench 8.0.1. If introns were present within the genomic regions, these were annotated in CLC and a separate genomic sequence was created with the intronic region omitted. Consensus sequences for the candidate *RxLR*s were then translated to their respective protein sequences in CLC, using a six-frame translation. The resulting translations were then searched manually for the RxLR motif to find the correct protein sequence – the sequence was selected based on the presence of a signal peptide upstream of the RxLR motif, and the inferred protein sequence contained only the part of that translation making up the open reading frame (ORF) of the peptide; that is from the Methionine (M) N-terminal residue to the first stop codon (indicated by a * in the CLC translations). The final protein sequences were then submitted to SignalP 4.1 to confirm the presence of a signal peptide in the N-terminal regions of the predicted sequences.

Comparison of new predictions to original peptide sequences

Once final protein predictions were obtained for the set of candidate effector genes, a database was created in Geneious v7.06 containing the new protein sequence predictions for each gene, based on manual annotation in IGV. The original set of protein sequences for the chosen effectors, as predicted by JGI, was then compared to the new list by using a Custom BLAST to the database of new peptide sequences in Geneious v7.06. Where the Custom BLAST did not yield the expected results for a specific effector, the respective protein sequences for that effector were aligned using a ClustalW alignment in Geneious v7.06, using a BLOSUM cost matrix, with a gap open cost of 10 and a gap extension cost of 0.1.

Determining possible functions for candidate RxLRs based on inferred protein sequences

Phylogenetic analysis of manually predicted RxLR protein sequences

Once final putative protein sequences had been obtained for each of the candidate effectors, they were subjected to a phylogenetic analysis to determine whether any possible functions for the candidate effectors could be inferred based on their relatedness to known RxLR effectors in other species.

A multiple sequence alignment of N-terminal regions of the *P. cinnamomi* RxLRs with the characterised effectors listed in Table 1 was conducted in Geneious Prime 2020.0.3 using the previously described method. To obtain an optimal phylogenetic tree, the alignment was fine-tuned by removing certain characterised effectors which were observed to interfere with groupings of the sequences in the alignment, and redundant *P. cinnamomi* RxLR protein sequences (which has identical N-terminal regions) were discarded. Once the final sequence list was obtained, the alignment was repeated and then manually edited to remove uninformative regions as before.

The edited alignment of the N-terminal peptide sequences was then subjected to Bayesian inference analysis in MrBayes 3.2.7a using the previously described method. Based on the results of the tree generated by Bayesian inference analysis, a second alignment was performed using only the sequences which did not form outgroups in the previous phylogenetic tree. This alignment was performed in Geneious Prime as before, and the edited output was analysed by MCMC analysis in Mr. Bayes.

Alignment of full-length peptides for groups with inferred relatedness

To investigate whether any of the effectors with a supposed shared evolutionary history were notably similar in their full-length protein sequences, the proteins which formed clades with posterior probability support values in the phylogenetic tree were aligned to each other. In this alignment, the full-length amino acid sequences for each protein were used, rather than only their N-terminal regions. The alignment was performed in Geneious v7.06, using a MUSCLE alignment algorithm as before, with a maximum of 10 iterations and 5 trees to build, and the default alignment parameters were used.

Results

Identification of candidate *RxLR*s in *P. cinnamomi* genome

Data mining of the P. cinnamomi genome

A total of 26,131 predicted peptides were obtained for the *P. cinnamomi* proteome from JGI and submitted to SignalP 3.0 to determine which were likely to be secreted. A total of 2,313 had a signal peptide with a HMM score above 0.9, 2,138 of which had a NN cleavage site present between residues 10 and 30. Sequences of peptides in the secretome were then searched for the presence of an RxLR motif within 30 to 60 residues of the sequences; 192 putative *RxLR* genes were thus identified in the *P. cinnamomi* genome.

Phylogenetic analysis of predicted RxLR effector genes

The 192 candidate RxLR proteins were subjected to a multiple sequence alignment to determine whether there were obvious similarities between the peptide sequences. Originally, full peptide sequences for all candidates were used in the alignment, but the quality of this alignment was insufficient for further analysis due to the highly divergent C-terminal regions of the candidate effectors. The results of the full alignment are thus not shown here. It was then decided that only N-terminal regions of the candidate effectors would be used, similar to the alignment method used by Goss *et al.* (2013), in order to optimize the alignment for phylogenetic analysis.

The trimmed peptides were then successfully aligned using the MUSCLE alignment algorithm (Edgar, 2004). The consensus sequence generated for the alignment was 52 amino acids in length, containing a clear RxLR (RFLR) and dEER motif, as shown in Figure 1, and showed that each of the 192 candidates had the RxLR motif, although 37 sequences did not align with their RxLR motifs at positions 37 to 40, and 127 sequences contained a clear dEER motif.

The final edited alignment of the 192 N-terminal sequences was subjected to Bayesian inference analyses to construct a phylogenetic tree (Figure 2) with no major clades being evident. Several sequences were not similar to other sequences in the tree and formed outgroups indicated by red blocks in Figure 2. The tree contained 30 small clades, each containing two to eight sequences; however, probabilities could not be calculated for relatedness among the smaller clades due to sequence divergence. Only 15 of the 30 clades had posterior probabilities above 0.8.

A second phylogenetic analysis was performed to investigate the possibility of a shared evolutionary history between the 192 candidate RxLRs in *P. cinnamomi* and those in other species. The tree constructed by Bayesian inference analysis (Figure 3) contained the N-terminal sequences of the *P. cinnamomi* candidate RxLRs which had been aligned with the N-terminal sequences of RxLRs found in other species (Table 1). As before, several sequences dissimilar to other sequences formed their own outgroups (highlighted in red

blocks in Figure 3). The tree contained 40 smaller groupings of two to ten protein sequences each without statistical support for relatedness among them.

Several clades contained characterized RxLRs from other species that were similar to some of the *P. cinnamomi* candidates: Phyci_146196, Phyci_255366, and Phyci_142664 were related to *P. sojae* effectors PsAvh240, PsAvr3c and PsPSR2 (*Phytophthora* suppressor of RNA-silencing 2), respectively, with posterior probabilities of more than 0.9 for each. Phyci_81261 grouped with *P. infestans* effector PiSFI4 (suppressor of Flg-22-induced immune response 4), with a high posterior probability of 0.997, and Phyci_243372 was related to *P. parasitica* RxLR PpPSE1 (penetration-specific effector 1), with a probability of 0.939.

A few clades revealed a more distant relation between *P. cinnamomi* RxLRs and known effectors, with posterior probabilities above 0.75 but below 0.8. These include Phyci_87160 and Phyci_566358, which were grouped with *P. sojae* effectors PsAvr3b and PsAvr1d, respectively. *P. cinnamomi* candidates Phyci_588247 and Phyci_588244 were related to *P. infestans* PiPSFI5, with a posterior probability of 0.791.

Identifying candidate *P. cinnamomi* RxLRs likely to play a role in avocado infection

Dual RNA-sequencing of avocado rootstocks infected by P. cinnamomi *GKB4*

To narrow the set of 192 candidate effector genes to those that were likely to play a role in infection of avocado, data generated from an unpublished dual RNA-seq experiment were used. In this trial, clonal plantlets of susceptible and partially resistant rootstocks were inoculated with *P. cinnamomi* zoospores and harvested at different timepoints. RNA that was extracted from the samples was sequenced by paired-end sequencing, and compiled into 15 libraries (Table 3), made up of 3 biological replicates for each of 5 sample groups. A total of 875,981,985 paired-end reads were obtained, of which 177,074,104 reads (20%) were mapped to the *P. cinnamomi* draft genome. Libraries with the highest percentage of *P. cinnamomi* reads were those extracted directly from pathogen mycelia,

designated MS13-15. The library containing reads for the partially resistant rootstock was not used for further analysis in this study.

Identification of expressed candidate RxLR*s from RNA-seq data*

The reads generated from the RNA-seq experiment were searched using the gene IDs of the 192 candidate RxLRs to determine which were expressed *in planta*. The normalized read count for each of the candidate genes, obtained from the Kallisto output as transcripts per million (TPM) are listed in Table 4. According to the initial analysis of RNA-seq data, 26 of the 192 candidate *RxLR* genes were expressed at some level during infection of the susceptible avocado rootstock, when counts at the different timepoints were compared to the counts in *P. cinnamomi* mycelia.

The expression of these genes was visualized in the form of heatmaps generated by the Shiny application tool (Figure 4). The relative expression levels of genes were visually represented in heatmaps in a colour range of dark red (lowest or absent expression) to light yellow (highest relative expression. The colours used in visualisation of the expression profiles were allocated relative expression values on a scale of 0-6 according to the Shiny application.

Phyci_97951 and *Phyci_227271* were expressed inconsistently, at low levels and only within a few biological replicates across the timepoints – indicated by sparsely distributed orange blocks (shown as a relative expression level of ~2 on the colour scale of 0-6 generated by Shiny) in the heatmap in Figure 4. The colours visualised for these genes in the heatmap coincide with the TPMs generated by Kallisto, where less than 7 TPM were counted for each of the libraries for transcripts of these genes (Table 4).

Phyci_313963 and *Phyci_88750* were consistently expressed at a low level in mycelia and during infection of the susceptible rootstock, shown by the dark orange colour in the heatmap (~2-3 on the Shiny expression scale) across almost all biological replicates. These low levels of expression can be seen in the TPMs for each gene, where all libraries had read counts between 0 and 35.113 TPM (Table 4). The genes *Phyci_97174,*

Phyci_63598, Phyci_81261 and *Phyci_22058* were consistently expressed at intermediate levels in mycelia and during infection, as shown by orange colour of all biological replicates (~2-3 on the Shiny expression scale). The average count across the biological replicates for each of these genes was between 2.968 and 100.423 TPM, as quantified in Kallisto (Table 4). One gene, *Phyci_86482,* was consistently expressed at high levels in mycelia and across all timepoints, indicated by the yellow coloured biological replicates on the heatmaps (~6 on the Shiny expression scale), which reflect the high average counts of 200.503-1002.4 TPM for each timepoint.

Two of the genes in the dataset, *Phyci_65337* and *Phyci_31697,* had lighter orange blocks in the heatmaps (~3-4 on the Shiny expression scale) which indicated a peak in their expression at 24 hours post inoculation (hpi) compared to the darker red of other replicates (~0-2 on the Shiny expression scale). This visual observation reflected the TPMs seen for the genes at these timepoints in Table 4. There was a peak in expression of *Phyci_16230, Phyci_80229, Phyci_646* and *Phyci_90101* at 5 days post inoculation (dpi), shown by the lighter orange colours of the relevant biological replicates (~3-4 on the Shiny expression scale) when compared to the other timepoints. This expression profile is shown in the Kallisto outputs in Table 4, where the average TPMs for the 5 dpi libraries are higher than those for the other libraries, for all four of these genes.

Of the 26 genes in the expressed dataset, *Phyci_325329, Phyci_100005* and *Phyci_220952* were the only three putative RxLRs where expression was shown to be upregulated during infection when compared to mycelial expression – illustrated by the much lighter orange and yellow blocks (~6 on the Shiny expression scale) represented by the biological replicates at all timepoints during infection, but especially at the 24hpi and 5dpi timepoints.

Three of the expressed genes, *Phyci_76605*, *Phyci_206852* and *Phyci_96120,* were not chosen for further analysis due to their low and inconsistent expression patterns when compared to expression within mycelial tissue. Of the *P. cinnamomi* RxLR genes predicted by Hardham & Blackman (2017), only two of the six putative genes, *Phyci_24296* and *Phyci_297058*, were expressed during infection of avocado, with peaks

in their expression at 24hpi and especially at 5dpi and so these were also included for further analysis.

Selection of expressed *RxLR*s for sequencing

Confirming the presence of required motifs within RxLR *genes*

The putative protein sequences for the remaining 25 genes that were selected based on their expression profiles were then screened for the presence of a signal peptide and R-X-L-R amino acid motif (Table 5). All 23 of our own predicted effectors contained an RxLR motif within the first 60 amino acids of their predicted protein sequences, which can be visualized in an alignment created using the peptide sequences of all the genes (Figure 5). The two genes predicted by Hardham and Blackman each contained two RxLx motifs within the first 60 residues of their protein sequences, but they were kept for further analysis due to the high likelihood that they are valid effectors. Of the 25 peptides that were screened, nine also contained possible dEER motifs downstream of the RxLR region (Table 5).

Both of the candidate genes predicted by Hardham and Blackman (2017), *Phyci_24296* and *Phyci_297058*, renamed *Hardham 5* and *Hardham 6*, respectively, were confirmed to contain a signal peptide at the start of their predicted protein sequence according to SignalP 4.1. Four of the effectors predicted in this study, *Phyci_63598*, *Phyci_65337*, *Phyci_97951* and *Phyci_227271,* did not contain a signal peptide in their JGI predicted protein sequence according to SignalP 4.1, even though they had previously been predicted to contain signal peptides according to SignalP 3.0. Longer sequences for these genes were then obtained from JGI, with up to 1000 extra nucleotides on either side of the predicted coding sequence, but these longer sequences still did not contain a signal peptide when the translated sequence was submitted to SignalP 4.1. These gene sequences were then submitted to both FGENESH and EumicrobeDB coding sequence prediction software to determine whether there were alternate protein sequences predicted which might contain a signal peptide. All the predicted proteins had the same

N-terminal sequence that lacked a signal peptide, except for *Phyci_227271*, for which no reliable predictions were possible.

Determining similarity of putative RxLR*s to proteins identified in other species*

The 25 candidate RxLRs expressed during infection were further investigated for possible similarity to characterized effector genes from other species by submitting peptide sequences to BLASTp for analysis, and the top hit for each gene was recorded as the hit with the lowest Expect value (E-value) (Table 6). None of the putative effectors were shown to be highly similar to previously characterized RxLR effectors with confirmed function. Twenty of the 25 candidate RxLRs were shown to be most similar (based on E-values) to hypothetical proteins from other *Phytophthora* species, with 13 of these matching to hypothetical proteins from *P. sojae*, five which were similar to proteins from *P. parasitica*, and two with the best matches being to *P. cactorum* proteins. Of the 20 sequences with hits to hypothetical proteins, 14 had a coverage of over 90% of the amino acid sequence. Two of the remaining sequences with matches to hypothetical proteins, *Phyci_70517* and *Phyci_88750*, scored with a coverage above 80%, while four sequences, *Phyci_22058, Phyci_16230, Phyci_31697* and *Phyci_297058,* scored a relatively low coverage of between 30-48%.

Five of the candidate RxLRs did not yield hits to hypothetical proteins, but three of these proteins, *Phyci_325329, Phyci_22723,* and *Phyci_30885* matched to predicted RxLR or avirulence-like effector proteins from *P. megakarya* and *P. sojae*, respectively. The two remaining proteins included in the search, *Phyci_63598* and *Phyci_97951*, were shown to be similar to *P. sojae* housekeeping proteins with 100% coverage for each sequence. *Phyci_63598* was 98% similar to a phosphoribosyldiphosphate synthetase, while *Phyci_97951* had 88% similarity to a subtilisin serine protease, making these two *P. cinnamomi* proteins unlikely RxLR effectors.

Annotation of possible coding regions for putative RxLRs

The 25 candidate *RxLR* genes were further analysed *in silico* to predict coding regions and putative peptide sequences using three different resources: FGENESH, EuMicrobeDB and the predictions provided by JGI itself. Several genes had different predicted coding regions based on the different software used, represented visually in Figure 6.

Of the 25 candidate genes that were subjected to gene prediction software, only nine had the same coding regions predicted by all three software. Eight of the genes had the same prediction made by two of the programs, while the third prediction differed. Seven genes had entirely different coding regions predicted by all three programs, while one of the genes did not have any possible predictions using FGENESH or EumicrobeDB. The results of the predictions for coding sequences of the candidate genes are summarised in Table 7.

Based on the analysis of the 25 candidate genes, 16 putative *RxLR*s were chosen for further analysis, and these were renamed *RxLR 1-14* and *Hardham 5* and *6*, respectively.

Statistical analysis of expression of candidate RxLR*s*

In order to determine which of the candidate *P. cinnamomi RxLR*s were significantly upregulated during infection of a susceptible avocado rootstock, expression data obtained for *P. cinnamomi* mycelia as well as data from three timepoints (12hpi, 24hpi and 5dpi) were subjected to a second method of statistical analysis. The mean TPM, standard error (SE) and standard deviation (SD) for each gene was recorded (graphically represented in Figure 7).

Five of the candidate genes, *RxLR 9*, *13*, and *14* and *Hardham 5* and *6*, were not significantly upregulated at any timepoints during the course of infection compared to the mycelia. *RxLR 4, 5, 6* and *7* were all significantly upregulated in mycelia compared to the time-course of infection.

Expression levels of *RxLR 1, 2, 3* and *10* all peaked at 24 hpi, although this was only statistically significant for *RxLR 2* at 24 hpi compared to all the other timepoints. Upregulation of *RxLR 1* at 24 hpi was significant compared to mycelia and 12 hpi, but not 5 dpi, and upregulation of *RxLR 10* at 24 hpi was significant compared to both other timepoints but not compared to mycelia. Four of the candidate genes, *RxLR 5, 6, 8* and *9* were all significantly upregulated 5 dpi compared to expression in mycelia and at earlier timepoints.

Confirming the genome sequence of the selected RxLRs

Extraction of P. cinnamomi *DNA*

DNA was extracted successfully from three replicates of the *P. cinnamomi* isolate GKB4 at concentrations of 1027.3 ng/µl, 1517.5 ng/µl and 1504.1 ng/µl, respectively, and extracted DNA was diluted to working solution concentrations of 100 ng/µl. PCR amplification of the *LPV3* putative storage protein gene region from all three GKB4 isolates produced the expected 450 base pair fragment on the agarose gel (Figure 8). The sequenced *LPV3* PCR product was shown to be 99.5% identical to the *P. cinnamomi* putative storage protein LPV (Genbank accession number AF315064.1), with 99% coverage and an E-value of zero. Successful extraction of *P. cinnamomi* GKB4 DNA was therefore confirmed.

Amplification of RxLR *genes from* P. cinnamomi *DNA*

The candidate *RxLR* genes were successfully amplified within the expected size range, with the exception of *RxLR 4, 6, 7*, and *10*, which were excluded from further analysis. In the case of *Hardham 6*, limitations imposed by incomplete genome information resulted in the inability to obtain a singular PCR product from the available primers. Optimal conditions for amplification and expected product size for each of the *RxLR*s is listed in Table 8.

Each of the genes was amplified and extracted from the agarose gel at relatively low concentrations (Table 9). Since a singular product could not be obtained for *Hardham 6*, two of the PCR products, found to be near the expected length of 887 bp, were both extracted from the gel (Figure 9), and designated *Hardham 6B* and *Hardham 6C*, respectively. Concentrations of eluted products were low for the majority of the candidate genes but were sufficient for subsequent cloning steps and so they were still deemed suitable for use in downstream analysis.

Sequencing of putative *RxLR* genes

Cloning of candidate RxLR*s and transformation of competent cells*

The *RxLR* genes were cloned into a sequencing vector and used to transform competent cells. Colonies of transformed *E. coli* cells were obtained for all of the candidate genes with the exception of *RxLR 11* and *RxLR 14*, which failed to produce colonies on LB/Kan50 plates after multiple cloning attempts.

Screening of putative RxLR *clones*

Successful cloning of the candidate *RxLR* genes was confirmed by a colony PCR using insert-specific primers, where five colonies were screened for each gene. For all of the putative recombinants that produced colonies on LB/Kan50 plates, the respective *RxLR* genes were successfully amplified in colony PCR screening steps, with *RxLR* inserts producing the expected band sizes on agarose gels. An example of the results of selected colony PCRs is shown in Figure 10.

Extraction of plasmids for sequencing

Plasmids were successfully extracted for at least three colonies for each cloned candidate *RxLR.* The concentration of extracted plasmid DNA is shown in Table 10.

To confirm that plasmids were successfully extracted and contained the desired *RxLR* inserts, the candidate genes were amplified using vector-specific M13 primers. Amplification using the M13 primers confirmed that plasmids containing *RxLR 1, 2, 3, 5, 8, 9, 12, 13, Hardham 5, Hardham 6B* and *Hardham 6C* were successfully extracted.

Sanger sequencing of candidate RxLRs

Full-length DNA consensus sequences were obtained for all candidate genes, after aligning the genes to the expected genomic sequences obtained from JGI. Consensus sequences were created for *RxLR 1, 2, 3, 5, 8, 9, 11 12, 13, 14, Hardham 5, Hardham 6B* and *Hardham 6C.* In most cases, the DNA sequences were similar to the sequences obtained from JGI, with the exception of several single nucleotide polymorphisms (SNPs) that indicated differences between the alleles of the strain used in this study (GKB4), or single nucleotide differences between strains (GKB4 compared to the JGI strain *P. cinnamomi var cinnamomi* v1.0). In the case of *RxLR 3* and *Hardham 5*, the DNA sequences obtained from the GKB4 isolate were missing nine bases and 21 bases respectively when compared to the JGI genomic sequence (Figure 11). Since these deletions were in multiples of three, they would not change the reading frame, although it is possible that the amino acid sequence of the final effector proteins would be altered.

Other notable sequencing results showed that *Hardham 6B* and *Hardham 6C* were nearly identical in sequence, except for five SNPs that differed between the two, and a 54 base deletion in *Hardham 6C* compared to *Hardham 6B* and the JGI genomic sequence. Subsequently, both *Hardham 6* sequences were maintained to be analysed further.

Manual annotation of candidate RxLR*s*

Once the confirmed genome sequences were obtained for each candidate *RxLR*, the genomic coordinates for these sequences could be obtained. The sequences were searched against the newly assembled *P. cinnamomi* genome and its putative annotations (ARP, unpublished data) using the custom BLAST service in Geneious v7.06.

Since the genomic sequence for *RxLR 4*, *RxLR 6*, *RxLR 7* and *RxLR 10* could not be experimentally confirmed by this study, the original DNA sequences for these genes, as obtained from JGI, were used in the custom BLAST.

The results of the custom BLAST are summarised in Table 11a for the PcGenome database, and Table 11b for the PcGenomeCDS database. Where a gene had multiple hits to either or both genome databases, all of the most likely hits were recorded in the results tables, based on the coverage of each hit and the similarity between each hit and the input DNA sequence. In Table 11a, the hit start and hit end are recorded, since these are the values that were used as the genomic coordinates to search the *P. cinnamomi* genome contigs in IGV.

In most cases, the "jg" transcript ID for the PcGenomeCDS hits could be used to find the location of the hit on the *P. cinnamomi* genome in IGV. In cases where there were more hits to the PcGenome than to transcript IDs in the PcGenomeCDS, such as for *RxLR 1*, *RxLR 3* and *RxLR 6*, the coordinates for hit start and hit end were used to search the contigs that did not have a corresponding CDS prediction. This was also the case for *RxLR 11*, which did not produce hits to any of the predicted protein sequences in the genome.

Once the prospective gene loci were found in IGV for each of the candidate genes, a consensus sequence could be generated based on the coverage of the genomic region by *P. cinnamomi* transcripts. A visual representation of the coverage analysis is shown in Figure 12, using the genomic location of *RxLR 1* on contig 14 as an example. Some of the regions that were considered possible loci for the selected candidate RxLRs according to the custom BLAST did not have any transcripts mapping to them in IGV – an example is shown in Figure 13. Where a predicted genomic locus for a candidate *RxLR* did not have RNA-seq reads mapping to the genomic location in question, a consensus sequence was not generated, and the region was not considered as a true genomic location for the candidate *RxLR* in this study. Of the 33 predicted loci based on BLAST hits to genomic contigs, 10 genomic loci did not have sufficient coverage of the region by RNA-seq reads - these are shaded in grey in Table 11a. These predicted genomic coordinates for *RxLR 1* (contig 32), *RxLR 2* (contig 38), *RxLR 3* (contigs 66, 37 and 76),

RxLR 6 (contigs 1, 27 and 8) and *RxLR 10* (both hits to contig 51) were discarded as true loci for the candidate genes in this study, since candidate *RxLR*s were originally selected based on the fact that they were shown to be expressed in infected avocado.

DNA consensus sequences were successfully generated for the 23 remaining predicted genomic locations of the candidate *RxLR*s, based on the coverage of the genomic region by RNA-seq reads. Regions that contained clear introns did not produce consensus sequences for the intronic regions – an example of visually observed introns is shown in Figure 14 for *RxLR 7*. The candidate *RxLR*s that were observed to contain introns were *RxLR 4* (contig 28 and contig 1 hits), *RxLR 6* (contig 47), *RxLR 7* (contig 14), *RxLR 14* (contig 12) and *Hardham 6B* (contig 50). Locations of introns are listed in Table 12.

Manual annotation of *RxLR*s using IGV confirmed that several of the candidate genes had their CDSs accurately predicted by BRAKER2 according to predicted genomic loci. Of the 23 remaining genomic coordinates that were considered possible loci for the candidate *RxLRs*, 10 genes had accurate BRAKER2 predictions. *RxLR 1* (contig 14 jg9436.t1), *RxLR 4* (contig 1 jg17626.t1 and contig 28 jg10431.t1), *RxLR 5* (contig 63 jg7968.t1), *RxLR 7* (contig 14 jg9835.t1), *RxLR 9* (contig 50 jg4237.t1), *RxLR 10* (contig 18 jg15182.t1), *RxLR 12* (contig 62 jg15774.t1), *RxLR 13* (contig 120 jg16635.t1) and *RxLR 14* (contig 35 jg1103.t1) all had their start and stop codons and intron/exon boundaries correctly predicted by putative genomic annotations.

Four of the prospective genomic loci, *Hardham 5* (contig 1 and 53), *RxLR 11* (contig 67) and *RxLR 13* (contig 68), did not have any CDS predictions according to BRAKER2, but genomic coverage by RNA-seq read confirmed that these regions were expressed, and consensus sequences were successfully generated. The genomic locus for *RxLR 2* on contig 123 had a long predicted CDS which included two introns and three exons spanning ~825 bp, but RNA-seq reads mapped to only the C-terminal part of this prediction, illustrated that no introns were present, and the consensus sequence generated was only 566 bp in length. For the predicted genomic locus for *RxLR 3* on contig 24, the predicted CDS, jg5887.t1 spanned over 22 kb, with four exons and three introns, one of which was over 21kb in length. When this locus was viewed in IGV, it was observed that RNA-seq reads mapped to only contig 24:315,155-314,586, part of the

predicted intron, and the consensus sequence generated for *RxLR 3* was only 583 bp long, with no introns observed according to genomic coverage by RNA-seq reads.

Manual annotation of *RxLR 4* contig 1 was identical to the annotation predicted by BRAKER2. It could be visually observed, however, that this transcript was alternatively spliced by *P. cinnamomi,* but alternatively spliced transcripts were not predicted by BRAKER2 for this genomic region. Manual annotation of alternative splicing according to IGV results is illustrated in Figure 15, which shows that the coverage track of *RxLR 4* contig 1 was interrupted by a short region with a number of reads that was reduced to approximately half the quantity of the reads adjacent to it. This region with reduced reads was annotated as an intron that was subject to alternative splicing, and both splice variants were predicted from the consensus sequence obtained for this region.

The IGV observed results for *RxLR 6* contig 47 were similar to those noted for *RxLR 2* contig 123, where the predicted CDS, jg7968.t1, had two introns and three exons, but reads only mapped to the C-terminal region of the prediction. Unlike the *RxLR 2* locus, however, the reads mapping for *RxLR 6* did not end where the predicted CDS ended. Instead, RNA-seq reads spanned the neighbouring CDS prediction, jg7969.t1, as well. Since the coverage of the region was not interrupted by breaks which would have indicated that separate genes were being expressed, the entire expressed region was used to produce the consensus sequence.

RxLR 8 had two hits to predicted CDSs which were both located on Contig 5 – the predicted transcripts were designated jg3373.t1 and jg3405.t1. For both of these hits, the RNA-seq reads did not only cover the predicted coding regions, but spanned the respective neighbouring predictions, jg3374.t1 and jg3404.t1, as well. Intron/exon boundaries were unclear by visual observations, however, and so no introns were taken into account for *RxLR 8* annotations. Since both of the hits for this gene spanned multiple CDs predictions, the entire region with RNA-seq read coverage was used to produce the consensus sequence for each hit.

The genomic hit for *RxLR 4* contig 12 had a CDS with one predicted intron, which coincided with manual annotation. The predicted coordinates of this intron differed from

those observed visually in IGV, however, and so the manually annotated intron was taken into account when the consensus sequence was generated for this region.

Finally, *Hardham 6B* and *Hardham 6C* both had two hits to the same regions of contig 50, but with different similarity scores. They also both produced hits to the same predicted CDS, jg4255.t1, which was predicted to span 3.9 kb, with an intron predicted to span 2.7 kb of the CDS. When the relevant region of contig 50 was visualised in IGV, it was apparent that the two regions of jg4255.t1 predicted as exons, instead formed two distinct coding regions – this is visually depicted by the coverage of the regions in question (Figure 16), where read numbers gradually reduce towards the end of transcribed regions, rather than suddenly decreasing as is seen when introns are present.

To test the hypothesis that these two regions represented *Hardham 6B* and *Hardham 6C, respectively*, the DNA sequences obtained for these candidate *RxLR*s from earlier experimental steps were assembled to the full consensus sequence of contig 50; 200,552 – 196,642 in Geneious v.7.0.6 (Figure 17).

Hardham 6B assembled to the upstream region of the genomic locus (contig 50: 200,552 -199,839) that had been predicted as the first exon of jg4255.t1, while *Hardham 6C* assembled to the downstream region of the locus (contig 50:197,100-16,642), which had been predicted as the second exon of jg4255.t1. Mapping of these two candidate *RxLR*s to the different sections of this region proved that they formed two distinct coding regions on the same genomic contig, rather than being two different exons of the same transcript, as had been predicted.

When the two regions of *Hardham 6* were inspected separately in IGV, it was evident that *Hardham 6C* produced a single transcript without introns. For *Hardham 6B*, however, regions of reduced coverage by RNA-seq reads showed the presence of two likely introns, which were subjected to alternative splicing. The intronic region at the coordinates contig 50:200,253-200,191 was easy to observe visually, and similar to the intron in *RxLR* contig 1, it consisted of an area where the number of RNA-seq reads was reduced by approximately 50%. A second putative intron could be observed as starting at contig 50:200,535, but no distinct end to this intron could be annotated according to RNA-seq

read coverage of the region, and so this potential intron was not taken into account when consensus sequences were generated for alternative splice variants.

Manual prediction of RxLR protein sequences from manually annotated genes

Once consensus sequences had been generated for the coding regions by manual annotation of each candidate *RxLR*, they were translated into six reading frames to predict the correct protein sequence for each putative effector. The six-frame translations were then manually searched for the presence of the RxLR motif to determine the correct reading frame. Protein sequences for each candidate were then saved and the relevant reading frame recorded in Table 12.

Final protein sequences could be predicted for 22 annotated CDSs of candidate *RxLR*s, which included two splice variants each for *RxLR 4* and *Hardham 6B*. The peptide sequence for the second splice variant of *RxLR 4* contig 1 was identical to the peptide sequence of *RxLR 4* contig 28. None of the translations for *RxLR 6* contig 47 contained the desired RxLR motif, and so this was taken as a falsely predicted *RxLR*. For *RxLR 13*, the protein predicted from translation of the contig 68 consensus sequence was truncated, since RNA-seq reads did not span the region which included a stop codon. *RxLR 14* contig 12 produced a protein sequence that was notably shorter than protein predicted from contig 35 of the same gene, having only 25 residues after the dEER motif before the translation termination site was found.

The protein prediction for *Hardham 5* contig 53 was truncated at the N-terminal region. For this candidate, the -2-frame translation contained a region identical to the C-terminal region of the original Hardham 5 prediction, but there was no Methionine residue or signal peptide in the N-terminal region of this prediction. The predicted locus of contig 53 was thus disregarded as the correct location for the expressed *Hardham 5* gene.

The final protein predictions were then submitted to SignalP 4.1 to confirm they could all be considered potential RxLR effector proteins. All of the manual protein predictions listed

in Table 12 contained a signal peptide, and so they were all maintained for further analysis.

Comparison of new predictions to original peptide sequences

In order to determine the similarity between the original protein predictions and the new predictions obtained by manual annotation, the new sequences were compared to the original sequences using a Custom BLAST. Results for the BLAST hits for each sequence are summarized in Table 13. Overall similarity was analysed according to the grade score provided by Geneious Prime 2020.0.3, since this score is weighted according to the pairwise similarity, coverage and E-value of each hit.

Only two of the candidate effectors were exact matches to the original protein predictions – these were RxLR 12 and both of the coding regions predicted for RxLR 8, which each had a grade of 100% for the BLAST hit to their respective original predictions. Seven of the new protein predictions were more than 95% (but less than 100%) similar to their original predictions; RxLR 2, RxLR 3, splice variant 1 of RxLR 4, RxLR 7, RxLR 11, and both coding regions for RxLR 13. For these candidate effectors the differences were mainly due to single amino acid substitutions which differed between the original proteins and the new predictions (Figure 18).

Interestingly, the new peptide sequence predicted for RxLR 5 had a BLAST score of 99.4% overall similarity to the original prediction for RxLR 4, while it did not have any matches to the original RxLR 5 protein prediction. To confirm this, the new predicted peptide sequence for RxLR 5 was aligned to the original sequence obtained from JGI. This alignment, shown in Figure 19, exhibited very little overall similarity between the sequences – individual identical residues were present, but these were separated by spans of residues that had no similarity, and a full consensus sequence could not be generated for the alignment.

For the coding region found on contig 1 for RxLR 4, the two predicted splice variants differed in their scores for BLAST hits to the original, but the coding region for RxLR 4 on

contig 28 had the same score as the second splice variant of RxLR 4_contig 1. When all the coding regions for RxLR 4 and the new prediction for RxLR 5 were aligned with the original RxLR 4 prediction (Figure 20), it was evident that the new RxLR 5 prediction and RxLR 4_contig 1_splice variant 1 were both nearly identical to the original prediction. RxLR 4_contig 1_splice variant 2 was identical to RxLR 4_contig 28, and both differed from the original sequence at their C-terminal ends.

In the case of Hardham 6, all new predictions were only a 60-66.5% match to the original predicted peptide. Notably, all the new Hardham 6 peptide predictions were revealed by the Custom BLAST to be more similar to the original Hardham 5 prediction than that of Hardham 6 (results not shown). Due to the diversity of the new Hardham 6 predictions, all the new peptide sequences for this effector were included in an alignment with the original prediction for visual observation (Figure 21). The alignment showed that though the sequences were similar enough to one another to generate a reliable consensus sequence, each of the new predictions was unique, and not one of the new sequence predictions was exactly the same as the original prediction.

Determining possible functions for candidate RxLRs based on inferred protein sequences

Phylogenetic analysis of manually predicted RxLR protein sequences

After the protein sequences of the candidate RxLRs had been determined following manual annotation, a phylogenetic analysis was performed to investigate whether the informed protein sequences could reveal possible functions for the candidate *P. cinnamomi* RxLRs based on their evolutionary relatedness to characterised effectors of other *Phytophthora* spp. To that end, a tree was constructed by Bayesian inference analysis of the alignment of N-terminal regions of *P. cinnamomi* candidates with known RxLRs. In this case, the redundant *P. cinnamomi* protein sequences RxLR 13_contig 68, RxLR 14_contig 12, RxLR 8_jg3405 and RxLR 4_contig 1 (both variants) were removed prior to phylogenetic analysis, since they had the same N-terminal region as RxLR 13_contig 120, RxLR 14_contig 35, RxLR 8_jg3373 and RxLR 4_contig 28, respectively.

The phylogenetic tree resulting from Bayesian inference analysis contained several sequences without similarity to any other sequence, which formed their own outgroups, all highlighted within the red block in Figure 22. The *P. cinnamomi* candidate effectors RxLR 3, 7, 8, 10 and 13 were included in this group, and are indicated by blue stars within the red block. No functional inferences could thus be made for these five candidate effectors, since they were not shown to be evolutionarily related to any of the characterised RxLRs included in this study.

For the sequences where evolutionary inferences could be made, the tree contained 10 smaller clades of two to ten protein sequences each, though there was no statistical probability to support relatedness between any of these small clades. Several of these clades contained characterized RxLRs from other species which exhibited similarity to some of the *P. cinnamomi* candidate effectors. Characterised RxLRs which grouped with *P. cinnamomi* candidates with a probability above 0.7 are indicated by green arrows in Figure 22.

The candidate *P. cinnamomi* effectors RxLR 9, Hardham 5 and Hardham 6 (all variants) formed a clade with *P. infestans* PiSNE1, with a posterior probability of more than 0.9 for their evolutionary relatedness. RxLR 1 was shown to be more distantly related to *P. sojae* Avr1b and *P. capsici* Avr3a12, with a posterior probability of 0.7811 for this grouping. Another candidate with more distant evolutionary inferences was RxLR 14, which formed a clade with *P. infestans* Avrblb1 and *P. parasitica* RxLR 2, with a supported probability of 0.7416.

Evolutionary grouping in this phylogenetic tree with lower probabilities include relatedness between RxLR 11 and *P. infestans* Pi14054, RxLR 2 with *P. sojae* Avh238 and the grouping of RxLR 4 and 5 with *P. infestans* Pi03192. The characterised RxLRs in these clusters are shown by yellow arrows in Figure 22, and posterior probability for each of these groups was below 0.7.

A second phylogenetic analysis was performed, using only the sequences which did not form outgroups indicated by the red block in Figure 22. It was hypothesised that this second analysis would provide better resolution for the clades which did form in the tree

in Figure 22, if there were no outgroups to interfere with the alignment. The resulting tree is shown in Figure 23.

The repeated phylogenetic analysis confirmed some of the evolutionary groupings shown in the previous tree (Figure 22), such as the grouping of RxLR 9, Hardham 5 and Hardham 6 with *P. infestans* PiSNE1, which had probability values above 0.9 for this cluster in the second tree as well. *P. cinnamomi* RxLR 14 formed a clade with *P. parasitica* RxLR 2 again, with a supported probability of 0.795. RxLR 2 was shown to be related to *P. sojae* Avh238 in this second tree as before, but again the support value was less than 0.7 – in this case the posterior probability was 0.513.

Anomalies in the second tree compared with the first include the complete separation of RxLR 4 and 5 from the rest of the tree, where they formed a divergent outgroup instead of grouping with *P. infestans* Pi03192 as they did previously. *P. cinnamomi* RxLR 12 formed an outgroup by itself rather than falling into a clade with any other sequences. Finally, RxLR 1 and RxLR 11 fell within the larger clade which contained all the ingroups in this tree but did not form smaller clades with known effectors as they did previously.

Alignment of full-length peptides for groups with inferred relatedness

To investigate whether the phylogenetic inferences made from N-terminal alignments of proteins were relevant when the full-length peptides were analysed, alignments were generated for the proteins which grouped within the smaller clades of the first phylogenetic tree (Figure 22) using their complete amino acid sequences. RxLR 11 was aligned with *P. infestans* Pi14054, RxLR 14 was aligned with *P. infestans* Avrblb1 and *P. parasitica* RxLR 2, RxLR 1 was aligned with *P. sojae* Avr1b and *P. capsici* Avr3a12, RxLR 4 and 5 were aligned with *P. infestans* Pi03192 and PiSFI4, and RxLR 9, Hardham 5 and Hardham 6 were aligned with *P. infestans* PiSNE1.

The alignment of RxLR 11 and Pi14054 showed little overall similarity between the two peptides, with only single residues generating a consensus, with spans of dissimilar residues between them. This was also the case for the alignment of RxLR 4 and 5 with

P. infestans Pi03192 and PiSFI4, and the alignment of RxLR 1 with *P. sojae* Avr1b and *P. capsici* Avr3a12, where the longest consensus could be found in the N-terminal region containing their signal peptides.

The alignment of RxLR 2 with *P. sojae* Avh238 had more similarity overall, with several short regions of consensus separated by short regions of divergent residues (Figure 24a). When RxLR 14 was aligned with *P. infestans* Avrblb1 and *P. parasitica* RxLR 2, low overall similarity was visible in the consensus (result not shown), but when this alignment was refined to contain only *P. cinnamomi* RxLR 14 and *P. parasitica* RxLR 2, more overall similarity was observed, with short regions of consensus separated by short regions of dissimilar amino acids (Figure 24b). Finally, the alignment of *P. cinnamomi* candidate effector proteins which had grouped in the clade with *P. infestans* PiSNE1 showed the highest overall similarity, with several regions in their C-terminal domains being conserved (Figure 24c).

Discussion

A total of 192 candidate *RxLR*s were identified in the *P. cinnamomi* genome using the pipeline developed by Win *et al.* (2007). This number of candidate genes was comparable to the 171 putative *P. cinnamomi RxLR*s predicted by Hardham & Blackman (2017), but was considerably higher than the 68 *P. cinnamomi RxLRs* predicted by McGowan & Fitzpatrick (2017). Nonetheless, the 192 predicted in the present study is much lower than predictions of 563 for *P. infestans*, 375 for *P. ramorum* and 396 for *P. sojae* (Jiang *et al.*, 2008, Haas *et al*., 2009). The large number of *RxLR*s predicted for *P. infestans* compared to *P. cinnamomi* is, perhaps, understandable given its larger genome (respectively, 240 Mb vs. 78) (Haas *et al.*, 2009). However, since the genomes of *P. ramorum* and *P. sojae (*65 Mb and 95 Mb, respectively) are comparable in size to that of *P. cinnamomi*, genome size may not impact effector number.

The smaller number of *P. cinnamomi* candidate *RxLR*s predicted in this study may be due to the method used. When *RxLR*s were predicted from the genomes of *P. sojae* and *P. ramorum* using the same pipeline used in this study, only 189 and 214 candidates were

predicted for each respective species when only candidates with a confirmed RxLR and dEER motif were considered (Win *et al.*, 2007). The number of *RxLR* effectors predicted above and for *P. cinnamomi RxLR* effectors, are likely underestimates of the true number of *RxLR* candidates, since this prediction method requires the presence of a RxLR motif in the N-terminal regions of the sequences (Win *et al.*, 2007). Other researchers have shown that variants to the traditional RxLR motif exist, and that all RxLR effectors might not contain the exact RxLR motif in their N-terminal regions (Fabro *et al.*, 2011, Kale & Tyler, 2011, Tian *et al.*, 2011, Stassen *et al.*, 2013). This means that the 192 candidate *P. cinnamomi RxLRs* identified in this study are not likely to be a true representation of all the RxLR effectors that make up this oomycete's effector arsenal, since variant motifs were not accounted for (Bozkurt *et al.*, 2012). Nonetheless, the 192 RxLRs that were predicted in the present study were considered sufficient and provided a starting point for the discovery of *P. cinnamomi* effectors that contribute to virulence during PRR development on avocado.

The 192 putative RxLR proteins were subjected to phylogenetic analysis to determine whether any of the candidates formed groups with similar sequences within the dataset. The predicted peptide sequences of the candidate RxLRs was used in phylogenetic analysis, since the sequences of the effector proteins were too divergent to enable the use of nucleotide sequences in the alignment. The original alignment of the sequences was generated using the full-length protein sequences of the RxLRs, but the C-terminal domains of the peptides were too divergent to be informative in terms of the evolutionary relatedness between proteins. The sequences were thus trimmed to only their N-terminal regions leading up to their dEER motifs, or up to 80 residues if there was no dEER motif present. These N-terminal sequences were then used in the alignment and phylogenetic analysis that followed, similar to the method used by Goss *et al.* (2013). It was postulated that using only the N-terminal regions of sequences would still provide sufficient information regarding the relatedness of the proteins, while providing more optimal alignments due to the removal of uninformative sequences.

Phylogenetic analysis of the 192 RxLRs was performed using a Bayesian inference analysis to provide the tree with the best resolution in terms of sequence grouping.

Original phylogenetic analysis used Maximum Likelihood methods as used by Goss *et al.* (2013), but these trees failed to produce clades with sufficient bootstrap values, and so the results have not been included for the purposes of this study. The phylogenetic tree produced by Bayesian inference analysis had the best resolution of clades compared to other tested methods, though posterior probability values could not be generated for relatedness between clades. This result was justifiable, since RxLR effectors are known to show high sequence divergence (Jiang *et al.*, 2008), and so it stands to reason that groups of related effectors might not be related to each other. Several sequences did not show similarity to any other sequences within the tree, and instead formed outgroups that did not fall within larger clades. This was again hypothesised to be due to the divergent nature of the dataset used. Despite the fact that not all of the sequences fell within clades on the phylogenetic tree, and that probability values could not be generated for relatedness between smaller clades, the tree produced by Bayesian inference analysis remained the optimal phylogenetic tree that could be produced for this dataset.

The phylogenetic analysis of the RxLRs was repeated with a larger dataset, to investigate the possibility of a shared evolutionary history between the 192 candidate *P. cinnamomi* RxLRs and characterised RxLR effectors in other species. Several of the candidate *P. cinnamomi* RxLRs were shown to be related to RxLRs with assigned functions, with a posterior probability value above 0.9. The candidate effectors Phyci_146196 was shown to be related to *P. sojae* Avh240 (posterior probability 0.996), which interacts with host aspartic proteases in the soybean host to interfere with defense responses (Guo *et al.*, 2019). Phyci_81261 formed a clade with *P. infestans* SFI4, which suppresses pattern-triggered immunity (PTI) in host plants during the biotrophic phase of infection (Zheng *et al.*, 2014). Phyci_255366 grouped with *P. sojae* Avr3c, an RxLR that interferes with pre-mRNA splicing in host cells to undermine immune responses. Phyci_142664 fell within the same clade a *P. sojae* PSR2, an effector which suppresses RNA-silencing to interfere with host defences. Another candidate RxLR, Phyci_243372, showed relatedness with the recently characterised *P. parasitica* PSE1, which has been shown to disrupt hormone signalling in the host by manipulating auxin physiology. The five examples listed here all had posterior probability support values above 0.9 for their relatedness to respective characterised RxLRs. It can therefore be hypothesised that these five candidates might

have functions related to those of the known effectors, but it should be taken into account that these effectors may only be expressed in certain hosts and may not have a virulence function in the avocado host.

A total of 25 candidate RxLRs were predicted to play a role during *P. cinnamomi* infection of avocado based on their expression profiles *in planta*, which is comparable to observations made in *P. sojae*, in which only a small subset of the predicted RxLR effectors appear to contribute to virulence in the soybean host (Wang *et al.*, 2011). It remains unclear, however, whether the subset of *P. cinnamomi* effectors expressed in our dataset represent a set of key effectors which contribute to virulence in all hosts, or whether the set of expressed RxLRs would differ depending on the host plant that is infected.

The 25 candidate RxLRs were further analysed to confirm their annotation as likely RxLR effectors, so that false predictions could be discarded for additional study. Four of the predicted effectors, Phyci_63598, Phyci_65337, Phyci_97951 and Phyci_227271, did not contain a signal peptide according to the latest algorithm, making them likely false positives in the effector prediction dataset. When this result was cross-referenced with the results of BLAST analysis of the predicted effectors, Phyci_63598 and Phyci_97951 showed high similarity to the *P. sojae* housekeeping proteins; a phosphoribosyldiphosphate synthetase and a subtilisin serine protease, respectively. The likelihood of these two candidates functioning as housekeeping proteins in *P. cinnamomi* supported the hypothesis that these candidates were falsely predicted RxLRs.

All of the 25 candidates were subjected to a manual search for RxLR and dEER motifs. All of the putative effectors identified from the genome contained the traditional RxLR, while the two candidates predicted by Hardham & Blackman (2017), Phyci_24296 and Phyci_297058, each contained three possible variants of the RxLR motif in their N-terminal regions; RNLA, RQLG and GGLR. The results of the multiple sequence alignments performed later in this study indicated that the GGLR variant was the most like “RxLR” motif in these two effectors, similar to the GKLR motif found in the effectors BLG01 and BLG03 in *Bremia lactucae* (Stassen *et al.,* 2013). This finding is in contrast to the hypothesis of Hardham and Blackman (2017), who suggested that the RxLR variant

motif in these genes was one of the RxLx (RNLA or RQLG) motifs in their N-terminal regions. The inference of GKLR as the “RxLR” motif in these effectors provided further evidence for the idea of a variant RxLR motif put forward by Kale and colleagues (2010). The discovery of the variant motif also has implications for prediction of RxLR effector genes by data mining methods, since methods which mine for proteins which have the exact RxLR motif in their N-terminal regions would likely underestimate the number of RxLRs present in a pathogen’s genome (Bozkurt *et al*., 2012).

Although only nine of the candidate RxLRs were shown to contain a dEER motif, for unclear reasons RxLR motifs are inconsistently followed by a dEER motif in other RxLR effectors (Tyler *et al.*, 2006, Jiang *et al.*, 2008, Tyler, 2011). Several RxLR effectors have been shown to contain WY domains in their C-terminal region, and these domains are believed to produce a conserved structural fold which signifies a function in suppression of PCD (Jiang *et al.*, 2008, Dou *et al.*, 2008b). Unfortunately, the presence of C-terminal domains was not analysed in this project. To elucidate possible functions based on the effector structures, the presence of WY domains in candidate *P. cinnamomi* RxLRs should be determined in the future.

Based on their expression profiles and further analysis, 25 candidate effectors analysed in the present study were reduced to a final number of 16, which were named RxLR 1-14 and Hardham 5 and 6 for the purposes of this study. Expression profiles of the 16 genes were then re-analysed to determine timepoints at which the genes were significantly upregulated. *RxLR 1, 2, 3* and *10* were all highly expressed at 24 hpi, which is similar to the expression profiles of *P. infestans Pi04089* and *P. sojae PsPSR2*, *Avr3b* and *Avr1b-1,* which were expressed during the biotrophic phase of infection (Shan *et al.*, 2004, Xiong *et al.*, 2014, Wang *et al.*, 2015). The expression patterns of *RxLR 1, 2, 3* and *10* thus support their annotation as candidate RxLRs, since they are expressed during the phase of infections when RxLR effectors function to suppress host immune responses. This finding makes sense in terms of our current understanding of infection by *P. cinnamomi*, since the hemibiotrophic pathogen initially grows as a biotroph in the host, where it produced feeding structures to obtain nutrients from host cells (Hardham, 2005, Hardham & Blackman, 2017). During this phase, hemibiotrophs avoid detection by host immune

signalling to prevent the induction of HR and subsequent limitation of pathogen growth (Jones & Dangl, 2006, Hein *et al.*, 2009, Dou & Zhou, 2012, Zeilinger *et al.*, 2016). Thus *P. cinnamomi* would be expected to utilize a subset of its effector arsenal to suppress the initiation of host defense responses.

The *P. cinnamomi* candidates *RxLR 5, 6, 8* and *9* were all significantly upregulated at 5 dpi, which coincided with the necrotrophic growth phase of the pathogen, during which it initiates the destruction of host cells, obtains nutrients and optimizes infection (Dou & Zhou, 2012, Mengiste, 2012, Hardham & Blackman, 2017). Although some *RxLR*s are expressed during later infection stages (Ye *et al.*, 2011), several *RxLR*s in other *Phytophthora* spp. were shown to decrease to basal levels at 4-5 dpi after peaking at 2-3 dpi (Haas *et al.*, 2009, Kelley *et al.*, 2010). The expression profiles of *P. cinnamomi RxLR 5, 6, 8* and *9* may indicate that they are unlikely RxLR effectors, or that they do not play a role in suppression of necrosis during the biotrophic phase of infection. However, it should be noted that Reitmann *et al.* (2017) reported that *RxLR 9* was expressed in the cysts and germinating cysts of *P. cinnamomi,* similar to *P. infestans Avr3a* (Armstrong *et al.*, 2005) and *P. sojae Avh241* (Yu *et al.*, 2012). Thus, expression patterns may not coincide with the expression of characterized effectors. Since the expression data were limited to three time-points, it would provide a limited understanding of the expression expression of these genes during infection of avocado.

Since the ultimate goal was to determine the protein sequences for the candidate effectors, it was imperative that correct DNA sequences were used for downstream analysis. Thus, after qualifying their expression, the 16 candidate genes were sequenced to confirm their presence in GKB4 (the isolate used in this study). In general, the sequences of cloned genes were similar to the original sequences obtained from JGI, with the exception of SNPs that differed between the alleles of GKB4, or between GKB4 and the JGI strain, *P. cinnamomi var cinnamomi* v1.0. For two of the candidates, there were missing base pairs compared to the original sequences, but multiples of three; consequently, their reading frames would not change and the predicted protein would not differ greatly in its amino acid sequence. Sequencing also revealed that both clones for *Hardham 6* (*Hardham 6B* and *C*) had nearly identical sequences except for a region in

Hardham 6C where 54 base pairs were missing. Both clones were analysed further to determine whether they were alleles of the same gene or repeats that represented two separate genes.

The confirmed genomic sequences were then used to determine the prospective loci for each gene using a custom BLAST search. Several of the genes had multiple hits, to different contigs of the *P. cinnamomi* genome. In most cases, the hits differed slightly in their overall similarity scores, leading to the hypothesis that the hits represented duplications of the candidate *RxLR*s, rather than errors in the genome assembly.

When these genomic coordinates were investigated for coverage by RNA-seq reads, several of the proposed duplications were not expressed. Even though they were similar to regions which were expressed, hits to genomic regions without corresponding transcriptome evidence were not used to generate final protein sequences. It is possible that the non-expressed duplications were located in regions of the genome that were not transcriptionally active at the timepoints used in this study or were not expressed during the infection of avocado specifically. Whether the RNA-seq reads were incorrectly mapped in the pseudo-alignment that was originally used to generate read counts could also be investigated, although this was outside the scope of this study.

RxLR 4, 8, 13, 14 and *Hardham 5* were each found at two genomic loci that were both expressed. When proteins were predicted for the expressed genes, the genomic coordinates for *Hardham 5* on contig 53 did not produce a full-length RxLR. While it was clear that the region was transcribed, no start codon was found before the region which shared amino acid identity with the Hardham 5 prediction; thus, this region was considered to be a pseudogene which did not produce a functional protein. The second proposed genomic locus for *RxLR 13* on contig 68 was transcribed to mRNA that did not contain a stop codon – it is therefore likely that the truncated protein generated by this transcript would not be functional. *Hardham 6B* and *Hardham 6C* were shown to be two distinct genes that were both expressed and found in a tandem repeat on the same contig.

The presence of *RxLR 4, RxLR 8, RxLR 14*, and *Hardham 6*, as multiple copies in the *P. cinnamomi* genome reflected their roles as putative effectors. Several *RxLR* effector

genes are present as multiple copies in the genome of *P. sojae,* the copy numbers of which can vary between different strains of the pathogen (Dong *et al.*, 2009, Qutob *et al.*, 2009). It has been hypothesized that the presence of some RxLR effector genes as multiple copies may contribute to the fitness of the pathogen (Qutob *et al.*, 2009). It has also been recognized that *RxLR* effectors of *Phytophthora* species tend to be located in plastic regions of the genomes, which are enriched in repeat regions and mobile genetic elements and evolve faster than the gene-rich regions where housekeeping genes are located (van Damme *et al.*, 2011, Jiang & Tyler, 2012).

The coverage of predicted genomic coordinates for the candidate *RxLR* genes was also used to determine intron/exon boundaries for the genes, to allow accurate prediction of protein sequences in downstream analysis. When the manual annotation was compared to the prediction of coding regions by BRAKER2, only 10 of the 23 genomic loci that were investigated were shown to have accurate predictions. For the remaining genes there were discrepancies in the prediction of coding regions, or in some cases no CDSs were predicted at all for transcripts which were observed to be expressed. When manual annotations and BRAKER2 predictions were compared with the original prediction software used for the candidate genes, no overall trend was evident. No single gene prediction algorithm was accurate for all candidate genes. However, this is probably due insufficient manual annotations that would inform gene prediction algorithms for oomycetes.

Manual annotation indicated that several candidate genes in the present study contained introns, which were taken into account for protein sequence prediction. No overall trend was observed for the presence/absence of introns, or the number of introns the candidate genes obtained. Thus, assumptions should not be made in future studies regarding the structure of oomycete effector genes. Rather, manual annotation should be performed when investigating protein sequences of oomycete effectors.

For a few candidate genes, intron/exon boundaries were unclear when using RNA-seq data. Consequently, all possible transcripts could not be taken into account when predicting the protein sequences for these genes. This limitation would ideally be solved by experimental approaches involving amplification and sequencing of the transcripts

from pathogen mRNA, which would provide the exact CDS for the relevant effectors. Manual annotation of the effector genes using molecular experimental evidence to predict the final peptides is suggested as an objective for future studies.

When manual annotations were used to predict protein sequences for the candidate RxLRs, it was revealed that the new predicted sequence for RxLR 6 did not have a translation start site followed by an RxLR motif in any of the translated reading frames, and so this was considered a false prediction that was not actually an effector gene expressed during avocado infection. The new protein predicted for RxLR 5 was in fact almost identical to RxLR 4 – the *RxLR 5* gene was therefore considered to be another copy of *RxLR 4* in the *P. cinnamomi* genome, although their expression profiles differed during infection, with RxLR 4 being upregulated only in mycelia while RxLR 5 was upregulated at 5 dpi. The opposing expression patterns of the two isoforms of *RxLR 4* are an interesting observation, since expression polymorphisms have been found in the *RxLR*s of other *Phytophthora* species (Qutob *et al.*, 2009). The differing expression patterns of gene isoforms in literature have only been shown for the same gene found in different isolates of a species (Qutob *et al.*, 2009). To our knowledge, there have been no expression polymorphisms observed to date for multiple copies of the same gene within a single *Phytophthora* isolate.

Another possibility is that *RxLR 4* and its isoform *RxLR 5* represent a recent duplication event, in which case the effector gene is in the process of undergoing subfunctionalization or neofunctionalization – where one of the copies is in the process of diverging functionally from its paralog (He & Zhang, 2005). In some cases, a duplication event can result in expression divergence between two copies of a gene if there are differences in their upstream promoter sequences (Zhang, 2003, Hahn, 2009). This type of functional divergence due to differing expression profiles has been observed for multi-copy genes in several plant species, though the expression divergence of these genes refers to the specific tissues in which they are expressed, rather than the timing of their expression (Duarte *et al.*, 2005, Renny-Byfield *et al.*, 2014, Qiao *et al.*, 2018). Although there have not yet been studies into this type of expression divergence for multi-copy genes in

oomycetes, it remains a possibility that *RxLR 4* and *RxLR 5* present paralogs that are in the process of evolving to carry out differing functions during infection by *P. cinnamomi*.

Subsequent to prediction of the final protein sequences, we wanted to investigate how much the new predictions differed from the original sequences obtained from JGI. Overall, the new peptides were 95-100% similar to their original sequences. The exceptions here were RxLR 5, which was more similar to RxLR 4, and Hardham 6B and C, which had showed more similarity to the original Hardham 5 prediction than its own. The low differences observed in general for the new protein sequences when compared to the original predictions means that conclusions drawn from studies using only the original predicted effectors may still be accurate in most cases. Using only the original peptide sequences would, however, possibly result in the accidental exclusion of other proteins present, which might not be taken into account if the DNA sequences of the relevant genes are not first mapped back to the genome.

Protein sequences of RxLR effectors were used in a phylogenetic analysis to investigate whether functional inferences could be made for the RxLRs during infection of avocado. A phylogenetics approach was used since RxLRs have been show to contain very little similarity to other RxLRs (Jiang *et al.*, 2008, Haas *et al.*, 2009), and so a simple BLAST search would not necessarily provide information as to the possible functions of the effectors. It was hypothesised that shared evolutionary history between RxLR effectors might provide clues as to their possible functions *in planta*.

The first phylogenetic tree produced by Bayesian inference contained several outgroups with sequences which were not demonstrated to be related to any other sequences; these included the *P. cinnamomi* candidate effectors RxLR 3, 7, 8, 10 and 13 in the present study. RxLR 7 and 13 may be false predictions or may not have virulence roles in avocado, since they were not related to other *Phytophthora* effectors and were not upregulated during the time-course study. RxLR 3 and 10 were both upregulated during the biotrophic phase of infection as is the norm for most RxLR effectors. These remain solid candidates for RxLR effectors which contribute to *P. cinnamomi* virulence, as is RxLR 8, due to its presence in multiple copies in the genome and its expression at a later stage of infection. It is possible that RxLR 3, 8 and 10 had evolved from other RxLRs

which were not included in the phylogenetic analysis. This is a distinct possibility, since only a fraction of the candidate RxLRs identified in other species have been assigned functions to date, and so it is conceivable that the related effectors in other species have not yet been characterised.

The outgroups from the first phylogenetic analysis were then removed, and a second analysis was performed, with the aim of producing a tree with better resolution within the clades. Sequences which grouped together in the first tree were also subjected to sequence alignments to determine whether there were regions of similarity within their C-terminal regions. When the results of both trees and the alignments were interpreted together, better conclusions could be drawn about possible functions of the candidate *P. cinnamom*i effectors.

Within the phylogenetic tree produced from the first analysis, several *P. cinnamomi* RxLR effectors, including RxLR 1, 4, 5 and 12 were shown to be distantly related to characterised effectors in other species, but these groupings were not consistent in the second tree that was produced. When the full-length proteins were aligned based on the clades in the first tree, the effectors were shown to be highly divergent across their sequences, providing evidence that the original relatedness predicted was not a true reflection of the relatedness between the effectors in question, as shown in the refined tree. Putative functions were thus not assigned for RxLR 1, 4, 5 and 12. These candidates maintained their putative assignations as RxLR effectors, since their expression profiles support their annotation as effector proteins, with *RxLR 1* being upregulated during the biotrophic phase of infection, and *RxLR 5* and *RxLR 12* being expressed at later stages. *RxLR 4* was not upregulated during the infection time-course presented here, but its presence in multiple copies in the genome supports its classification as a likely effector of *P. cinnamomi.* It is possible that the absence of expression data for *RxLR 4* is due to the specific time-points chosen for this experiment, and that RNA-seq data was not obtained for the specific time at which this gene's expression peaked. It is also conceivable that this *RxLR* might be expressed in another plant host, but does not play a specific role during avocado infection.

Putative functions could be assigned to a subset of the candidate effectors based on their consistent grouping with characterised effectors across all phylogenetics analyses. RxLR 14 showed relatedness to *P. parasitica* RxLR 2 in both trees, and an alignment of these protein sequences revealed short regions of consensus between the peptides. It is therefore possible that RxLR 14 may have similar functions to *P. parasitica* RxLR 2, which suppresses PCD in host cells (Dalio *et al.*, 2018). This hypothesis is not supported by the expression profile of RxLR 14 shown in this study, since the gene was not significantly upregulated at any of the infection timepoints, but RxLR 14 (Phyci_30885) has previously been shown to be expressed in pre-infection structures of the pathogen (Reitmann *et al.*, 2017). It is therefore possible that this gene was upregulated at much earlier infection timepoints which were not included in the RNA-seq experiment used for this study.

In both the first and second tree produced by Bayesian inference analysis *P. cinnamomi* RxLR 2 formed a clade with *P. sojae* Avh238, but both groupings had relatively low posterior probability values of less than 0.7. It was hypothesised that these two effectors were more distantly related. Alignment of their full-length protein sequences revealed several short regions of identity between the effectors, further supporting this hypothesis. Based on the results of this phylogenetic analysis, and significant upregulation of *RxLR 2* at 24 hpi, it is highly likely that this effector plays a role in suppression of host immune responses during the biotrophic phase of infection. It would be interesting to investigate whether the host targets for Avh238, Type2GmACSs (*Glycine max* 1-aminocyclopropane-1-carboxylate synthases) (Yang *et al.*, 2017, Yang *et al.*, 2019), have homologs in avocado which are targeted by this RxLR effector during infection. This would provide a novel avenue for future studies into functions of *P. cinnamomi* RxLRs.

Finally, the clade with the highest supporting probability values was the group containing RxLR 9, Hardham 5, Hardham 6 and *P. infestans* PiSNE1. These effectors were shown to be related with posterior probability values of more than 0.95 for both trees generated by phylogenetic analysis. Alignment of the full-length protein sequences revealed that several blocks of conserved residues were present in these effectors, supporting the fact that they likely have similar functions *in planta*. This is supported by observations made in literature, where Reitmann *et al.* (2017) proposed RxLR 9 (Phyci_16230) as a putative

ortholog of PiSNE1, and Hardham & Blackman (2017) predicted Hardham 5 (Phyci_24296) and Hardham 6 (Phyci_297058) as candidate *P. cinnamomi* RxLRs based on their homology to PiSNE1. It can therefore be postulated that Hardham 5, Hardham 6 and RxLR 9 all function to suppress necrosis during the biotrophic phase of host infection(Kelley *et al.*, 2010). This hypothesis is disputed by the expression profiles of the genes in question, since not one of these genes was significantly upregulated during infection of avocado at the time points sequenced in this study. Based on the expression profile of PiSNE1 in tomato, it would be expected that these genes would be highly expressed during the biotrophic phase of infection, with expression peaking at 2-3 dpi (Kelley *et al.*, 2010). It was, however, shown that RxLR 9 was expressed in cysts and germinating cysts of *P. cinnamomi* (Reitmann *et al.*, 2017), and so it is possible that these genes were expressed at earlier timepoints of infection that were not analysed in this study. It is also possible that these genes represent viable RxLR effectors that play a role during infection of other host plants, and are not specifically upregulated in avocado.

Conclusion

Several candidate *RxLR* effector genes have been predicted in *P. cinnamomi*, but it remains unclear how many are actually used by the pathogen during infection of various hosts. To our knowledge, this is the first study to identify candidate *P. cinnamomi RxLRs* that play a role in infection of avocado. We used transcriptome data of a susceptible avocado rootstock infected with *P. cinnamomi* to identify 16 putative RxLRs expressed at some level during infection, which enabled the manual annotation of these genesand accurate prediction of their protein sequences. The final protein sequences of these genes were compared to known effector proteins, and putative functions could be assigned for four of the candidate RxLRs. Expression profiles suggested that four additional effectors play a probable role during avocado infection. Although functional assays for *P. cinnamomi* RxLRs were not performed, several candidate effectors were identified that could help focus future studies to elucidate virulence functions of the pathogen during its infection of avocado.

Acknowledgements

The RNA-seq data used in this study was produced by members of the Avocado Research Programme and sequencing was performed by Novogene (Novogene Corporation Inc., Chula Vista, California, USA). Quality control and read quantification was performed by Dr. Sarah Mwangi. The prediction of *RxLR* genes in the *P. cinnamomi* genome was done by Dr. Tuan Duong. The new *P. cinnamomi* genome used in this study was produced by Ms. Juanita Engelbrecht. This study was supported by funding provided by the Hans Merensky Foundation and the National Research Foundation.

Tables

Table 1. List of functionally characterized RxLR effectors in oomycete species as used in this study.

Protein name	Species	Genbank accession number	Reference
PiAVRblb1	*Phytophthora infestans*	UniProt D0P3S7_PHYIT*	(Senchou *et al.*, 2004, Bouwmeester *et al.*, 2011, Chen *et al.*, 2012)
PiAvr1	*P. infestans*	ATG23596.1	(Du *et al.*, 2015)
PiAVR3a	*P. infestans*	AEH27535.1	(Bos *et al.*, 2010, Gilroy *et al.*, 2011b)
Pi04089	*P. infestans*	UniProt D0N0I9_PHYIT*	(Wang *et al.*, 2015)
PiSNE1	*P. infestans*	ABI74673.1	(Kelley *et al.*, 2010)
PiAVRblb2	*P. infestans*	EEY54134.1	(Bozkurt *et al.*, 2011)
PiPexRD2	*P. infestans*	EEY62542.1	(King *et al.*, 2014)
Pi03192	*P. infestans*	EEY65678.1	(McLellan *et al.*, 2013)
Pi04314	*P. infestans*	EEY67311.1	(Boevink *et al.*, 2016)
PiPSR2	*P. infestans*	EEY63424.1	(Xiong *et al.*, 2014, de Vries *et al.*, 2017)
PexRD54	*P. infestans*	XP_002903599.1	(Dagdas *et al.*, 2016)
PiAvr4	*P. infestans*	ABV66276.1	(van Poppel *et al.*, 2008)
PiAvr-vnt1	*P. infestans*	D0NTY1.1	(Stefańczyk *et al.*, 2017)
PiSFI1	*P. infestans*	D0N0J7.1	(Zheng *et al.*, 2014)
PiSFI2	*P. infestans*	D0N0N6.1	(Zheng *et al*., 2014)
PiSFI3	*P. infestans*	D0N6D2.1	(Zheng *et al.*, 2014, He *et al.*, 2019)

PiSFI4	*P. infestans*	D0NCC1.1	(Zheng *et al*., 2014)
PiSFI5	*P. infestans*	D0NMF6.1	(Zheng *et al.*, 2014, Zheng *et al.*, 2018)
PiSFI6	*P. infestans*	D0NN72.1	(Zheng *et al*., 2014)
PiSFI7	*P. infestans*	D0NXM3.1	(Zheng *et al*., 2014)
PiSFI8	*P. infestans*	D0P1B2.1	(Zheng *et al*., 2014)
Pi02860	*P. infestans*	D0MXE4.1	(Yang *et al.*, 2016)
Pi22926	*P. infestans*	EEY57148.1	(Ren *et al.*, 2019)
Pi14054	*P. infestans*	D0NNI8.1	(Vetukuri *et al.*, 2017)
Pi17316	*P. infestans*	D0NVS7.1	(Murphy *et al.*, 2018)
PsPSR1	*Phytophthora sojae*	UniProt E0W5Q5_PHYSO*	(Qiao *et al.*, 2013, Qiao *et al.*, 2015)
PsPSR2	*P. sojae*	AEK80747.1	(Qiao *et al.*, 2013, Xiong *et al.*, 2014)
PsAvr1k (PsAvh331)	*P. sojae*	AGC95073.1	(Cheng *et al.*, 2012)
PsAvr1b-1	*P. sojae*	AAM20937.1	(Dou *et al.*, 2008a, Dou *et al.*, 2008b)
PsAvr3b (PsAvh307)	*P. sojae*	AEI75279.1	(Dong *et al.*, 2011b)
PsAvr3c	*P. sojae*	ACN59479.1	(Huang *et al.*, 2017)
PsAvh262	*P. sojae*	AEK81064.1	(Jing *et al.*, 2016)
PsAvh241	*P. sojae*	AEK81009.1	(Yu *et al.*, 2012)
PsAvr1d	*P. sojae*	G4ZLE6.1	(Na *et al.*, 2013, Yin *et al.*, 2013)
PsAvr4/6	*P. sojae*	ADI72736.1	(Dou *et al.*, 2010, Fang & Tyler, 2016)
PsAvr3a/5	*P. sojae*	AEA51000.1	(Qutob *et al.*, 2009, Dong *et al.*, 2011a)
PsAvr1a	*P. sojae*	ABO47652.1	(Qutob *et al.*, 2009, Na *et al.*, 2014)

PsAvh23	*P. sojae*	G4ZJX4.1	(Kong *et al.*, 2017)
PsAvh238	*P. sojae*	AEK81001.1	(Yang *et al.*, 2019)
PsAvh240	*P. sojae*	G5A8M1.1	(Guo *et al.*, 2019)
PcapRxLR48	*Phytophthora capsici*	N/A (Sequence obtained from Supplementary information from Li *et al.,* 2019a)	(Li *et al.*, 2019a)
PcapAvr3a12	*P. capsici*	jgi|Phyca11|114071**	(Fan *et al.*, 2018)
PcapRxLR207	*P. capsici*	A0A411NN20.1	(Li *et al.*, 2019b)
PcapAvh1	*P. capsici*	MF975713	(Chen *et al.*, 2019)
PpPSE1	*Phytophthora parasitica*	P0CV74.1	(Evangelisti *et al.*, 2013)
PpRxLR2	*P. parasitica*	W2MBG9.1	(Dalio *et al.*, 2018)
PpE4	*P. parasitica*	W2RE78.1	(Huang *et al.*, 2019)
HaATR1	*Hyaloperonospora arabidopsis*	M4B6G6.1	(Chou *et al.*, 2011)
HaATR5	*H. arabidopsis*	M4C699.1	(Bailey *et al.*, 2011)
HaATR13	*H. arabidopsis*	Q5G7K8.1	(Leonelli *et al.*, 2011)
HaATR39-1	*H. arabidopsis*	H9BPR8.1	(Goritschnig *et al.*, 2012)
PIAP_11895	*Pilasporangium apinafurcum*	N/A (Sequence obtained from Supplementary information from McGowan & Fitzpatrick, 2017)	(McGowan & Fitzpatrick, 2017)

*Peptide sequences that were not available on NCBI Genbank were obtained from the UniProtKB database (https://www.uniprot.org)

**Sequence obtained from the Joint Genome Institute (JGI)

Table 2. Sequences of primers designed for amplification of candidate RxLR genes from *Phytophthora cinnamomi* DNA.

Sequence ID	Primer ID	Primer sequence 5'-3'	No. of bp
RxLR 1 - 325329	RxLR1-F1	CAGCCCTCTTCCTCATTCAC	20
	RxLR1-R1	CTGTATTGCTTCACACCAGG	20
RxLR 2 – 100005	RxLR2-F1	CAGACGTAAGAATGGTTGGC	20
	RxLR2-R2	ACTATCACCAACCGCTCCAG	20
RxLR 3 – 220952	RxLR3-F1	GGTGCTTGTAGTACCCAGTTCC	22
	RxLR3-RM	CGCTAGTAACGGGGAAAAGC	20
RxLR 4 – 97174	RxLR4-F1	TGCTCTTCACTTGGCTTCAC	20
	RxLR4-R1	CGTAGCAGCACGAATGACAC	20
	RxLR4-R3	CGTCCACTGCATGCTTCAC	19
RxLR 5 – 81261	RxLR5-F2	CGTGGTACACTGTACCTGAAAGC	23
	RxLR5-R1	TCAGCTTGCTTTCGTCCTTCAG	22
RxLR 6 – 22058	RxLR6-F1	GGTGGCACTGTATCCCGAAG	20
	RxLR6-R1	TGCAACGCGAATTGCTACTG	20
	RxLR6-R4	GTTTGGCCTCAGCCACTCTC	20
RxLR 7 – 313963	RxLR7-F1	CAAGTCAAACCGGCCTTTCG	20
	RxLR7-R1	CGCTGCTGGACACTACCAAC	20
	RxLR7-R2	GCCGAGATGGTGGTCAAGTACG	22
RxLR 8 - 86482	RxLR8-F1	TCATGAGGCGTCGCAGAAGC	20
	RxLR8-R1	TCCCTCCTGCCTTCCTTCTG	20
RxLR 9 - 16230	RxLR-9-F	GAGAGCTCACTCTCCAATTCC	21

	RxLR-9-R1	GGCTTGATAACTGTACGCACG	21
RxLR 10 - 31697	RxLR-10-F	TCCAGAACGGCAACGCTAAG	20
	RxLR-10-R	ACGGCGCAAACACTGGTAAG	20
	R-10-In-F	CGGGTTACCACACTGAAGAG	20
	R-10-In-R	GCATCGCCCACTTCTCAAAC	20
RxLR 11 - 80229	RxLR-11-F	TCCGGGTATGGACACTTCAC	20
	RxLR-11-R	CATGCATCCAAGGCGTTGTG	20
RxLR 12 - 90101	RxLR-12-F	GCGCTCACTTCGGCTTTGTC	20
	RxLR-12-R	GTGACTCCCTCATCGTGAAC	20
RxLR 13 - 88750	RxLR-13-F	AGACTGGCATCGCGTCACTC	20
	RxLR-13-R	CGACAACCGTATCCGACTCTC	21
RxLR 14 - 30885	RxLR-14-F	GAACCTGTGGCGACAAACTG	20
	RxLR-14-R	CTTCAGAACAGCACGATCAG	20
HARDHAM 5 - 24296	Hardham-5-F	ATTTGCCGAGCCAAGCATAC	20
	Hardham-5-R	CTAGAAGGAAGTGCGGCATC	20
HARDHAM 6 - 297058	Hardham-6-F	CCAACTCAAGCACCATGAAG	20
	Hardham-6-R	TGCGGTATCGCCAGCATTC	19

Table 3. Number of raw reads obtained from RNA-sequencing of a susceptible avocado infected with *Phytophthora cinnamomi* across all timepoints sampled, and percentage of reads mapping to the draft genome of *P. cinnamomi.*

Rootstock type	**Timepoint**	**Library name**	**Total number of paired reads**	**Alignment to *P. cinnamomi* draft genome**	**Total number of reads aligned to genome**
N/A	*P. cinnamomi* Mycelia	MS13	95,071,217	86.02%	0
		MS14	51,388,341	83.48%	0
		MS15	52,439,583	82.71%	0
Susceptible (R0.12)	12 hpi*	MS4	42,715,042	0.06%	93873.3884
		MS5	44,667,268	0.23%	232895.81
		MS6	58,180,645	0.17%	89704.1398
Susceptible (R0.12)	24 hpi*	MS7	42,669,722	0.22%	1187008.495
		MS8	93,158,324	0.25%	2272025.514
		MS9	40,774,609	0.22%	989554.5764
Susceptible (R0.12)	5 dpi**	MS10	44,457,247	2.67%	81780260.86
		MS11	103,273,887	2.20%	42898987.07
		MS12	69,686,942	1.42%	43372779.1
Partially Resistant (Dusa®)	5 dpi**	MS19	44,413,579	2.03%	901595.6537
		MS20	52,611,639	3.54%	1862452.021
		MS21	40,473,940	2.89%	1169696.866

*Hours post-inoculation

**Days post-inoculation

Table 4. Read counts for individual *Phytophthora cinnamomi* transcripts, quantified and normalized as transcripts per million (TPM) in Kallisto.

Timepoint	**Mycelia**			**12 hours post-inoculation**			**24 hours post-inoculation**			**5 days post-inoculation**		
Replicate number	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
Library name	**(MS13)**	**(MS14)**	**(MS15)**	**(MS4)**	**(MS5)**	**(MS6)**	**(MS7)**	**(MS8)**	**(MS9)**	**(MS10)**	**(MS11)**	**(MS12)**
Phyci97951 **read count (TPM)**	0.83871	1.58018	1.47988	0	0	0	0	3.4939	0	4.07451	0.61869	0.61869
Phyci227271 **read count (TPM)**	6.36465	6.54077	4.70631	0	0	0.37545	0	4.09408	0	4.37224	1.62461	1.62461
Phyci16230 (RxLR 9) **read count (TPM)**	0.80074	0.61153	0.64944	0	0	0	0	24.0732	0	53.8426	10.7293	10.7293
Phyci65337 **read count (TPM)**	3.32039	5.88689	14.0264	0	9.38346	0.94898	11.3826	21.4129	6.65911	1.78684	9.68699	9.68699
Phyci31697 (RxLR 10) **read count (TPM)**	6.57068	4.41139	5.51522	0	0	0	12.1565	3.59326	16.5227	0.62922	2.28461	2.28461
Phyci80229 (RxLR 11) **read count (TPM)**	1.33365	1.41487	0.57617	0	0	0	0	0	0	27.6052	88.5679	88.5679
Phyci646 **read count (TPM)**	1.7051	2.0594	5.87609	0	1.89595	0.71010	1.08055	0	0	14.8633	10.946	10.946

Phyci90101 (RxLR 12) **read count (TPM)**	26.0272	20.2698	7.10281	0	1.44908	0	3.27948	0	1.55438	127.108	21.3038	21.3038
Phyci76605 **read count (TPM)**	52.6375	46.1916	41.6374	0	40.2465	7.4727	5.81989	0	5.40055	8.61356	0	0
Phyci206852 **read count (TPM)**	19.0639	16.8399	12.2928	0	6.97667	6.46435	0	0.85073	2.7476	13.5834	8.07678	8.07678
Phyci325329 (RxLR 1) **read count (TPM)**	1.91186	0.56701	2.61875	7.75453	14.027	2.66331	197.524	506.206	323.802	463.176	123.324	123.324
Phyci100005 (RxLR 2) **read count (TPM)**	3.25335	2.12212	8.79204	12.432	18.3431	0	119.763	601.683	446.481	45.6139	5.43907	5.43907
Phyci220952 (RxLR 3) **read count (TPM)**	0	0	0	18.9338	0	0	0	145.735	169.031	112.978	16.6666	16.6666
Phyci99179 **read count (TPM)**	3.86647	3.73066	5.02759	0	0.83149	0	0	5.62003	0.89523	8.8859	3.94403	3.94403
Phyci70517 **read count (TPM)**	2.39449	2.32009	2.92197	0	5.35452	0.25067	0.37613	0.64756	0.71190	3.86732	5.92699	5.92699
Phyci90847 **read count (TPM)**	2.90875	2.82112	3.27094	0	0.20675	1.4809	0	1.96333	1.7825	0	1.68723	1.68723
Phyci22723 **read count (TPM)**	1.07762	3.426	2.39426	1.15288	11.811	0	13.8359	0	0	3.72766	4.61955	4.61955

Phyci96120 **read count (TPM)**	12.8538	15.2756	14.9768	1.16316	8.64586	3.26725	1.70489	1.47919	1.58868	19.558	11.4335	11.4335
Phyci313963 (RxLR 7) **read count (TPM)**	24.3887	35.1133	31.0439	2.32326	7.78439	5.19284	6.71284	9.80821	2.08338	7.74444	3.30987	3.30987
Phyci88750 (RxLR 13) **read count (TPM)**	6.57928	7.62009	8.11672	10.2832	0	5.52521	14.3357	2.46792	0	5.34987	8.3213	8.3213
Phyci30885 (RxLR 14) **read count (TPM)**	2.16611	0.47367	6.30595	13.1167	37.3857	0	0	0	13.5678	1.51215	19.7124	19.7124
Phyci97174 (RxLR 4) **read count (TPM)**	75.3153	104.927	73.2703	4.63239	8.10542	16.7994	6.29615	8.1117	17.3302	24.1484	9.88004	9.88004
Phyci63598 **read count (TPM)**	78.0836	108.533	114.652	1.63233	14.0975	22.3909	19.9862	46.1885	28.0794	37.5325	42.5574	42.5574
Phyci81261 (RXLR 5) **read count (TPM)**	37.4082	40.3247	32.5759	0	17.2793	11.9032	3.29542	8.45358	6.09332	23.0607	20.9731	20.9731
Phyci22058 (RxLR 6) **read count (TPM)**	28.2161	43.108	49.5602	0	3.96762	4.9362	5.29598	21.9227	21.0144	38.2758	10.4064	10.4064
Phyci86482 (RxLR 8) **read count (TPM)**	458.882	411.559	385.888	163.933	734.241	96.5231	334.795	186.15	80.5634	965.641	1020.78	1020.78
Phyci87160 **read count (TPM)**	0.08522	0.19355	0	0	0	0	0	0	0	0	0	0

Phyci568683 **read count (TPM)**	0	0.12359	0	0	0	0	0	0.77768	0	3.04073	0	0
Phyci129006 **read count (TPM)**	4.59456	3.98152	3.72969	0	0	0	0	0.74694	0	0.54053	0.20932	0.20932
Phyci324245 **read count (TPM)**	0.63172	1.06909	1.83823	0	0.59862	0	0.69035	0	0	0	2.69862	2.69862
Phyci24296 (Hardham 5) **read count (TPM)**	0	0	0.29814	4.4114	0	0	0	46.9813	0	112.464	2.53855	2.53855
Phyci297058 (Hardham 6) **read count (TPM)**	1.10061	1.70641	12.9686	0	5.35795	0	95.0419	172.635	59.9744	214.984	30.5419	30.5419

Table 5. Characteristics of predicted protein sequences for the selected RxLRs from *Phytophthora cinnamomi* and position of relevant motifs.

Gene name	Peptide length (number of residues)	Signal Peptide	Signal peptide cleavage site	RxLR motif	RxLR motif position	dEER motif	dEER motif position
RXLR 1 - Phyci325329	150 aa	Yes	24	RNLR	33-36	None	-
RXLR 2 – Phyci100005	134 aa	Yes	23	RFLR	49-52	DEER	65-68
RXLR 3 – Phyci220952	153 aa	Yes	23	RRLR	39-42	EKDDDADEER	70-79
RXLR 4 – Phyci97174	162 aa	Yes	24	RRLR	34-37	None	-
Phyci63598	426 aa	No	-	RDLR	39-42	None	-
RXLR 5 – Phyci81261	337 aa	Yes	20	RGLR	49-52	None	-
RXLR 6 – Phyci22058	637 aa	Yes	21	RMLR	53-56	DERNEER	60-66
Phyci65337	692 aa	No	-	RVLR	31-34	None	-
RXLR 7 – Phyci313963	451 aa	Yes	24	RRLR	40-43	DIER	64-67
RXLR 8 – Phyci86482	89 aa	Yes	23	RQLR	31-34	None	-
Phyci97951	1208 aa	Yes	24	RGLR	31-34	None	-
Phyci227271	401 aa	Yes	27	RYLR	40-43	None	-
RXLR 9 - Phyci16230	158 aa	Yes	23	RSLR	46-49	None	-
RXLR 10 – Phyci31697	397 aa	Yes	22	RLLR	46-49	DANDEER	54-60
RXLR 11 – Phyci80229	109 aa	Yes	23	RNLR	48-51	None	-

Phyci646	462 aa	No	-	RTLR	42-45	None	-
RXLR 12 – Phyci90101	205 aa	Yes	28	RSLR	45-48	None	-
Phyci99179	489 aa	Yes	21	RHLR	54-57	None	-
Phyci70517	1155 aa	Yes	23	RLLR	55-58	EDDVNDDDSSADEEER	64-79
Phyci90847	620 aa	No	-	RALR	35-38	None	-
Phyci22723	277 aa	Yes	23	RFLR	45-48	DDDTEDLKSAEEER	57-70
RXLR 13 – Phyci88750	172 aa	Yes	23	RLLR	46-49	DEDEER	55-60
RXLR 14 - Phyci30885	160 aa	Yes	21	RMLR	56-59	ENNEER	71-76
HARDHAM 5 - Phyci24296	183 aa	Yes	23	RNLA RQLG GGLR		None	-
HARDHAM 6 - Phyci297058	262 aa	Yes	23	RNLA RQLG GGLR		None	-

Table 6. Results of Protein BLAST searches for similarity of predicted *Phytophthora cinnamomi* RxLR protein sequences to known proteins in other species.

Gene name	Top pBLAST Hit	Organism	% similarity	% coverage	E-value	Accession number
RXLR 1 -Phyci325329	RxLR effector protein	*Phytophthora megakarya*	54%	26%	0.013	OWZ06359.1
RXLR 2 – Phyci100005	Hypothetical protein PPTG_00121	*Phytophthora parasitica* INRA-310	52%	97%	2e-34	XP_008889734.1
RXLR 3 – Phyci220952	Hypothetical protein PHYSODRAFT_287036	*Phytophthora sojae*	50%	96%	2e-33	XP_009533161.1
RXLR 4 – Phyci97174	Hypothetical protein PHYSODRAFT_551054	*Phytophthora sojae*	84%	93%	3e-83	XP_009538323.1
Phyci63598	Phosphoribosyldiphosphate synthetase	*Phytophthora sojae*	98%	100%	0.0	XP_009518996.1
RXLR 5 – Phyci81261	Hypothetical protein PHYSODRAFT_361266	*Phytophthora sojae*	95%	91%	0.0	XP_009531151.1
RXLR 6 – Phyci22058	Hypothetical protein PHYSODRAFT_547128	*Phytophthora sojae*	93%	48%	0.0	XP_009531015.1
Phyci65337	Hypothetical protein PHYSODRAFT_547952	*Phytophthora sojae*	91%	95%	0.0	XP_009532454.1
RXLR 7 – Phyci313963	Hypothetical protein PHYSODRAFT_342126	*Phytophthora sojae*	91%	100%	0.0	XP_009538787.1

RXLR 8 – Phyci86482	Hypothetical protein PPTG_06368	*Phytophthora parasitica* INRA-310	78%	100%	2e-41	XP_008898521.1
Phyci97951	Subtilisin serine protease	*Phytophthora sojae*	88%	100%	0.0	XP_009518600.1
Phyci227271	Hypothetical protein PHYSODRAFT_325639	*Phytophthora sojae*	75%	96%	2e-177	XP_009519816.1
RXLR 9 - Phyci16230	Hypothetical protein PHYSODRAFT_279137	*Phytophthora sojae*	87%	43%	3e-17	XP_009520776.1
RXLR 10 – Phyci31697	Hypothetical protein PHYSODRAFT_332765	*Phytophthora sojae*	94%	39%	2e-100	XP_009528124.1
RXLR 11 – Phyci80229	Hypothetical protein PHYSODRAFT_259375	*Phytophthora sojae*	63%	100%	7e-27	XP_009523525.1
Phyci646	Hypothetical protein F443_11359	*Phytophthora parasitica* P1569	86.67%	94%	0.0	ETI43745.1
RXLR 12 – Phyci90101	Hypothetical protein PPTG_22495	*Phytophthora parasitica* INRA-310	71%	100%	3e-120	XP_008902595.1
Phyci99179	Hypothetical protein F442_10908	*Phytophthora parasitica* P10297	79%	94%	0.0	ETP42156.1
Phyci70517	Hypothetical protein PHYSODRAFT_469522	*Phytophthora sojae*	88%	81%	0.0	XP_009516952.1
Phyci90847	Hypothetical protein PHYSODRAFT_486668	*Phytophthora sojae*	92.15%	91%	0.0	XP_009520900.1
Phyci22723	Avr1b-1 avirulence-like protein	*Phytophthora sojae*	60%	100%	5e-107	XP_009522698.1
RXLR 13 – Phyci88750	Hypothetical protein PHYSODRAFT_286360	*Phytophthora sojae*	55%	89%	3e-45	XP_009529451.1

RXLR 14 - Phyci30885	Avr1b-1 avirulence-like protein	*Phytophthora sojae*	46%	100%	8e-21	XP_009516309.1
HARDHAM 5 - Phyci24296	Hypothetical protein PC110_g2418	*Phytophthora cactorum*	54.59%	98%	2e-33	RAW41426.1
HARDHAM 6 - Phyci297058	Hypothetical protein PC110_g7140	*Phytophthora cactorum*	71.43%	30%	7e-32	RAW36597.1

Table 7. Comparative results of gene prediction according to three different prediction programmes, showing the relative nucleotide location for each coding region of the *RxLR* genes.

	Joint Genome Institute			**FGENESH**				**EumicrobeDB**	
Gene name	**Gene Length**	**CDS***	**Number of Exons**	**TSS****	**CDS***	**Number of Exons**	**Poly- A site**	**CDS***	**Number of Exons**
RXLR 1 - Phyci325329	660 bp	1 - 327 463 - 588	2	-121	1 - 199 314 - 588	2	658	1 - 588	1
RXLR 2 – Phyci100005	451 bp	47 - 451	1	Not predicted	-89 – (-36) 78 - 451	2	546	-89 – (-36) 78 - 451	2
RXLR 3 – Phyci220952	462 bp	1 - 462	1	-443	1 – 242 316-322	2	395	1 - 462	1
RXLR 4 – Phyci97174	661 bp	35 – 535	1	Not predicted	35 – 490 583 - 732	2	850	Gene 1: -200 – (-104) Gene 2: 35 – 490 583-732	Gene 1: 1 Gene 2: 2
Phyci63598	1359 bp	1 – 503	2	-169	1 – 503	2	1373	1 – 503	2

		585 - 1359			585 - 1359			585 - 1359	
RXLR 5 – Phyci81261	1014 bp	1 - 814	1	-141	1 - 814	1	1189	1 - 814	1
RXLR 6 – Phyci22058	2714 bp	1 – 956 1757 - 2714	2	1115	1776 - 2714	1	2752	1 – 956 1775 - 2714	2
Phyci65337	2079 bp	1 - 2079	1	-388	1 - 2079	1	2148	1 - 2079	1
RXLR 7 – Phyci313963	1575 bp	24 – 1004 1075 – 1323 1400 - 1525	3	-106	24 – 1004 1075 – 1261 1322 - 1323 1400 - 1525	4	1633	24 – 1004 1075 – 1323 1400 - 1525	3
RXLR 8 – Phyci86482	270 bp	1 - 270	1	-178	1 - 270	1	433	1 - 270	1
Phyci97951	3830 bp	63 - 3689	1	-166	63 - 3689	1	4342	81 - 3689	1
Phyci227271	1206 bp	1 - 1206	1	No predictions possible				No predictions possible	
RXLR 9 - Phyci16230	477 bp	1 - 477	1	-344	1 - 477	1	559	1 - 477	1
RXLR 10 – Phyci31697	1742 bp	23 – 718 1010 - 1507	2	-272	23 – 152 294 – 718	3	1712	23 – 718 1034 - 1507	2

					1010 - 1507				
RXLR 11 – Phyci80229	330 bp	1 - 330	1	Not predicted	-995 - (-537)	1	-167	-999 – (-537)	1
Phyci646	1579 bp	1 – 226 336 – 1026 1108 - 1579	3	-161	1 – 226 336 – 1026 1108 - 1579	3	1603	1 – 226 336 – 1026 1108 - 1579	3
RXLR 12 – Phyci90101	977 bp	57 - 824	1	-327	57 – 716 825 - 842	2	2509	55 - 824	1
Phyci99179	1580 bp	42 - 1511	1	-94	42 - 1511	1	1594	42 - 1511	1
Phyci70517	4085 bp	1 – 502 942 – 1019 1139 – 1576 1645 - 4085	4	-240	1 – 276 367 – 470 996 – 1019 1501 – 1576 1746 - 4085	5	Not predicted	<u>Gene 1:</u> 1 – 519 <u>Gene 2:</u> 1075 – 1576 1645 - 4085	<u>Gene 1:</u> 1 <u>Gene 2:</u> 2
Phyci90847	2131 bp	53 – 1419 1489 – 1756 1840 - 2067	3	-135	53 – 1419 1489 – 1756 1840 - 1968	3	Not predicted	53 – 1419 1489 – 1756 1840 - 2067	3

Phyci22723	834 bp	1 - 834	1	-190	1 – 755 895 - 904	2	955	1 - 834	1
RXLR 13 – Phyci88750	667 bp	133 - 651	1	56	133 - 651	1	856	133 - 651	1
RXLR 14 - Phyci30885	483 bp	1 - 483	1	-282	1 - 483	1	519	1 - 483	1
HARDHAM 5 - Phyci24296	673 bp	1 – 472 494 - 573	2	-347	1 - 573	1	722	1 - 573	1
HARDHAM 6 - Phyci297058	937 bp	1 – 693 727 - 822	2	Not predicted – no upstream sequence available	1 – 818 957 – 968 1009 - 1018	3	1036	1 - 822	1

*Coding sequence

**Transcription start site

Table 8. Optimised PCR conditions for each of the respective candidate *RxLR* genes from *Phytophthora cinnamomi*.

Gene name	Primer annealing temperature (°C)	Reaction extension time (seconds)	Expected product size (base pairs)
RxLR 1	65	40	739
RxLR 2	65	30	557
RxLR 3	65	30	611
RxLR 5	69	60	1161
RxLR 8	68	20	414
RxLR 9	64	30	623
RxLR 11	65	30	517
RxLR 12	64	40	990
RxLR 13	64	30	673
RxLR 14	64	40	736
Hardham 5	64	40	722
Hardham 6	67	30	887

Table 9. Concentration of DNA extracted from PCR amplification of each of the candidate *RxLR*s from *Phytophthora cinnamomi*.

Extracted PCR Product	Replicate number	DNA concentration (ng/µl)
RxLR 1	1	4.0
	2	5.1
RxLR 2	1	8.4
	2	10.8
	3	13.7
RxLR 3	1	7.5
	2	7.7
RxLR 5	1	4.0
	2	4.4
RxLR 8	1	4.1
	2	5.9
RxLR 9	1	12.3
	2	8.1
RxLR 11	1	6.5
	2	9.4
RxLR 12	1	13.9
	2	14.0
RxLR 13	1	11.3
	2	10.6
	3	17.5
RxLR 14	1	11.8
	2	12.1
	3	21.3

Hardham 5	1	3.3
	2	34.2
Hardham 6B	1	3.5
	2	2.7
Hardham 6C	1	8.5
	2	4.1

Table 10. Concentration of plasmid DNA extracted from transformed *Escherichia coli* for each successfully cloned candidate *RxLR*.

Extracted PCR Product	**Clone from which plasmids extracted**	**Elution**	**Plasmid DNA concentration (ng/µl)**
RxLR 1	1.4	1	125.0
		2	19.3
	1.5	1	250.8
		2	29.3
	1.8	1	248.2
		2	82.9
RxLR 2	2.2	1	359.4
		2	50.6
	2.6	1	355.0
		2	89.3
	2.7	1	385.3
		2	36.5
RxLR 3	3.6	1	414.8
		2	37.1
	3.11	1	281.9
		2	28.6
	3.12	1	238.0
		2	22.0
RxLR 5	5.1	1	260.8
		2	44.9
	5.3	1	357.8
		2	89.7
	5.6	1	302.9
		2	88.4

RxLR 8	8.1	1	213.6
		2	34.0
	8.3	1	234.4
		2	45.3
	8.5	1	336.0
		2	55.7
RxLR 9	9.1	1	320.3
		2	38.3
	9.4	1	558.5
		2	95.2
	9.8	1	462.2
		2	65.5
RxLR 12	12.3	1	421.4
		2	140.4
	12.6	1	441.0
		2	61.3
	12.11	1	44.8
		2	125.7
RxLR 13	13.6	1	212.3
		2	26.6
	13.7	1	109.3
		2	43.6
	13.10	1	289.1
		2	51.9
Hardham 5	H5.6	1	49.8
		2	23.7
	H5.10	1	56.0
		2	15.4
	H5.12	1	138.2
		2	42.6

Hardham 6B	H6B.4	1	99.7
		2	10.8
	H6B.5	1	228.3
		2	45.5
	H6B.9	1	168.2
		2	40.6
Hardham 6C	H6C.2	1	196.5
		2	20.5
	H6C.4	1	87.1
		2	14.3
	H6C.7	1	103.3
		2	33.8

Table 11a. Results of custom BLAST hits to the PcGenome database created in Geneious v7.06, shown for each of the chosen candidate *Phytophthora cinnamomi RxLRs.*

Gene name	Custom BLAST contig hit	Grade score for hit*	E-value	Hit start site	Hit end site
RxLR 1	Contig 14	98.7	0	64,623	63,882
	Contig 32	96.3	0	253,587	252,872
RxLR 2	Contig 123	99.9	0	55,961	56,517
	Contig 38	90.1	9.80E-159	425,046	425,567
RxLR 3	Contig 24	97.4	0	314,547	315,148
	Contig 66	100	0	228,497	227,896
	Contig 37	92.4	3.74E-173	1,278,796	1,278,210
	Contig 76	91.8	2.89E-174	53,239	52,994
RxLR 4	Contig 28	99.4	0	15,715	14,660
	Contig 1	99.4	0	1,334,417	1,333,358
RxLR 5	Contig 63	98.6	0	548,759	547,601
RxLR 6	Contig 47	99.6	0	160,828	163,928
	Contig 1	65.3	0	1,353,067	1,351,792
	Contig 27	69	0	36,200	37,651
	Contig 28	65.7	0	34,306	33,019
RxLR 7	Contig 14	98	0	2,369,911	2,371,821
RxLR 8	Contig 50	99.6	0	343,937	344,361
RxLR 9	Contig 50	99.9	0	148,395	149,026

RxLR 10	Contig 18	87.9	0	477,106	475,157
	Contig 51	75.8	0	1,472	10
	Contig 51	70	0	29,562	28,662
RxLR 11	Contig 67	99.9	0	461,318	461,813
RxLR 12	Contig 62	100	0	731,789	730,800
RxLR 13	Contig 120	99.9	0	96,240	96,912
	Contig 68	76.8	0	59,713	59,351
RxLR 14	Contig 35	99.9	0	679,340	678,624
	Contig 12	99.9	0	16,090	16,805
Hardham 5	Contig 1	98	0	884,907	885,583
	Contig 53	97.7	0	839,680	839,008
Hardham 6B	Contig 50	96.3	0	197,336	196,591
	Contig 50	78	0	200,222	199,773
Hardham 6C	Contig 50	98.3	0	197,336	196,591
	Contig 50	86.1	4.82E-173	200,659	199,945
	Contig 50	73.1	1.04E-169	200,222	199,773

* The grade provided by Custom BLAST is the weighted score for the hit based on the pairwise identity, coverage and E-value. Pairwise identity and query coverage are not shown in this table.

Note: Custom BLAST hits shaded in grey were later rejected as the true genomic coordinates for the candidate RxLRs based on observations in IGV.

Table 11b. Results of custom BLAST hits to the PcGenomeCDS database created in Geneious v7.06, shown for each of the chosen candidate *Phytophthora cinnamomi RxLRs.*

Gene name	PcGenomeCDS Hit ID	Grade score for hit* (%)	Pairwise identity (%)	Query coverage (%)	E-value
RxLR 1	jg9436.t1	86.2	92.0	80.46	1.63E-128
RxLR 2	jg19254.t1	81.9	99.2	64.63	5.49E-77
	jg6034.t1	66.4	70.8	61.94	7.27E-46
RxLR 3	jg191.t1	70.8	70.9	70.76	2.63E-55
RxLR 4	jg17626.t1	68.6	99.3	37.89	1.04E-89
	jg10431.t1	68.6	99.3	37.89	1.04E-89
RxLR 5	jg13020.t1	88.1	95.5	80.76	0
RxLR 6	jg7968.t1	60.7	91.2	30.25	0
	jg8301.t1	53.9	77.7	30.06	4.65E-165
	jg10438.t1	53.7	77.4	30.06	2.24E-161
RxLR 7	jg9835.t1	70.3	98.9	41.77	0
RxLR 8	jg3373.t1	64.8	100	46.81	4.90E-41
	jg3405.t1	64.8	100	46.81	4.90E-41
RxLR 9	jg4237.t1	22.3	69.4	25.16	9.23E-15
RxLR 10	jg3714.t1	56.9	75.3	38.52	1.01E-142
	jg15182.t1	51.9	65.5	38.38	2.31E-110
RxLR 11	No hits to CDS				
RxLR 12	jg15774.t1	84.2	100	68.48	2.01E-164

RxLR 13	jg16635.t1	78.6	89.9	67.31	2.16E-89
RxLR 14	jg1103.t1	79.6	98.7	60.67	2.03E-102
	jg17846.t1	69.6	78.6	60.67	6.13E-74
Hardham 5	jg4255.t1	27.4	95.5	9.42	7.37E-04
Hardham 6B	jg4255.t1	52.8	73.4	32.25	8.49E-25
	jg4255.t1	29.4	66.1	34.13	8.28E-21
Hardham 6C	jg4255.t1	61.3	91.3	31.29	1.21E-31
	jg4255.t1	57.1	84.9	29.34	4.06E-31

* The grade provided by Custom BLAST is the weighted score for the hit based on the pairwise identity, coverage and E-value. Pairwise identity and query coverage are not shown in this table.

Table 12. Final protein sequences inferred for each candidate RxLR based on manual annotation.

GENE ID	Genome Contig Number	Read signal start site	Read signal end site	Intron position	Reading frame with motif	Inferred protein sequence
RxLR 1	14	64,589	63,845	None	(-2)	MRFSVFLALAIATFVASISFTNAESVAHPESVRNLRTQXI PQGLVGKADDVAEKFKAKDIVAKFLESTRTNSLLKRAA ALEANGLKGDAAXLRANAAGISKLKPAQKQELGRMVA EIAKKDPEKLAQAVVQGAKNDPSRLAQAMGRVAKEEP EKLAQMANEVAKKNPKKWSTFQMVLVGLVAGGATYL GTKWLSSXST
RxLR 2	123	56,002	56,566	None	(+1)	MRCVYFVAFAVAVLARXSVVAAFTNADESKLLLKTTPD LAADALIGQKRFLRVXDPEDDDLTTGDEERGGXXKFAS LSAIIKKLDEQDLTHAAEVLKNMKQTHKENAEKAIKAAL DEGRIKAXDVDAVKALIGI
RxLR 3	24	315,155	314,586	None	(+2)	MRLXQVLVIAAASFLFASEAIAVTTDSEITVARDSLSQR RLRSYSKPAKEDDSDDSDASAKSGFTAEKDDDADEER GGFDHFSPEQIERLRKKAADLGYNFKHIESGTARFAAE DLKAWQEHLGQIIREKRSAGTAAHNAEWIARQRPHQ
RxLR 4	28	15,493	14,643	15,025-14,939	(-2)	MKVTKVLAALCVALLATAAGSEGVADENRDVSSRRLR QEFAAKPGETGGVLKDSTNPLRRRDQALVSAHRVYDP ASGLACSLVGDCVACPRSERDESFCRETGYRQELDCP RPKDPKDAALLTRPEDERETRFKACSPGDTSRPGVAV

						VKFEALMTVVLVVSFVLLRRERGNHMSSFDLRKDPRQ RAGLLGGGASSDKGSD
	1	1,334,192	1,333,348	1,333,727-1,333,637	(-1)	MKVTKVLAALCVALLATAAGSEGVADENRDVSSRRLR QEFAAKPGETGGVLKDSTNPLRRRDQALVSAHRVYDP ASGLACSLVGDCVACPRSERDESFCRETGYRQELDCP RPKDPKDAALLTRPEDERETRFKACSPGDTSRPGVAV VKFEVRGEREQKR
					(-1)	MKVTKVLAALCVALLATAAGSEGVADENRDVSSRRLR QEFAAKPGETGGVLKDSTNPLRRRDQALVSAHRVYDP ASGLACSLVGDCVACPRSERDESFCRETGYRQELDCP RPKDPKDAALLTRPEDERETRFKACSPGDTSRPGVAV VKFEALMTVVLVVSFVLLRRERGNHMSSFDLRKDPRQ RAGLLGGGASSDKGSD
RxLR 5	63	548,701	547,664	None	(-1)	MKVTKVLAALCVALLATAAGSEGVADENRDVSSRRLR QEFAAKPGETGGVLKDSTNPLRRRDQALVSAHRVYDP ASGLACSLVGDCVACPRSERDESFCRETGYRQELDCP RPKDPKDAALLTRPEDERETRFKACSPGDTSRPGVAV VKFEVRGEREQKR
RxLR 6	47	162,678	164,142	162,692-162,770	-	None inferred with RxLR and start codon - false prediction

RxLR 7	14	2,370,026	2,371,623	2,371,053-2,371,121 2371371-2371446	(+3)	MLSRLTLLRALGAHGVAAGGASAAVLSASFSSPAANRRARRLRRAKASAQAAAPDPRTPSLLADIERLSRNGLAQRVEMPAELLQRLXSVVRSRTHSQLETLRQKHVGDXRNTRQLPLDMSKTPLGWTMDRSQQIPPFAYGPAETLAFLAFEMEATYACTHAVFTELHKRLPDFKPTSVLDFGAGPGTASWVAKDFYDQSLDKYRVVEPSQSMVDAAEVLLDGFPGLSVRRNIADLSRDINAGNKYDLIVVSYVFSDITNDFERVATTSALWELLSENGCLVIVDRGSPWGSHHVRSARQFVLDSVKEDENGQEDVRIIAPCPHHFECPAAGSTWCHFVQRSPVVNRPREATTKRWHGQKGSKFSYMIMQKTRKGSDEDAAAKKKKPIARMLRSPLLATRHVHLDLCTPEGTMERRSVTRGKAIRDVYRASRKAHWGALWPADESSYLKDE
RxLR 8	5	204,598	202,656	None	(-1)	MKLPTLFLLALVSLSQLGGGAANEATDIMRRQLRVGKAVASLFENQHQTARDLQERIMQDEDNDDEVQAEPTKFRMRRLRSPNYVEIIE
	5	342,552	344,488	None	(+2)	MKLPTLFLLALVSLSQLGGGAANEATDIMRRQLRVGKAVASLFENQHQTARDLQERIMQDEDNDDEVQAEPTKFRMRRLRSPNYVEIIE
RxLR 9	50	149,001	148,377	None	(-1)	MKLSYLIAFAAVVVASTAVPASASTGLTTNNLAEDFQVAPELAALRSLRGANQDSSKGDDKKPKGEHGDKKKGGDKKKDGDKKKDGDKKKEKKEKKEKKEKRKEEEREKKHGKKDGDKKHGKKDGDKKHGKKDGEKKHGDKKHEQKDGKKDHKKTDGSKK

RxLR 10	18	476,924	476,034	None	(-1)	MRLGYLLLTAAVAPLLTNGAASKAATPEQVESFSTTIVQ NNEHGARLLRREGVDANDEERATVVLGGLIKSKEESK ALISSWVSNKDSVEVVAKRLGVNTALSPEKAVLQLNWA AFKRFERLKRVEETGPKRNYAYFGTGXQTKKTTEDALL VGHINGNTLDDVASLLRIEKGLSDHNLILHPNFRALKQY VNMQVKFQXMLNPIXFSQFDSAGGRKMFEQWAMEGT SLEVVAQTLHLAGLTGRALMEHKNYGAFMLYAKYYLD YAPRRAMKVGRARSAV
RxLR 11	67	461,337	462,155/ 462,953	None	(-1)	MKFNHCCAVAVLLLSFAASSTMATPNYNMRSADHPTS NAGSSTGSATRNLRGARALATTSSGTIVYTENGGDTSS SSTSTSSHGNDASQVVGXFATTAGLVIXMSAYLL
RxLR 12	62	731,770	730,850	None	(-1)	MWTSTSSPRLAALGLLVVLVLAVTTSSAMEHVGQRRL ESTLEDARSLRIHFKVKRASMAMYGHSEFDAFANPVVS SDKTTVRYDGYATFMENGTNHTISMVDGIAYFITTTAD GSVTAECTTSSSLALLDYIIPALNEATAISSATVDDKKVT CQSGDLFKVTLGDAIFVLCASGFSGFVVYGSDLNIIVKY LDSPVPIAAPTLSPDEARDCVAIVTPSPVTTTTLALLTGE PIAYESTTQSLARYLSHYFSYAR
RxLR 13	120	96,238	96,834	None	(+3)	MRRNLARLLLAAALLTVASEALSTVTDEEQAQISELGEE VIIRPKRLLRAVSKIDEDEERAILGEQVLLKWSKLAEKHN LQTLSKKLADKAWLTGKKADHKAWLKAGKNPQDIFKE YGLKGKTLEELKKDPRYARYDGFGDAWLKKQXKKGTI HTADNWIKLWNEQQKKVAG

	68	59,358	59,686	None	(-2)	MRRNLARLLLAAALLTVASEALSTVTDEEQAQISELGEE VIIRPKRLLRAVSKIDEDEERAILGEQVLLKWSKLAEKHN LQTLSKKLADKAWLTGKKAN
RxLR 14	35	678,651	679,285	None	(+3)	MRVLLWVLLVALVTLLSSTDALSTNNSDKKQVVQPNSE EVATRMLAANYENNNDKRMLRGESKMTYATNAENDE ERAVFIKFQGSVQRLREKXRXINPFSTKNIEKRFQKLAD KGRTPDYYFKKYQIGTFNSRHWNRRFYKKYEEWYKRT HPDWVSEITK
	12	16,772	16,145	16457-16376	(-3)	MRVLLWVLLVALVTLLSSTDALSTNNSDKKQVVQPNSE EVATRMLAANYENNNDKRMLRGESKMTYATNAENDE ERAVFIKFQGRQRENPRLLLQKVSNWYV
Hardham 5	1	884,928	885,623	None	(+3)	MKLSYVIAFAAVVVASTTVPASAATGLATRNLAEELQV GRQLGAGGLRGASQQGDHTQNTPAPGNTSQDDGKQ QSQDDGKQRSQDDGKQQDKKAEKEKKQQEKKEKQE KKDQEKKEKQEKKKSDNKDSGKQSEQQGEQSEKQGD KSEKQGDKSEKQGEQSEKQGSTEHPQNQSRK
	53	839,112	839,439	None	(-2)	Protein sequence is truncated – prediction not viable.
Hardham 6B	50	200,708	199,726		(-1)	MKLSYVIAFAAVVVASITVPASATTGLTTRNLAEELQVG RQLGAGGLRGASQQGDETQNTPTPGNTSQDDGKQQ SQDDGKQQSQDDGKQQDKKDDKEKKDDGDKKKDKK DDGDKKKDQEAEKEKKQQEKEKKQQEKEKKQQEKEK KQQEKEKKQQEKEKKQQEKKEKQEKKDQEKKEKQEK KKSDNKDSGKQSEQQGDKSEKQGDKPEKQGDKSEK

						QGDKPEKQGDKSEKQGSTQGDKSEKQGSTQGDKSEKQGEQSEKQGSTEPPQQKVGLQ
				200,253-200,191	(-1)	MKLSYVIAFAAVVVASITVPASATTGLTTRNLAEELQVGRQLGAGGLRGASQQGDETQNTPTPGNTSQDDGKQQSQDDGKQQSQDDGKQQDKKDDKEKKDDGDKKKDKKDDGDKKKDQEAEKEKKQQEKEKKQQEKEKKQQEKKEKQEKKDQEKKEKQEKKKSDNKDSGKQSEQQGDKSEKQGDKPEKQGDKSEKQGDKPEKQGDKSEKQGSTQGDKSEKQGSTQGDKSEKQGEQSEKQGSTEPPQQKVGLQ
				200,535-?	?	Can't predict third splice variant since second exon boundary could not be clearly visualized in IGV.
Hardham 6C	50	197,378	196,542	None	(-2)	MKLSYVIAFAAVVVASITVPASATTGLTTRNLAEELQVGRQLGAGGGLRGASQQGDHTQNTPTPGNTSQDDGKQHDKKDDKEKKDDGDKKKDDGDKKKDKKDDGDKKKDQEAEKEKKQQEKEKKQQEKEKKQQEKKEKQEKKDQEKKEKQEKKKSDNKDSGKQSEQQGDKSEKQGDKSEKQGDKPEKQGDKSEKQGSTQGDKSEKQGDKSEKQGSTEPPQQKVGLQ

Table 13. Summary of results obtained from Custom BLAST to analyse similarity between the final RxLR protein sequence predictions and the original sequences predicted by the Joint Genome Institute (JGI).

New protein prediction	Original JGI protein ID	% Pairwise identity	Query coverage (%)	E-value	Grade* (%)
RxLR_1_contig_14_consensus_seq_(-2)	325329_protein_RXLR_1	67.7	64.14	1.28e-85	65.9
RxLR_2_contig_123_consensus_seq_(+1)	100005_protein_RXLR_2	92.0	100	1.11e-87	96.0
RxLR_3_contig_24_consensus_seq_(+2)	220952_protein_RXLR_3	91.5	100	2.02e-92	95.8
RxLR_4_contig_28_consensus_seq_(-2)	97174_protein_RXLR_4	100	75.62	1.11e-109	87.8
RxLR_4_contig_1_consensus_seq_splicevariant1 (-1)	97174_protein_RXLR_4	98.8	100	2.81e-115	99.4
RxLR_4_contig_1_consensus_seq_splicevariant2 (-1)	97174_protein_RXLR_4	100	75.62	1.11e-109	87.8
RxLR_5_contig_63_consensus_seq_(-1)	97174_protein_RXLR_4	98.8	100	2.81e-115	99.4
RxLR_7_contig_14_consensus_seq_(+3)	313963_protein_RXLR_7	98.9	100	0	99.5
RxLR_8_contig_5_jg3373_cons_seq_(-1)	86482_protein_RXLR_8	100	100	6.42e-64	100
RxLR_8_contig_5_jg3405_cons_seq_(+2)	86482_protein_RXLR_8	100	100	6.42e-64	100
RxLR_9_contig_50_consensus_seq_(-1)	16230_protein_RXLR_9	100	40.99	1.37e-31	70.5

RxLR_10_contig_18_consensus_seq_(-1)	31697_protein_RXLR_10	68.7	81.07	2.04e-104	74.9
RxLR_11_contig_67_consensus_seq_(-1)	80229_protein_RXLR_11	98.2	100	1.06e-74	99.1
RxLR_12_contig_62_consensus_seq_(-1)	90101_protein_RXLR_12	100	100	0	100
RxLR_13_contig_68_consensus_seq_(-2)	88750_protein_RXLR_13	98	100	1.49e-54	99.0
RxLR_13_contig_120_consensus_seq_(+3)	88750_protein_RXLR_13	94.2	100	7.90e-105	97.1
RxLR_14_contig_35_consensus_seq_(+3)	30885_protein_RXLR_14	95.9	90	4.41e-103	92.9
RxLR_14_contig_12_consensus_seq_(-3)	30885_protein_RXLR_14	98.5	66.67	5.42e-45	82.6
Hardham_5_contig_1_consensus_seq_(+3)	Protein_24296_Hardham_5	83.1	100	1.27e-83	91.5
Hardham_6_contig_50_first_CDS_consensus_seq_(-2)	Protein_297058_Hardham_6	88.6	31.86	3.92e-35	60.2
Hardham_6_contig_50_second_CDS_consensus_seq splicevariant1_(-1)	Protein_297058_Hardham_6	100	30.4	1.27e-46	65.2
Hardham_6_contig_50_second_CDS_consensus_seq splicevariant2_(-1)	Protein_297058_Hardham_6	100	32.94	1.98e-46	66.5

*The grade provided by Custom BLAST is the weighted score for the hit based on the pairwise identity, coverage and E-value.

Figures

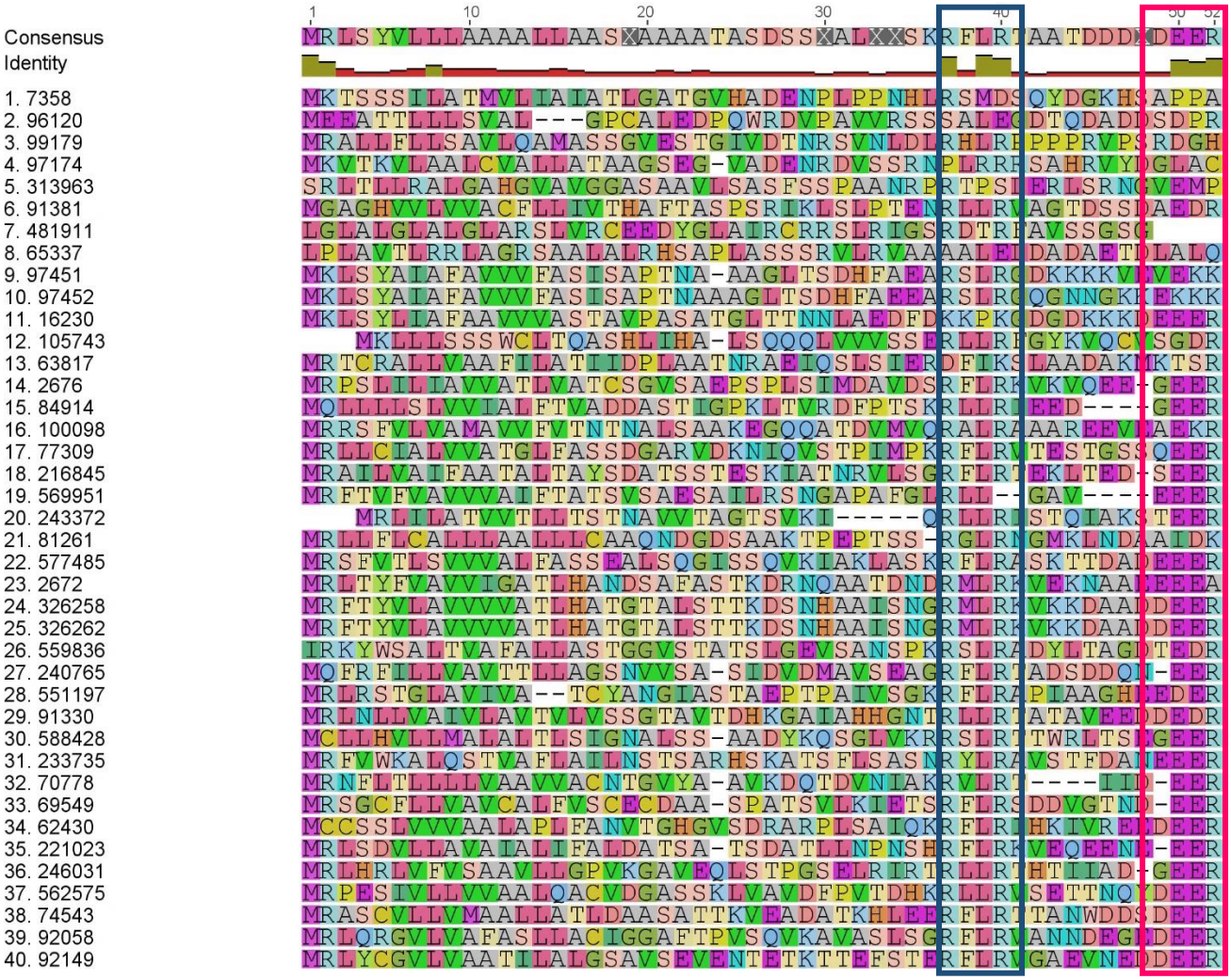
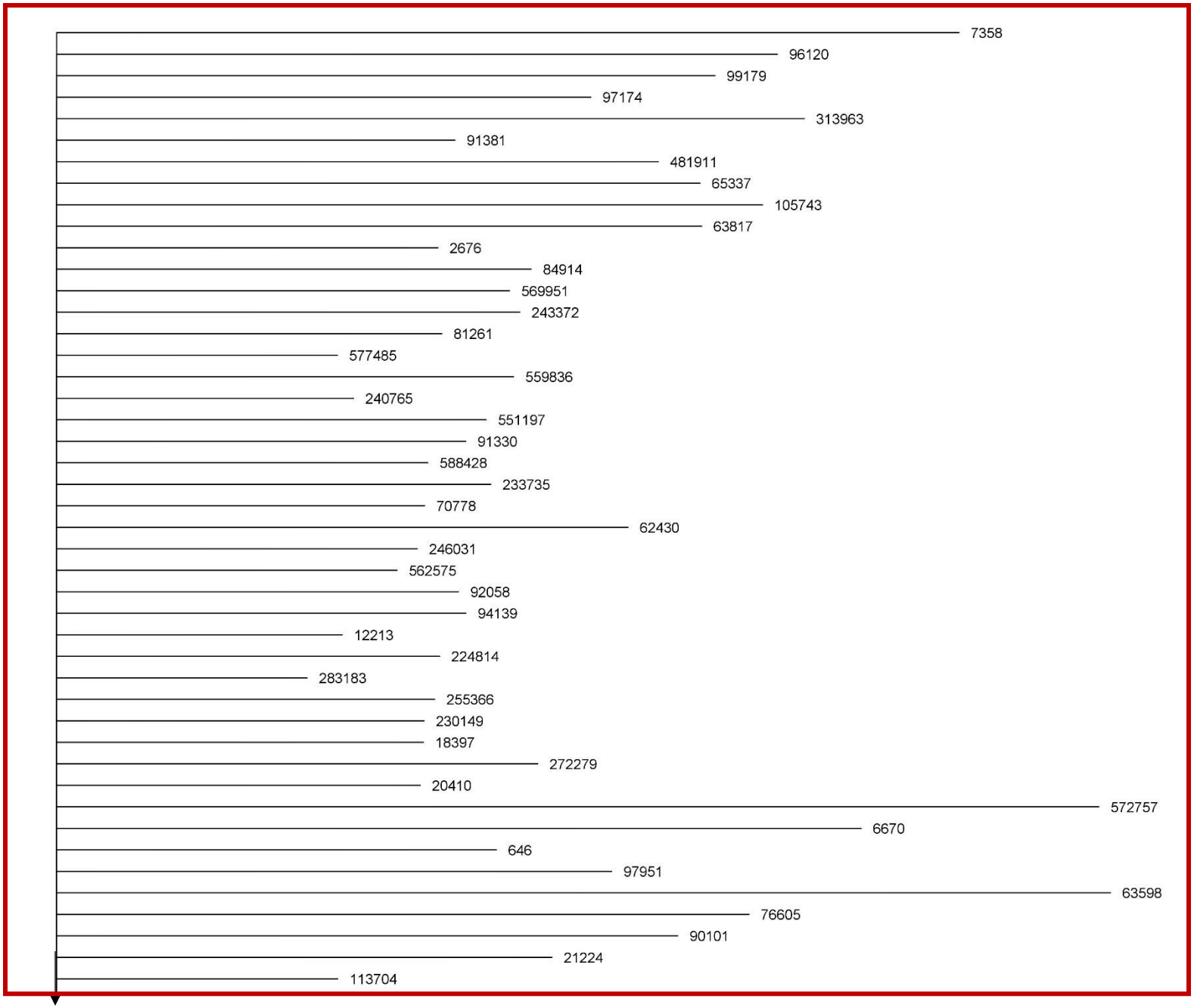
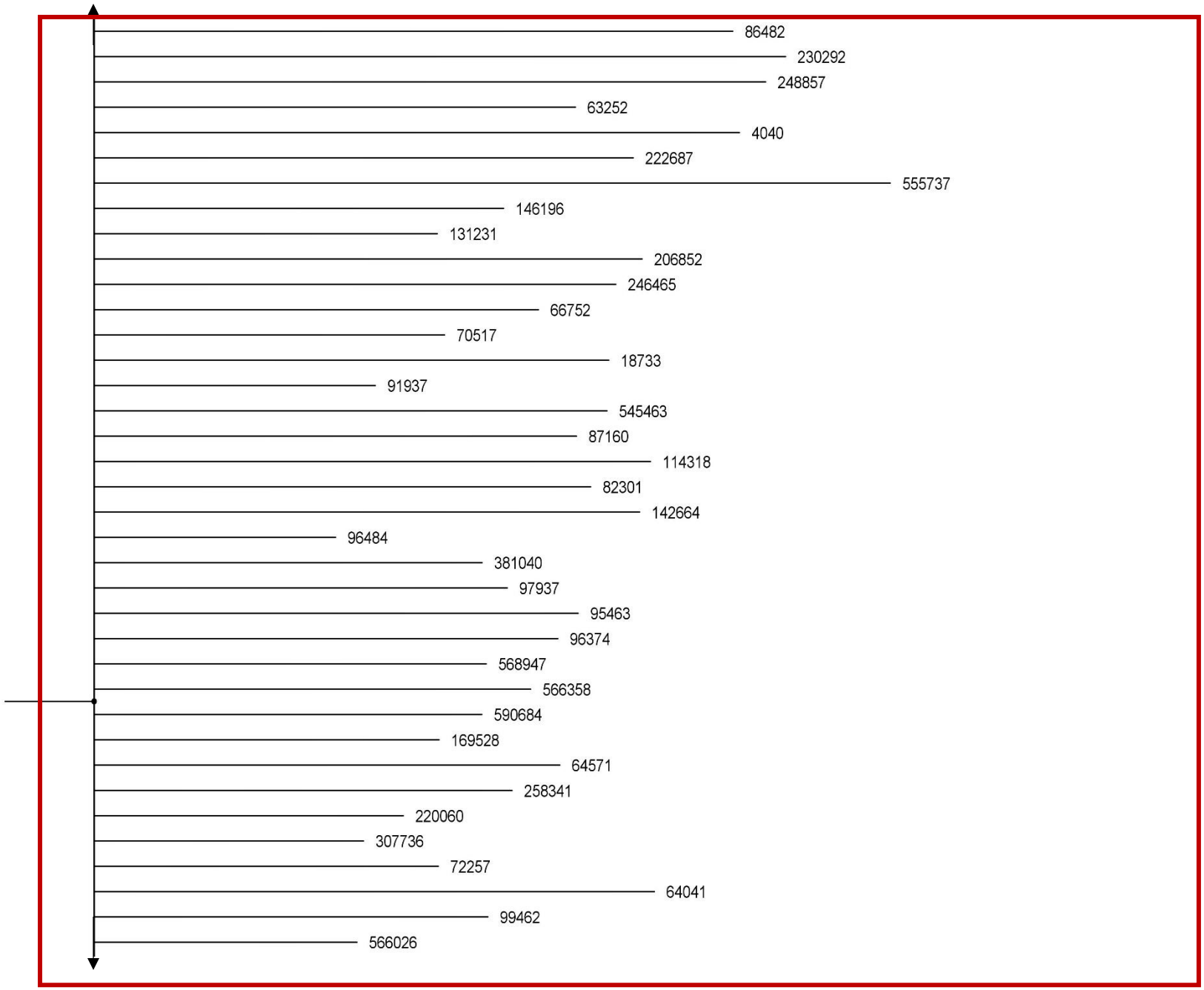
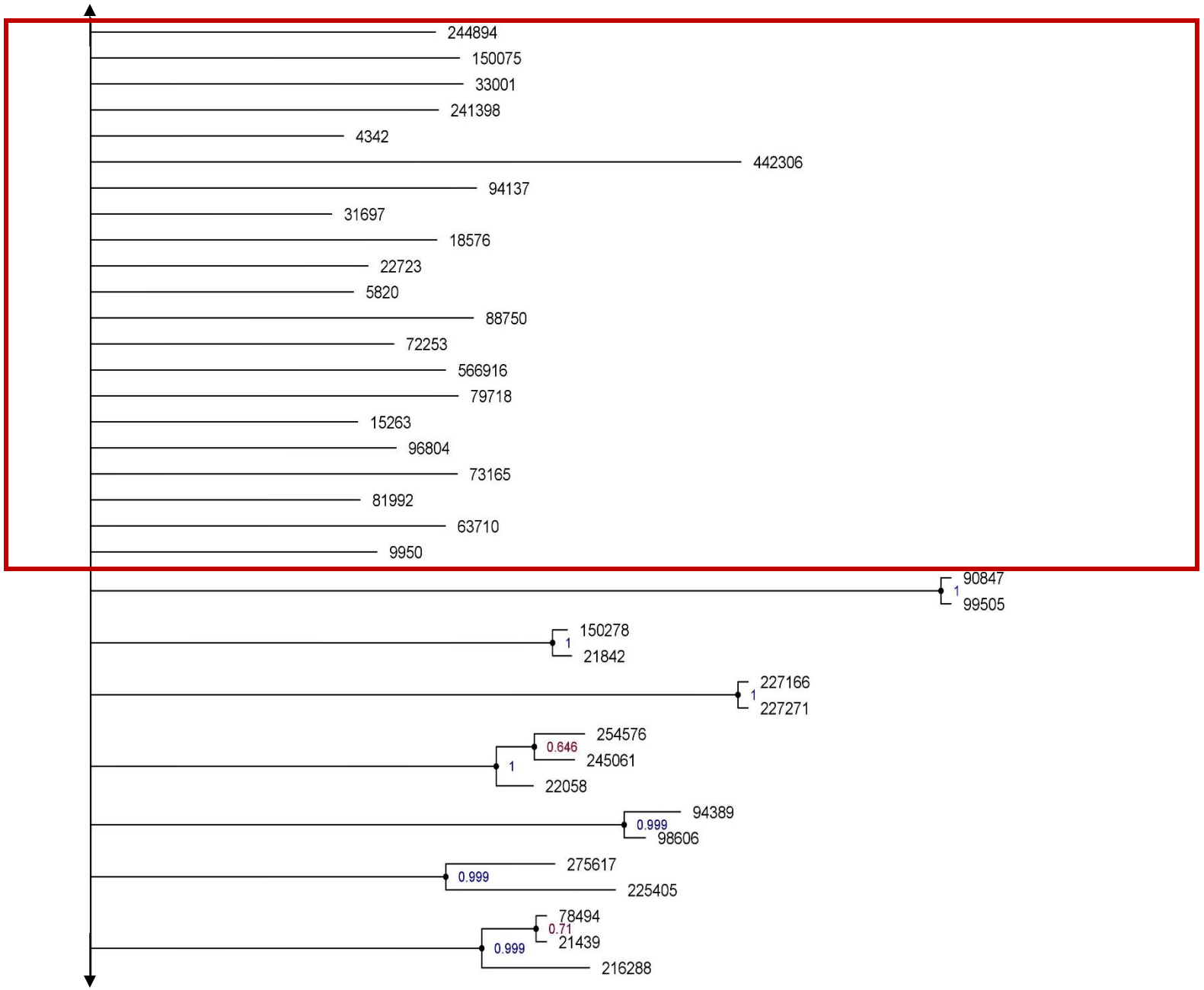
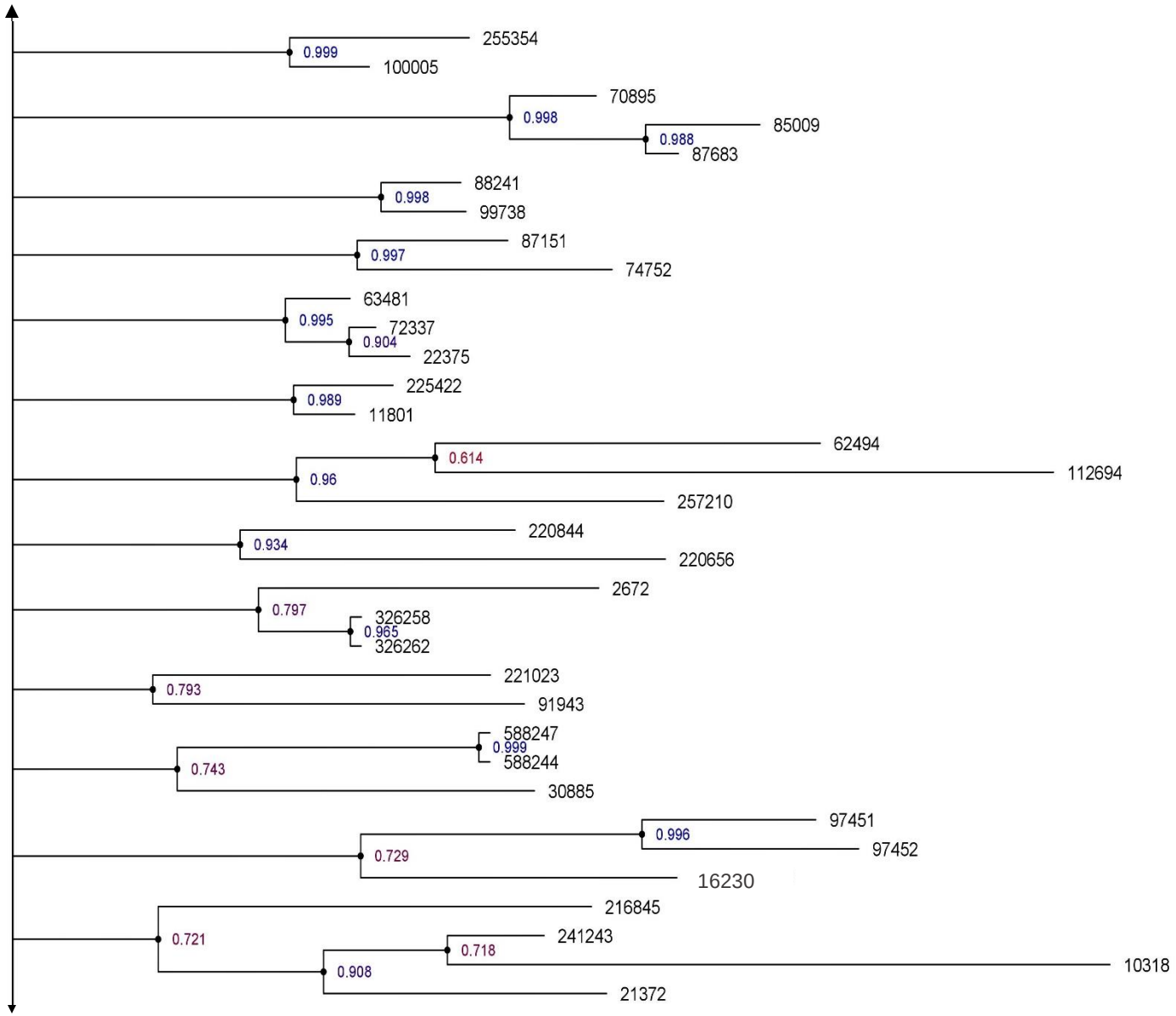


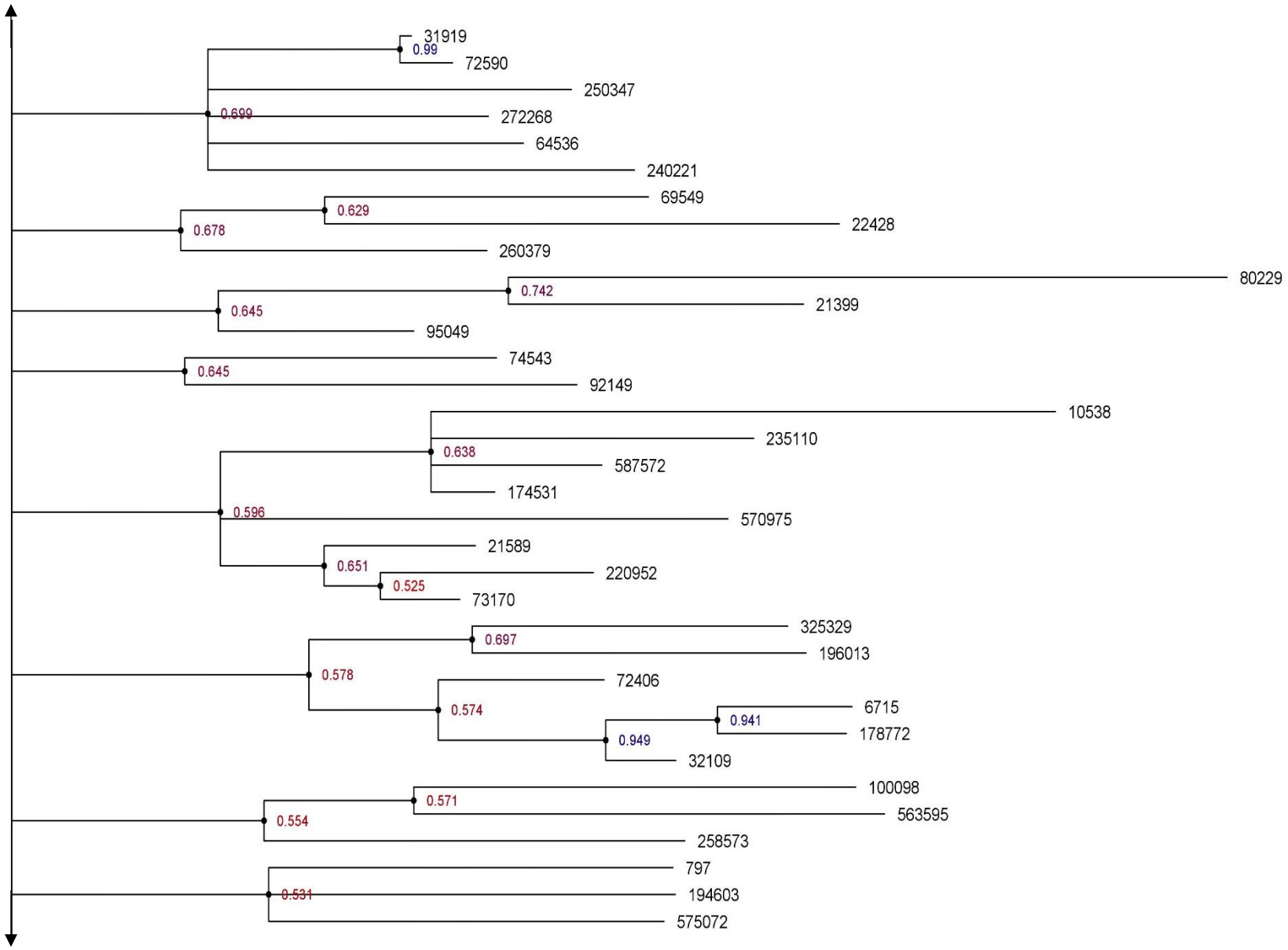
Figure 1. Protein sequence alignment of the N-terminal regions of the 192 candidate RxLRs predicted from the *Phytophthora cinnamomi* genome. Alignment was performed using MUSCLE protein alignment, uninformative sequences were removed, and the alignment was visualized in Geneious Prime 2020.0.3. The first 40 sequences of the alignment and the consensus generated by the alignment are shown in this image. The RxLR motif is highlighted by the blue block, while the dEER motif is indicated by the pink block in the figure.


7358
96120
99179
97174
313963
91381
481911
65337
105743
63817
2676
84914
569951
243372
81261
577485
559836
240765
551197
91330
588428
233735
70778
62430
246031
562575
92058
94139
12213
224814
283183
255366
230149
18397
272279
20410
572757
6670
646
97951
63598
76605
90101
21224
113704


86482
230292
248857
63252
4040
222687
555737
146196
131231
206852
246465
66752
70517
18733
91937
545463
87160
114318
82301
142664
96484
381040
97937
95463
96374
568947
566358
590684
169528
64571
258341
220060
307736
72257
64041
99462
566026


244894
150075
33001
241398
4342
442306
94137
31697
18576
22723
5820
88750
72253
566916
79718
15263
96804
73165
81992
63710
9950
90847
1
99505
150278
1
21842
227166
1
227271
254576
0.646
245061
1
22058
94389
0.999
98606
275617
0.999
225405
78494
0.71
21439
0.999
216288


255354
0.999
100005
70895
0.998
85009
0.988
87683
88241
0.998
99738
87151
0.997
74752
63481
0.995
72337
0.904
22375
225422
0.989
11801
62494
0.614
112694
0.96
257210
220844
0.934
220656
2672
0.797
326258
0.965
326262
221023
0.793
91943
588247
0.999
588244
0.743
30885
97451
0.996
97452
0.729
16230
216845
0.721
241243
0.718
10318
0.908
21372


31919
0.99
72590
250347
0.699
272268
64536
240221
69549
0.629
22428
0.678
260379
80229
0.742
21399
0.645
95049
74543
0.645
92149
10538
235110
0.638
587572
174531
0.596
570975
21589
0.651
220952
0.525
73170
325329
0.697
196013
0.578
72406
0.574
6715
0.941
178772
0.949
32109
100098
0.571
563595
0.554
258573
797
0.531
194603
575072

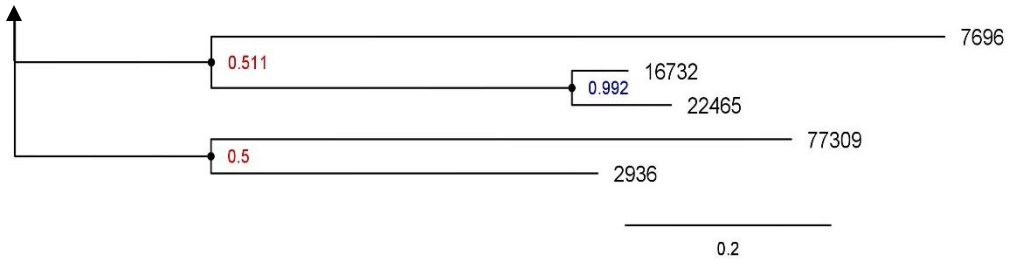
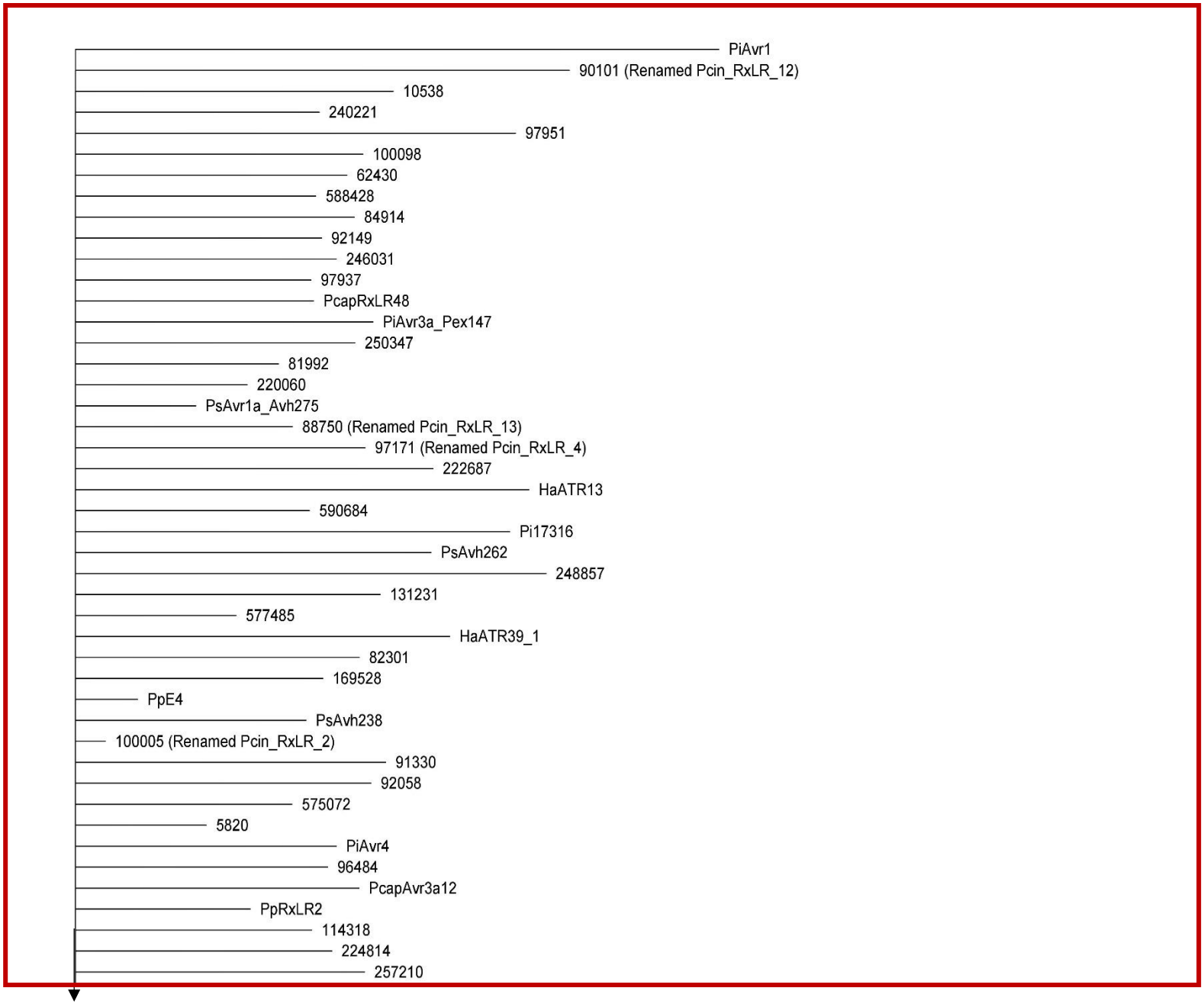
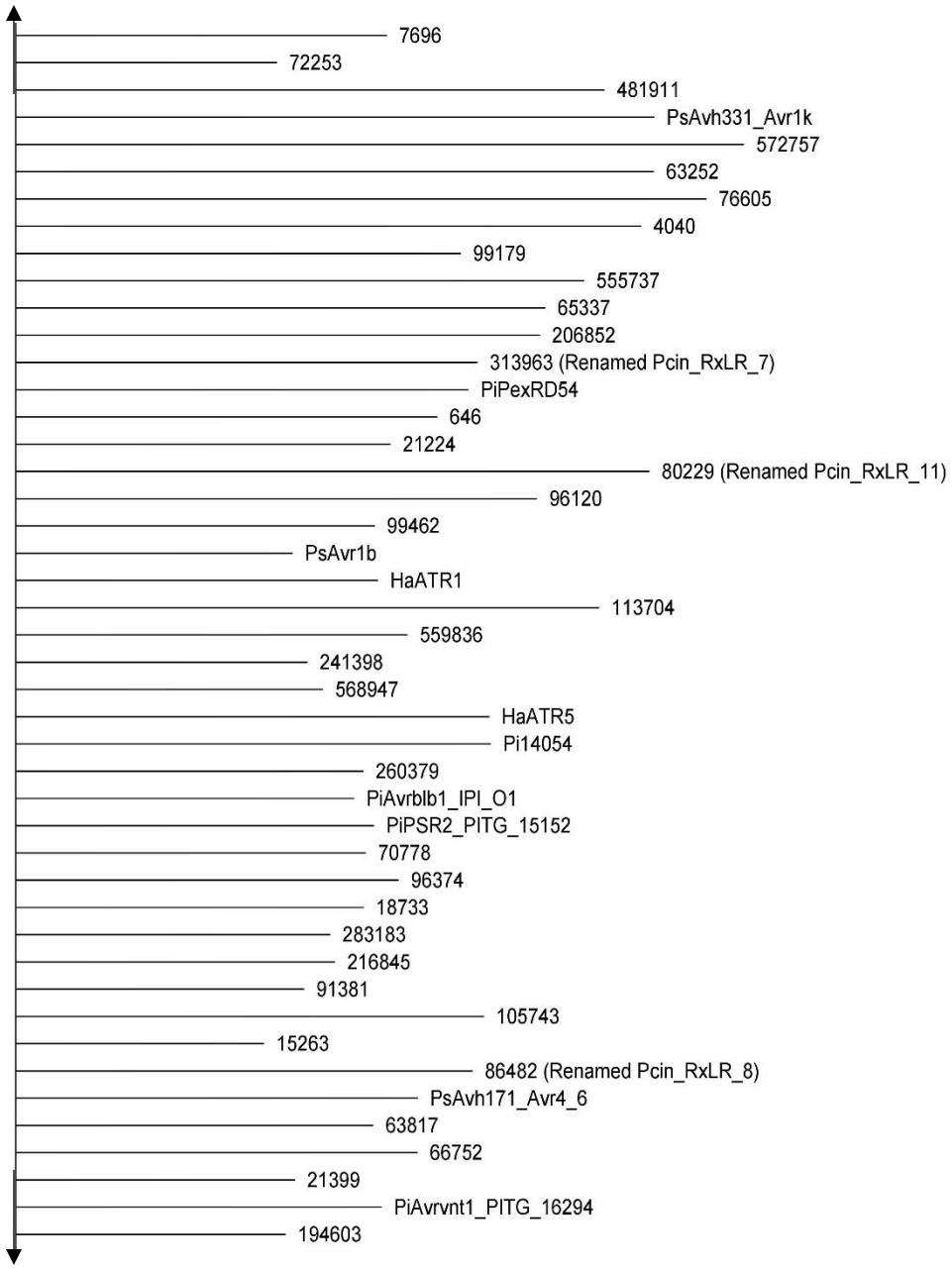
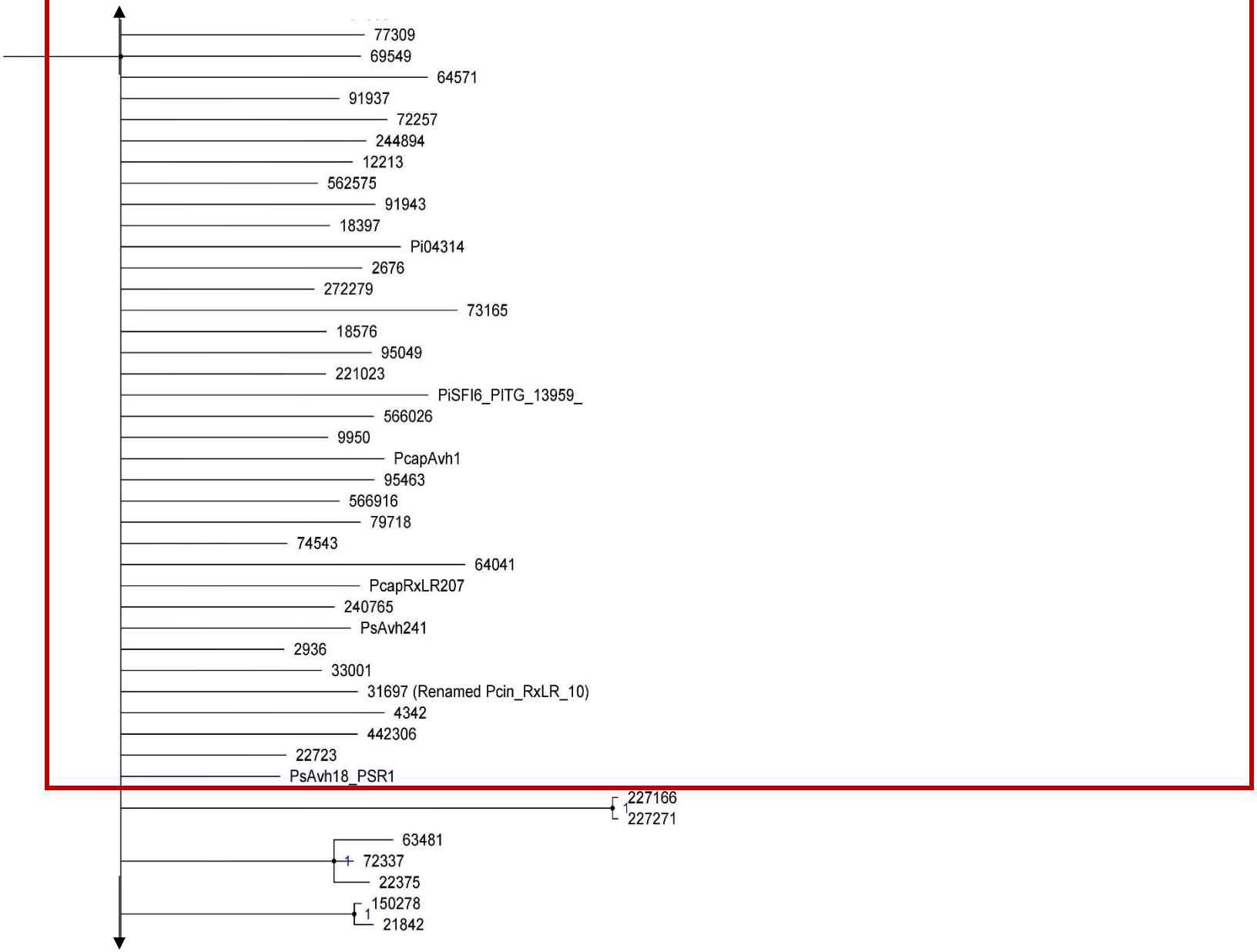
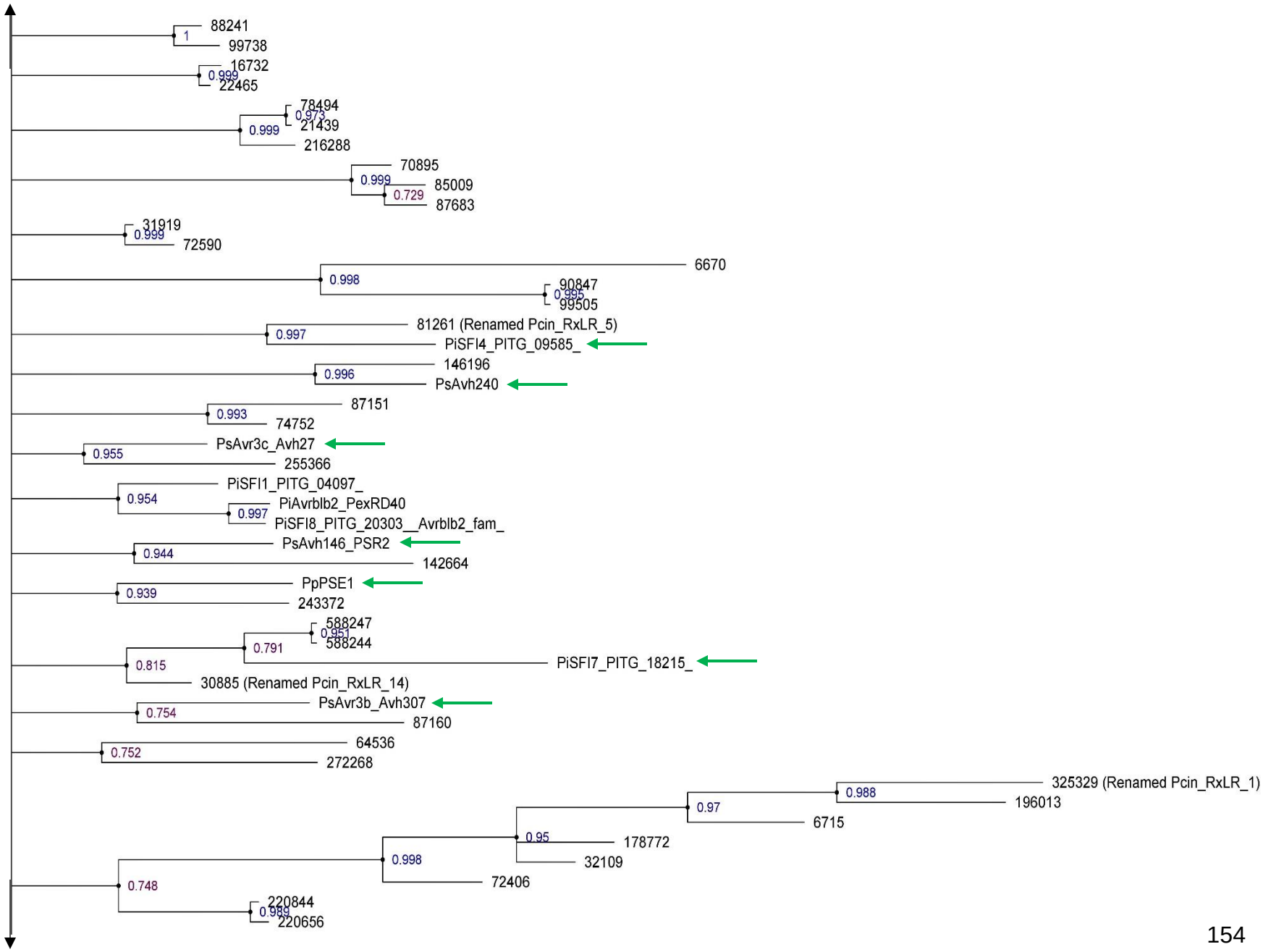


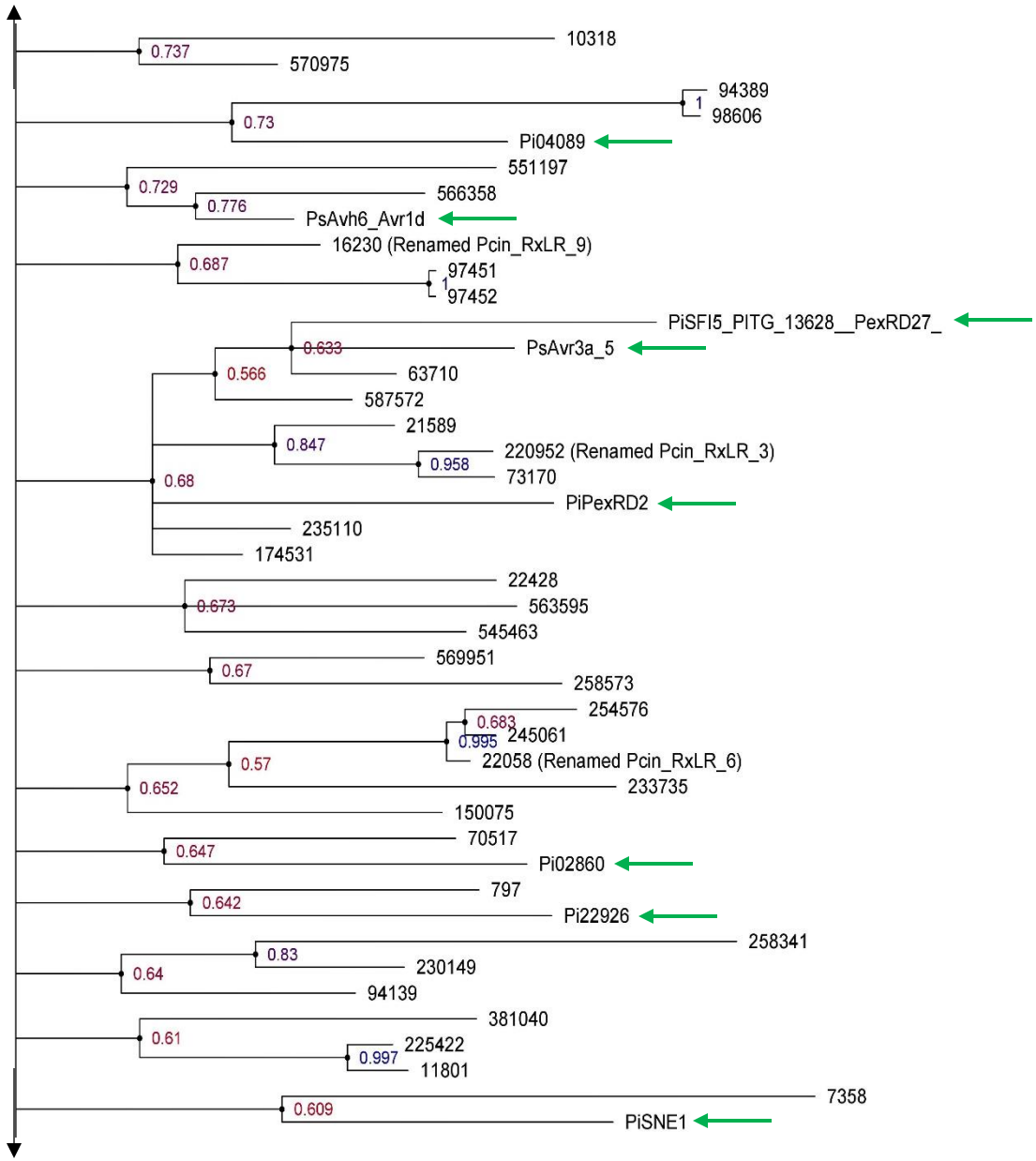
Figure 2. Phylogenetic tree resulting from Bayesian inference analysis of the aligned N-terminal regions of the *Phytophthora cinnamomi* candidate RxLR effectors. Branch tips of the tree are labelled with the protein IDs for each effector, as obtained from the Joint Genome Institute (JGI). Posterior probability values are shown up to the third significant digit for each node, coloured according to the probability value. Probability value colours fall on a spectrum ranging from red (lowest probability values) to blue (highest probability values). The sections indicated by red blocks represent sequences that were not inferred to be evolutionarily related to any other sequences and can all be perceived as outgroups.


PiAvr1
90101 (Renamed Pcin_RxLR_12)
10538
240221
97951
100098
62430
588428
84914
92149
246031
97937
PcapRxLR48
PiAvr3a_Pex147
250347
81992
220060
PsAvr1a_Avh275
88750 (Renamed Pcin_RxLR_13)
97171 (Renamed Pcin_RxLR_4)
222687
HaATR13
590684
Pi17316
PsAvh262
248857
131231
577485
HaATR39_1
82301
169528
PpE4
PsAvh238
100005 (Renamed Pcin_RxLR_2)
91330
92058
575072
5820
PiAvr4
96484
PcapAvr3a12
PpRxLR2
114318
224814
257210


7696
72253
481911
PsAvh331_Avr1k
572757
63252
76605
4040
99179
555737
65337
206852
313963 (Renamed Pcin_RxLR_7)
PiPexRD54
646
21224
80229 (Renamed Pcin_RxLR_11)
96120
99462
PsAvr1b
HaATR1
113704
559836
241398
568947
HaATR5
Pi14054
260379
PiAvrblb1_IPI_O1
PiPSR2_PITG_15152
70778
96374
18733
283183
216845
91381
105743
15263
86482 (Renamed Pcin_RxLR_8)
PsAvh171_Avr4_6
63817
66752
21399
PiAvrvnt1_PITG_16294
194603


77309
69549
64571
91937
72257
244894
12213
562575
91943
18397
Pi04314
2676
272279
73165
18576
95049
221023
PiSFI6_PITG_13959_
566026
9950
PcapAvh1
95463
566916
79718
74543
64041
PcapRxLR207
240765
PsAvh241
2936
33001
31697 (Renamed Pcin_RxLR_10)
4342
442306
22723
PsAvh18_PSR1
1
227166
227271
63481
1
72337
22375
1
150278
21842


88241
1
99738
16732
0.999
22465
78494
0.973
21439
0.999
216288
70895
0.999
85009
0.729
87683
31919
0.999
72590
6670
0.998
90847
0.985
99505
81261 (Renamed Pcin_RxLR_5)
0.997
PiSFI4_PITG_09585_
146196
0.996
PsAvh240
87151
0.993
74752
PsAvr3c_Avh27
0.955
255366
PiSFI1_PITG_04097_
0.954
PiAvrblb2_PexRD40
0.997
PiSFI8_PITG_20303__Avrblb2_fam_
PsAvh146_PSR2
0.944
142664
PpPSE1
0.939
243372
588247
0.951
588244
0.791
PiSFI7_PITG_18215_
0.815
30885 (Renamed Pcin_RxLR_14)
PsAvr3b_Avh307
0.754
87160
64536
0.752
272268
325329 (Renamed Pcin_RxLR_1)
0.988
196013
0.97
6715
0.95
178772
0.998
32109
72406
0.748
220844
0.988
220656


10318
0.737
570975
94389
1
98606
0.73
Pi04089
551197
0.729
566358
0.776
PsAvh6_Avr1d
16230 (Renamed Pcin_RxLR_9)
0.687
97451
1
97452
PiSFI5_PITG_13628__PexRD27_
0.633
PsAvr3a_5
0.566
63710
587572
21589
0.847
220952 (Renamed Pcin_RxLR_3)
0.958
73170
0.68
PiPexRD2
235110
174531
22428
0.673
563595
545463
569951
0.67
258573
254576
0.683
245061
0.995
22058 (Renamed Pcin_RxLR_6)
0.57
233735
0.652
150075
70517
0.647
Pi02860
797
0.642
Pi22926
258341
0.83
230149
0.64
94139
381040
0.61
225422
0.997
11801
7358
0.609
PiSNE1

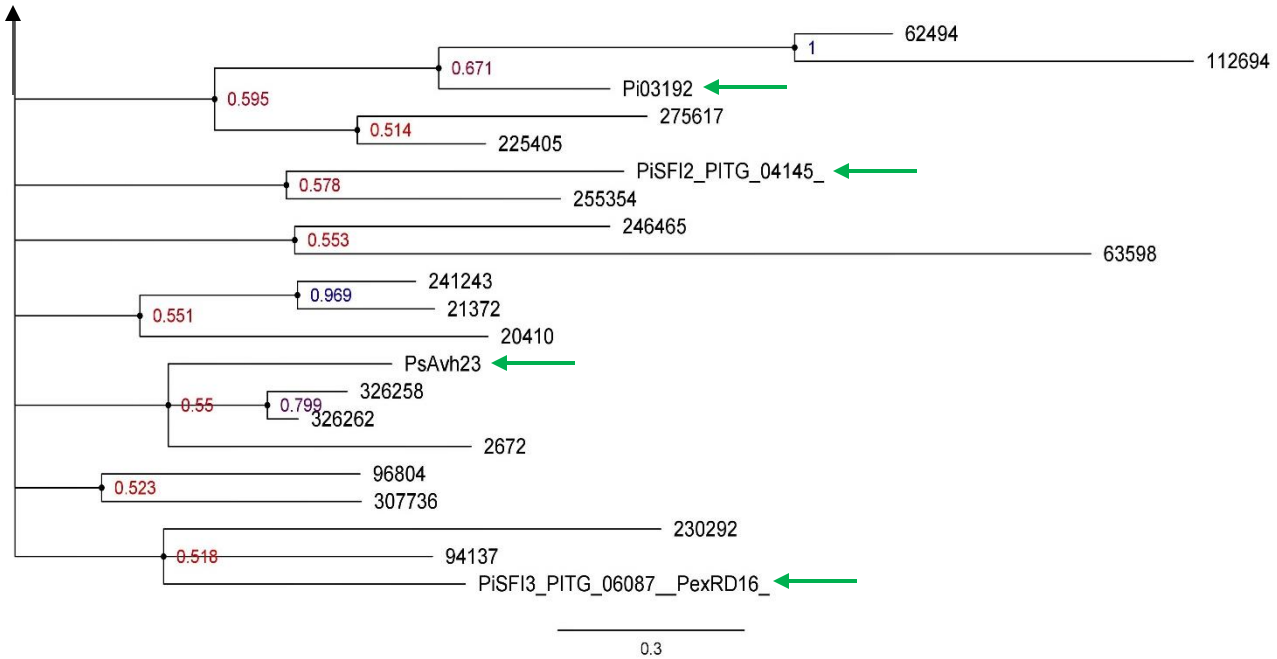


Figure 3. Phylogenetic tree resulting from Bayesian inference analysis of the N-terminal regions of the *Phytophthora cinnamomi* candidate RxLR effectors aligned with the N-terminal regions of functionally characterised RxLRs in other species. Branch tips of the tree are labelled with the protein IDs for *P. cinnamomi* effectors, as obtained from the Joint Genome Institute (JGI), or the names of confirmed RxLRs in other species, as obtained from literature. Where *P. cinnamomi* candidates were renamed in downstream analysis of this project, their new names are indicated in brackets next to original protein IDs. Posterior probability values are shown up to the third significant digit for each node, coloured according to the probability value. Probability value colours fall on a spectrum ranging from red (lowest probability values) to blue (highest probability values). The sections indicated by red blocks represent sequences that were not inferred to be evolutionarily related to any other sequences and can all be perceived as outgroups. Proteins indicated by green arrows are characterised RxLRs from other species which grouped with candidate *P. cinnamomi* RxLRs.

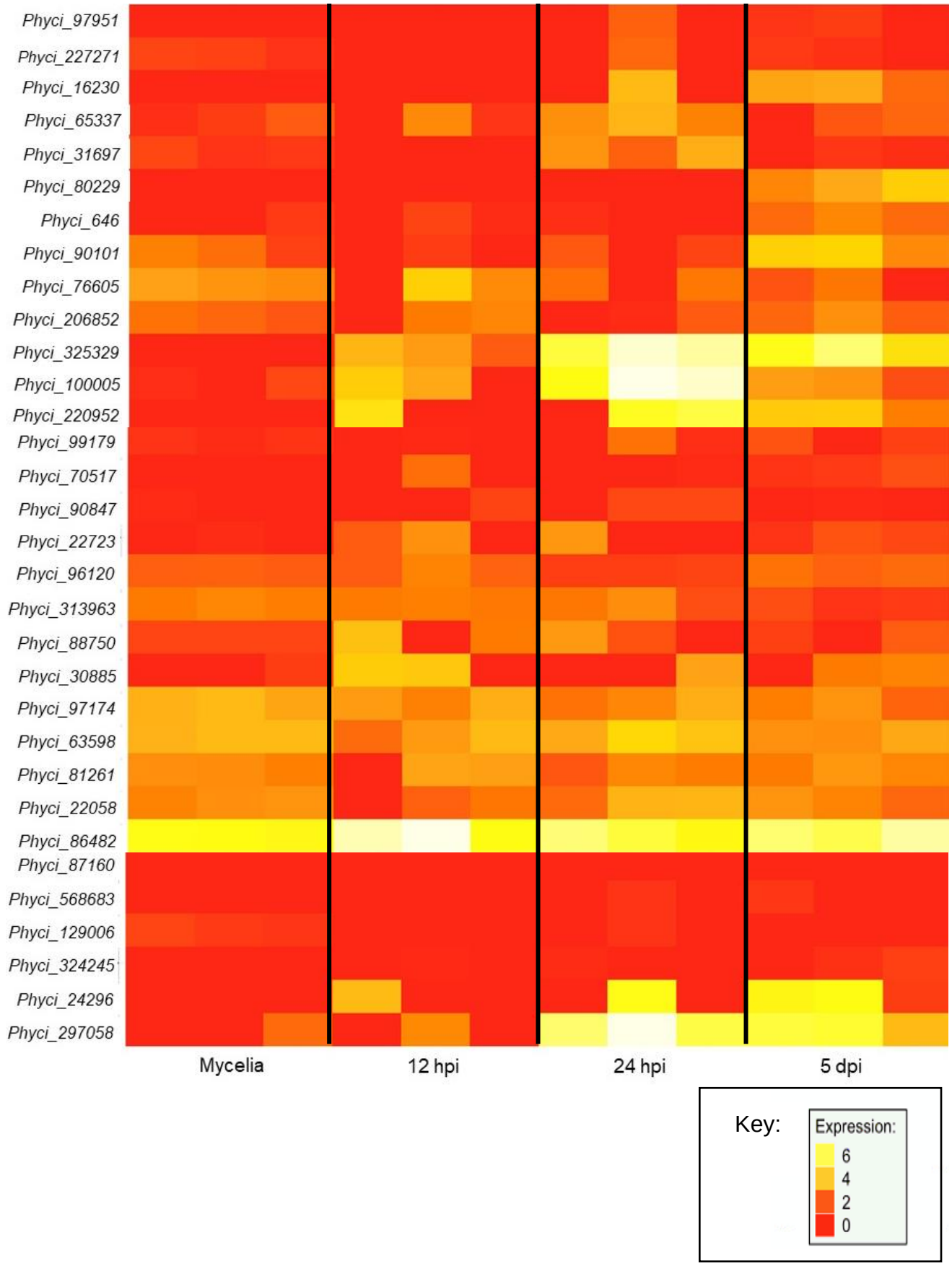

Phyci_97951
Phyci_227271
Phyci_16230
Phyci_65337
Phyci_31697
Phyci_80229
Phyci_646
Phyci_90101
Phyci_76605
Phyci_206852
Phyci_325329
Phyci_100005
Phyci_220952
Phyci_99179
Phyci_70517
Phyci_90847
Phyci_22723
Phyci_96120
Phyci_313963
Phyci_88750
Phyci_30885
Phyci_97174
Phyci_63598
Phyci_81261
Phyci_22058
Phyci_86482
Phyci_87160
Phyci_568683
Phyci_129006
Phyci_324245
Phyci_24296
Phyci_297058
Mycelia
12 hpi
24 hpi
5 dpi
Key:
Expression:
6
4
2
0

Figure 4. Expression patterns of *Phytophthora cinnamomi* candidate *RxLRs* visually represented in a heatmap, generated in Shiny, from Sleuth analysis of Kallisto outputs. Relative expression levels are indicated by colour on a scale from 0-6, with red (0 on the scale) being an indication of low or no expression, while light yellow or white (6 on the scale) indicates high expression. The relative levels of expression were colour-coded according to the scale generated by Shiny for each library. Each block signifies one biological replicate, and the three biological replicates for each timepoint are grouped together. Relative expression of each gene is shown for biological replicates from mycelium, as well as a susceptible avocado rootstock R0.12 harvested 12 hours post inoculation (hpi), 24hpi and 5 days post inoculation (5dpi).

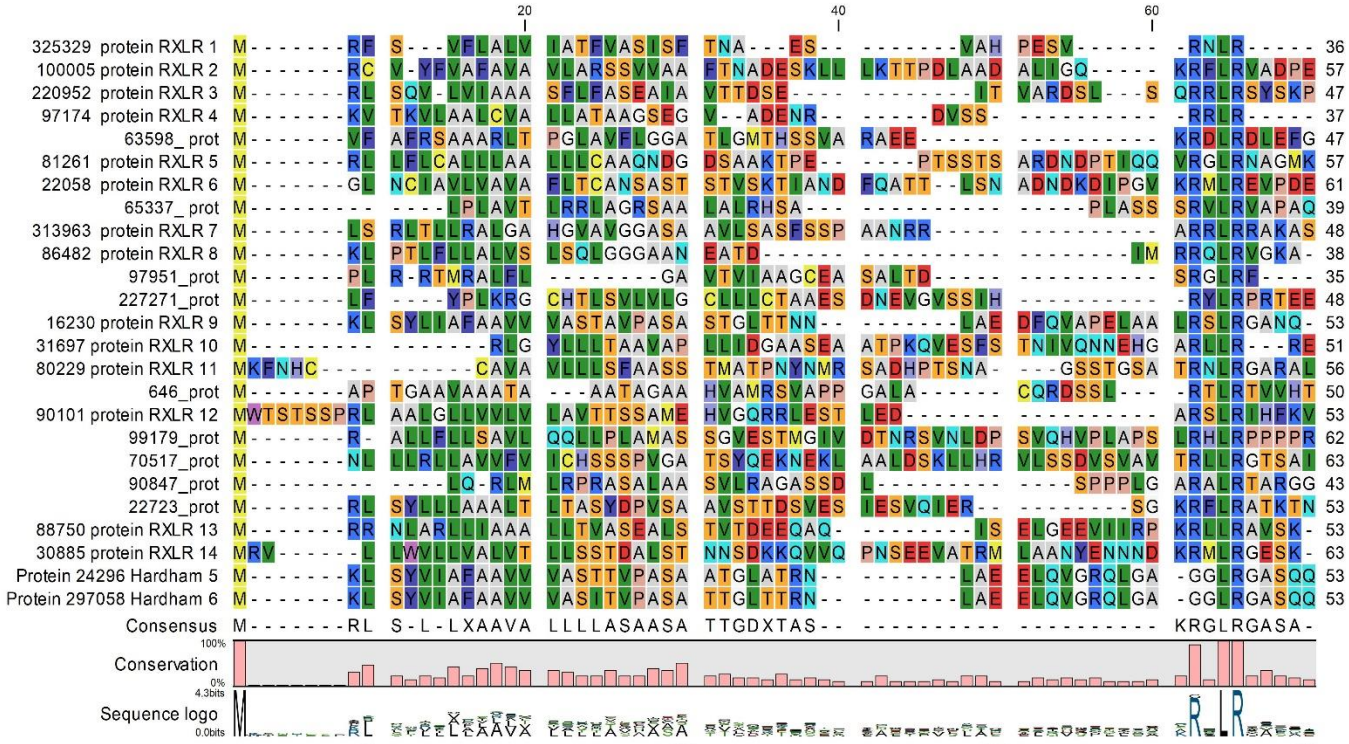


Figure 5. N-terminal region of the alignment of the peptide sequences of the 25 candidate genes. Alignment using RxLR motifs as alignment fixpoints for the relevant genes highlights the presence of the motif in all sequences except for Hardham 5 and Hardham 6. Since these two effectors had two possible RxLx motifs, no alignment fixpoints were selected for these sequences. Interestingly, neither of the RxLx motifs aligned with the RxLR consensus; the motif “GGLR” in both Hardham 5 and Hardham 6 aligned with the RxLR fixpoint in the other genes.

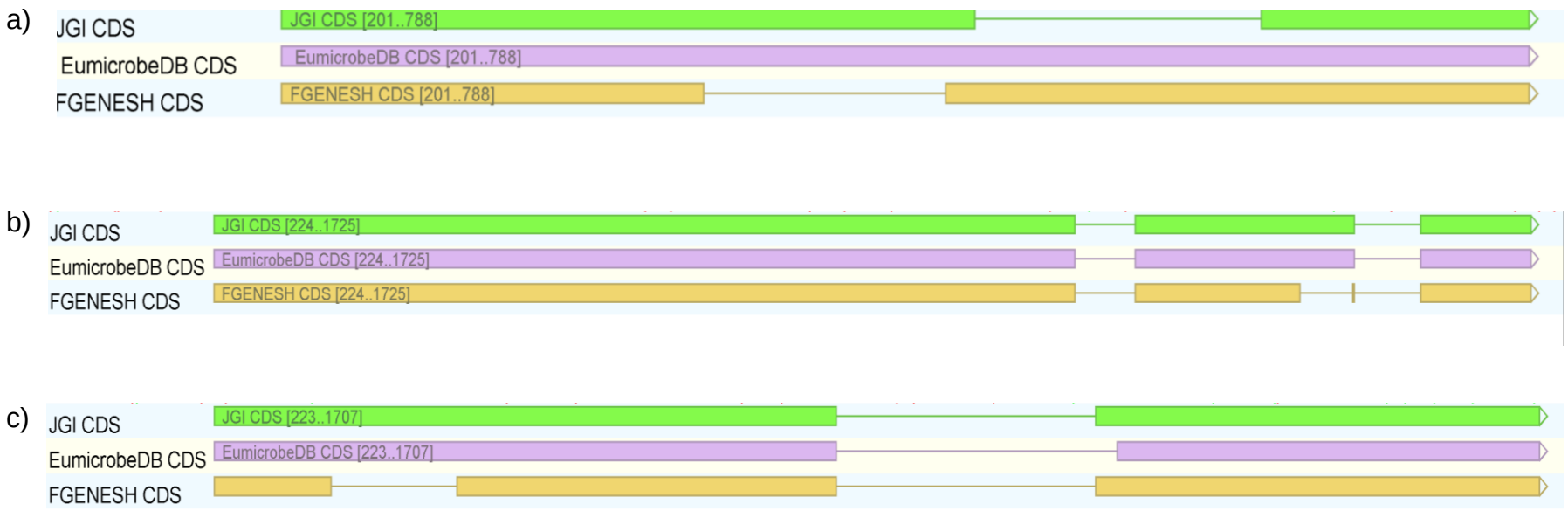
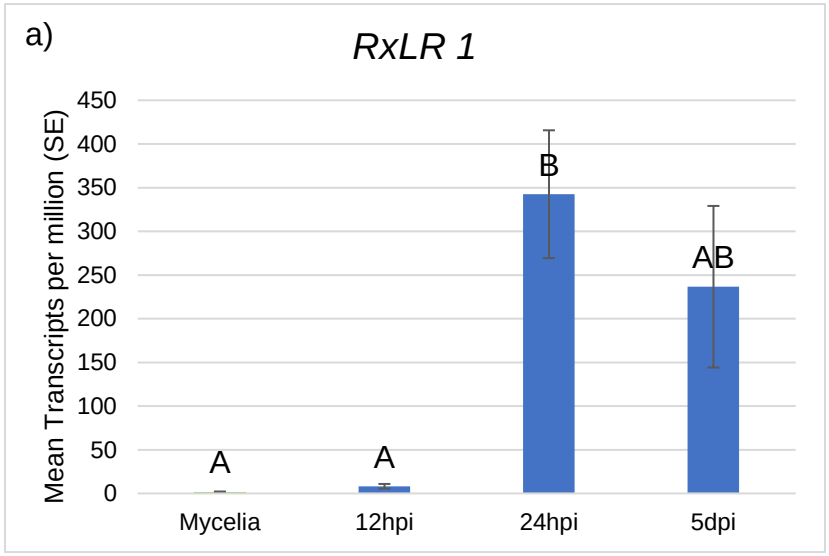
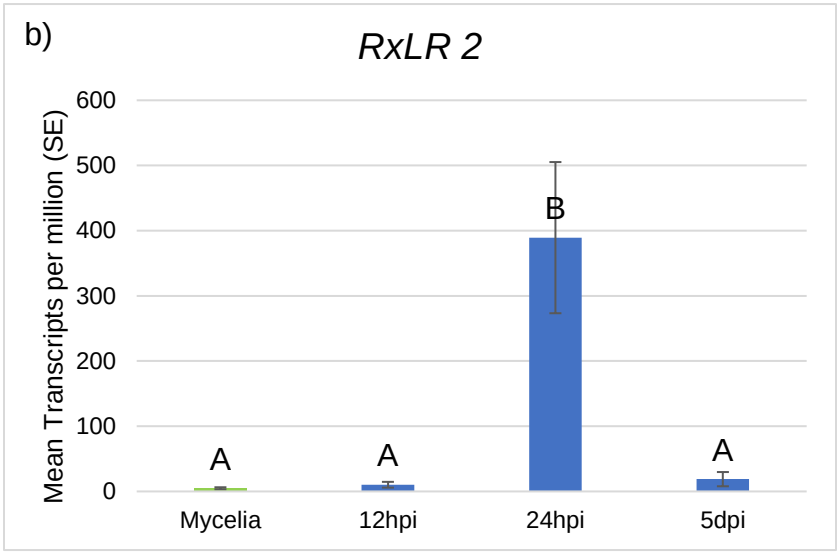
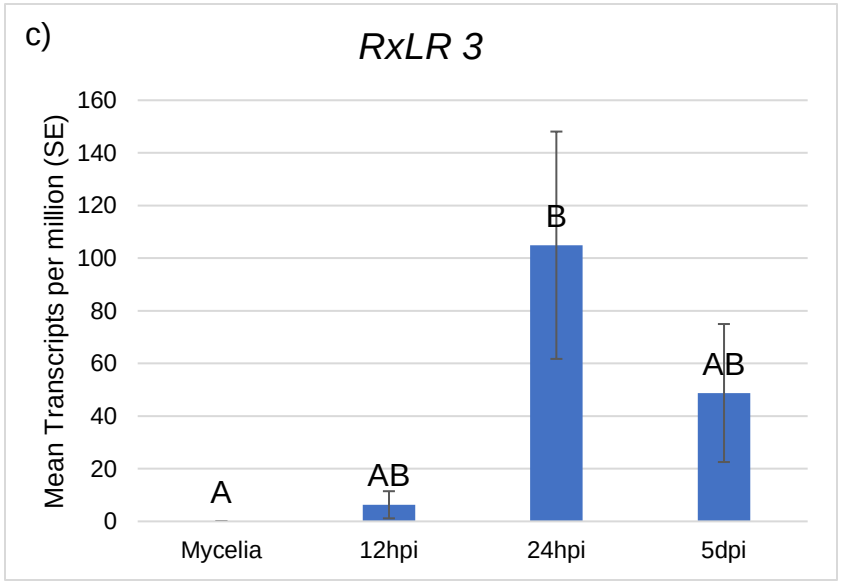
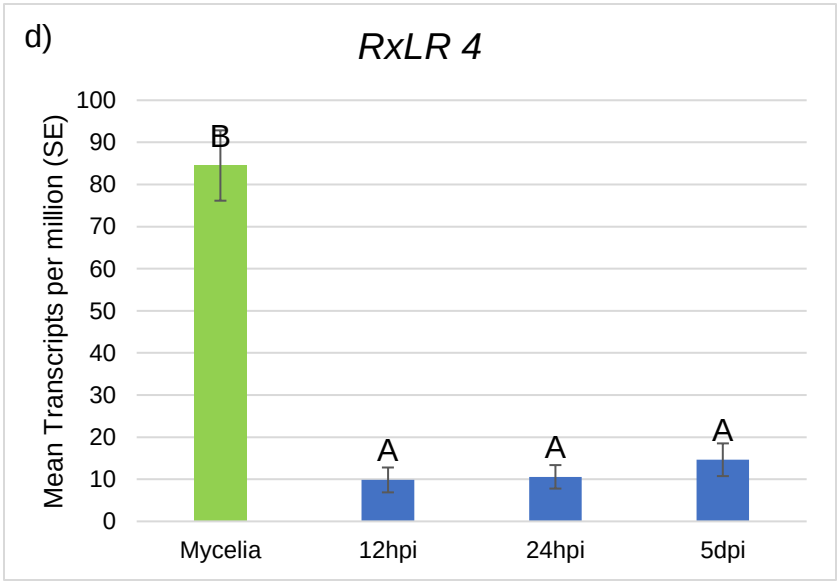


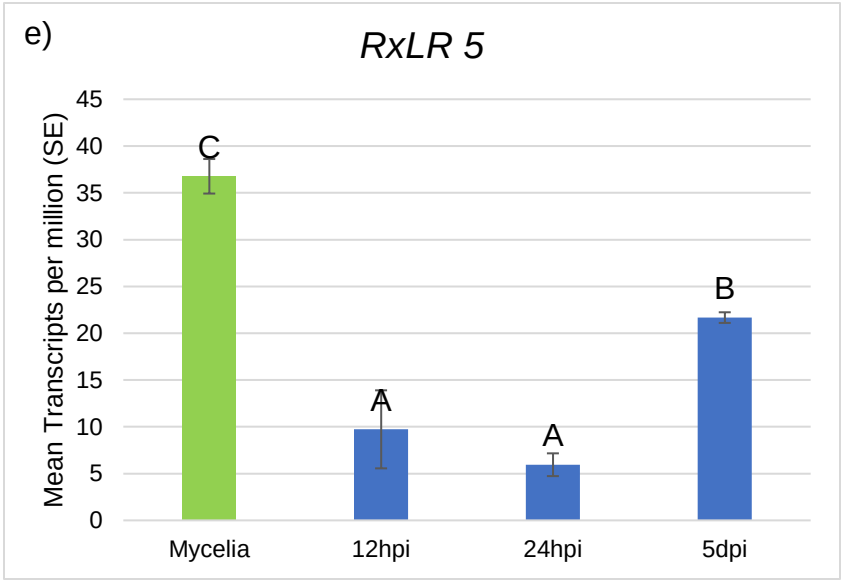
Figure 6. Examples of candidate RxLRs that had different gene predictions according to different predictions programmes used, illustrated using GenomeView and colour-coded according to the prediction software used. Solid bars represent predicted exons while lines represent introns within the prediction. Green bars show predicted regions according to JGI, purple bars show predictions obtained from EumicrobeDB and orange bars show FGENESH predictions. The examples shown in this figure, *RxLR 1* (a), *RxLR 7* (b), and *RxLR 10* (c), had different intron/exon boundaries predicted by the different software, which can be seen as differences in where bars and lines start and end for each example in the figure above. Bars in the different examples are not drawn to scale.

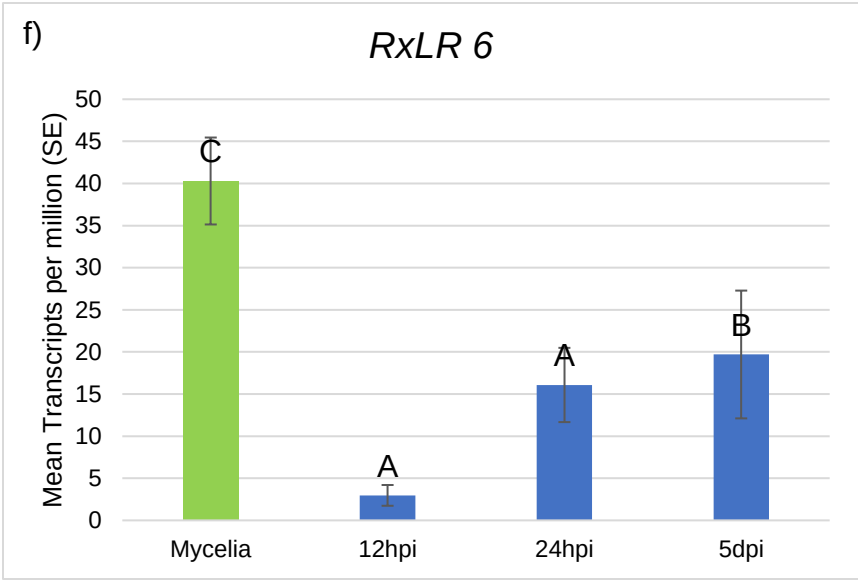

a)
RxLR 1
Mean Transcripts per million (SE)
450
400
350
300
250
200
150
100
50
0
A
A
B
AB
Mycelia
12hpi
24hpi
5dpi

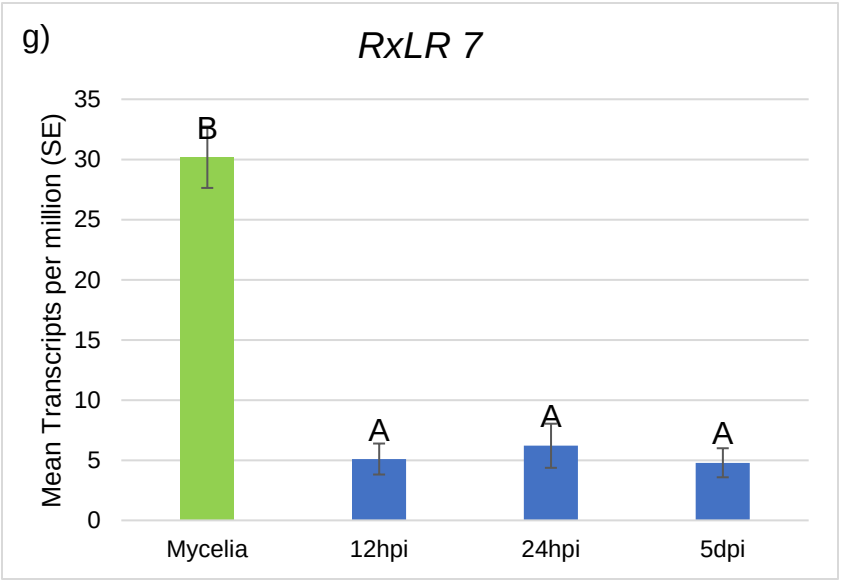

b)
RxLR 2
Mean Transcripts per million (SE)
600
500
400
300
200
100
0
A
A
B
A
Mycelia
12hpi
24hpi
5dpi

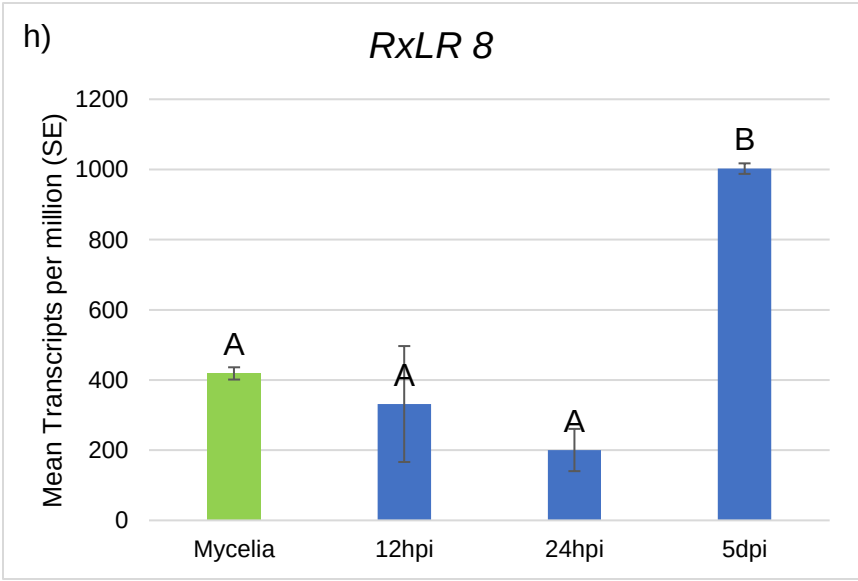

c)
RxLR 3
Mean Transcripts per million (SE)
160
140
120
100
80
60
40
20
0
A
AB
B
AB
Mycelia
12hpi
24hpi
5dpi

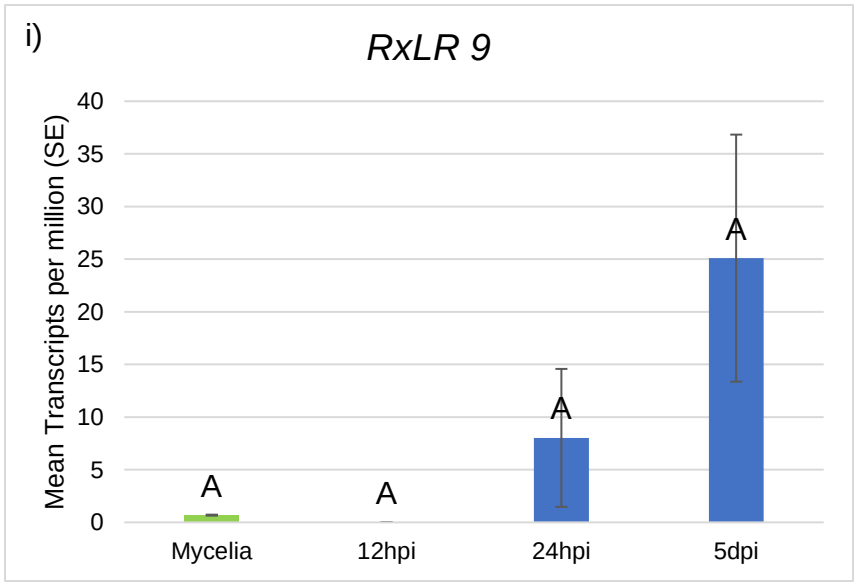

d)
RxLR 4
Mean Transcripts per million (SE)
100
90
80
70
60
50
40
30
20
10
0
B
A
A
A
Mycelia
12hpi
24hpi
5dpi

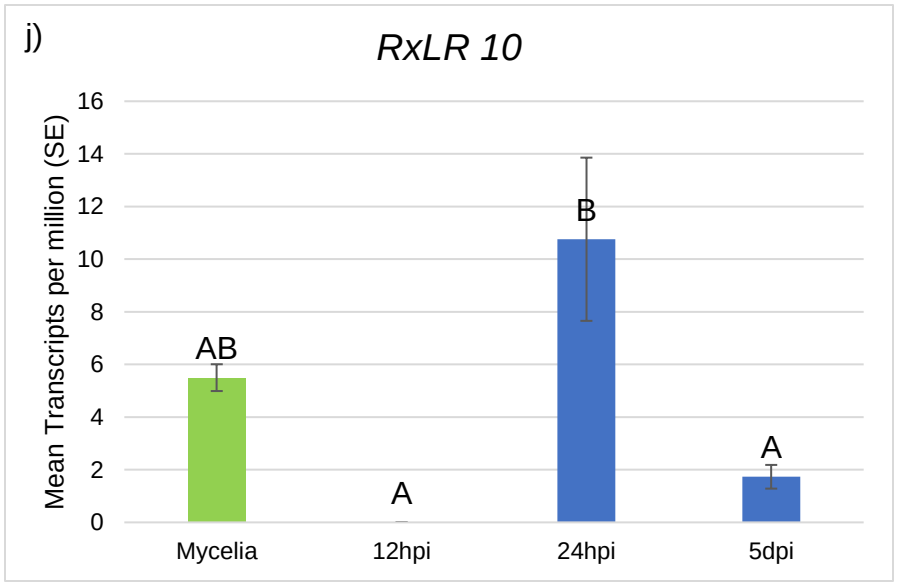

e)
RxLR 5
Mean Transcripts per million (SE)
45
40
35
30
25
20
15
10
5
0
C
A
A
B
Mycelia
12hpi
24hpi
5dpi

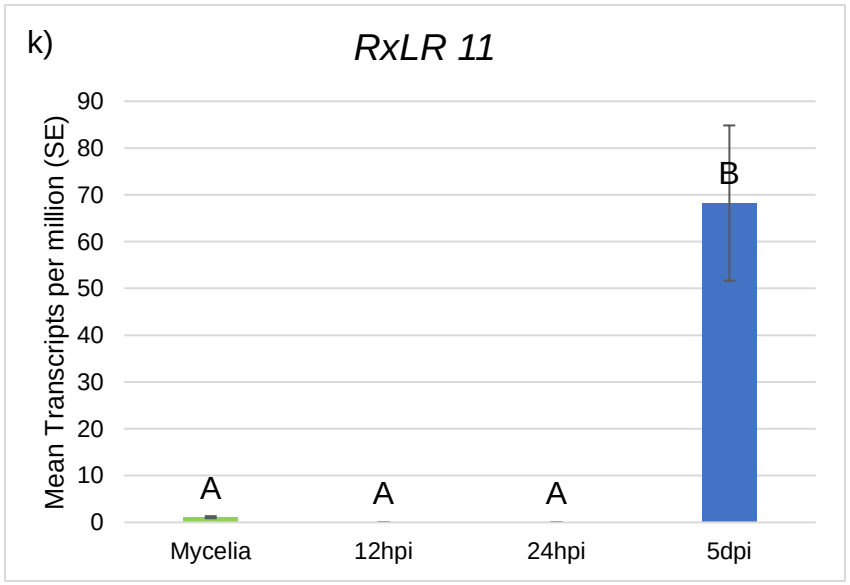

f)
RxLR 6
Mean Transcripts per million (SE)
50
45
40
35
30
25
20
15
10
5
0
C
A
A
B
Mycelia
12hpi
24hpi
5dpi

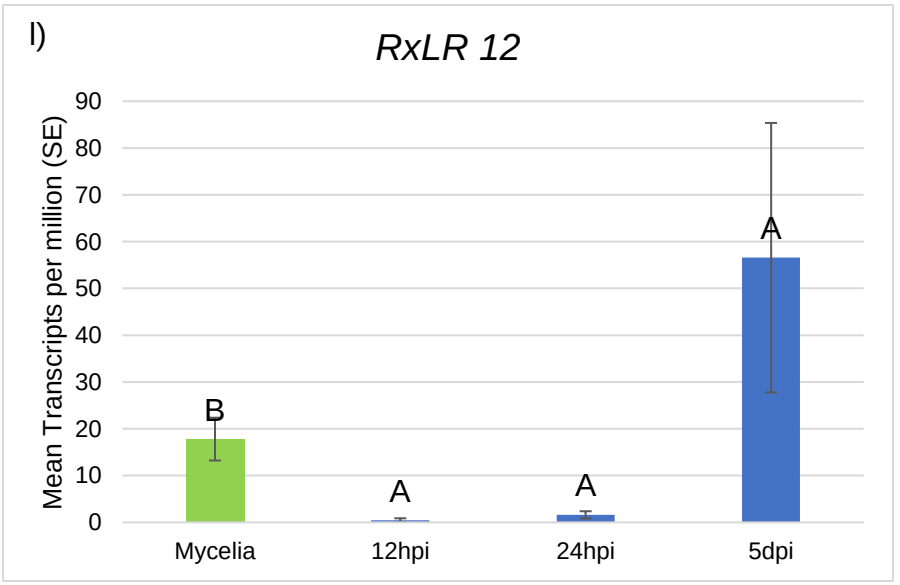

g)
RxLR 7
Mean Transcripts per million (SE)
35
30
25
20
15
10
5
0
B
A
A
A
Mycelia
12hpi
24hpi
5dpi

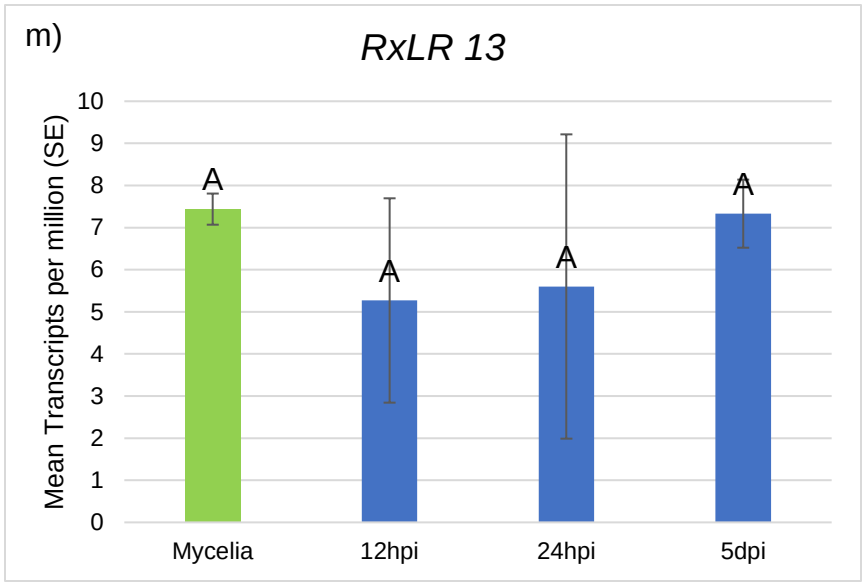

h)
RxLR 8
Mean Transcripts per million (SE)
1200
1000
800
600
400
200
0
A
A
A
B
Mycelia
12hpi
24hpi
5dpi

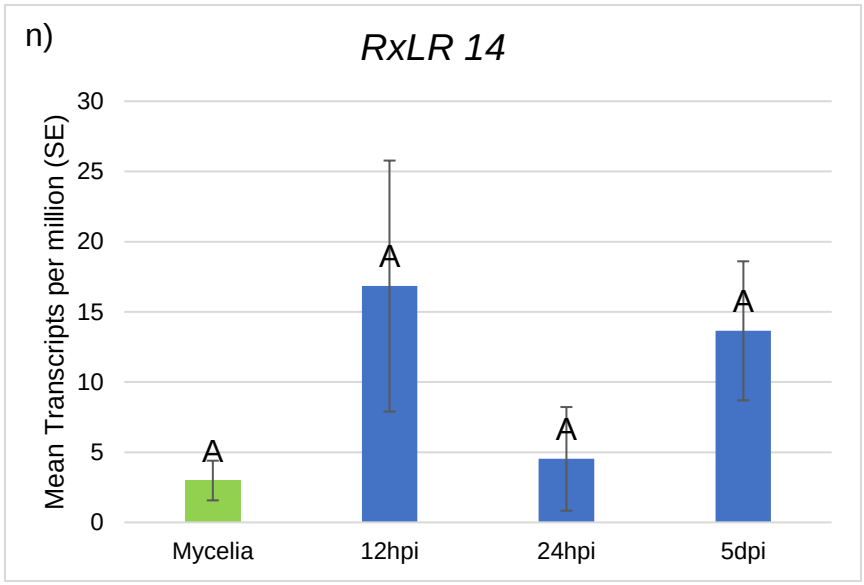

i)
RxLR 9
Mean Transcripts per million (SE)
40
35
30
25
20
15
10
5
0
A
A
A
A
Mycelia
12hpi
24hpi
5dpi

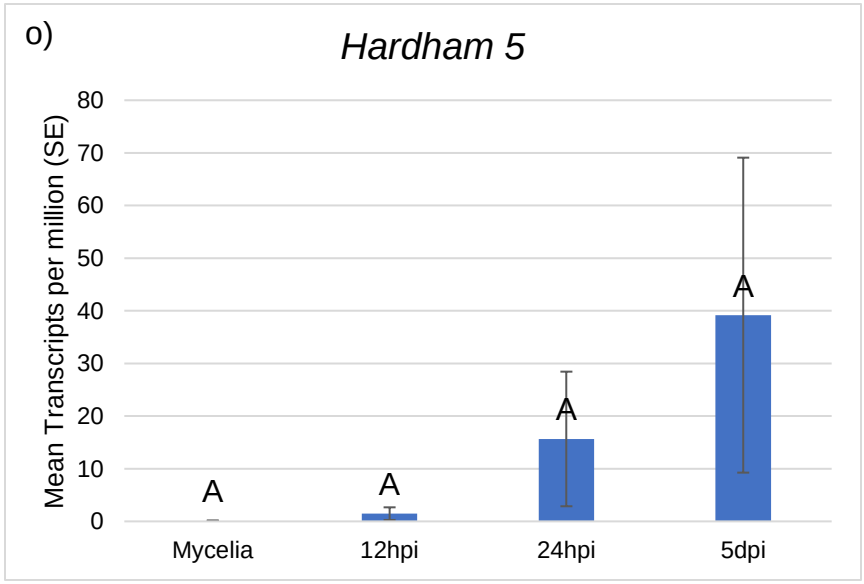

j)
RxLR 10
Mean Transcripts per million (SE)
16
14
12
10
8
6
4
2
0
AB
A
B
A
Mycelia
12hpi
24hpi
5dpi


k)
RxLR 11
Mean Transcripts per million (SE)
90
80
70
60
50
40
30
20
10
0
A
A
A
B
Mycelia
12hpi
24hpi
5dpi


l)
RxLR 12
Mean Transcripts per million (SE)
90
80
70
60
50
40
30
20
10
0
B
A
A
A
Mycelia
12hpi
24hpi
5dpi


m)
RxLR 13
Mean Transcripts per million (SE)
10
9
8
7
6
5
4
3
2
1
0
A
A
A
A
Mycelia
12hpi
24hpi
5dpi


n)
RxLR 14
Mean Transcripts per million (SE)
30
25
20
15
10
5
0
A
A
A
A
Mycelia
12hpi
24hpi
5dpi


o)
Hardham 5
Mean Transcripts per million (SE)
80
70
60
50
40
30
20
10
0
A
A
A
A
Mycelia
12hpi
24hpi
5dpi

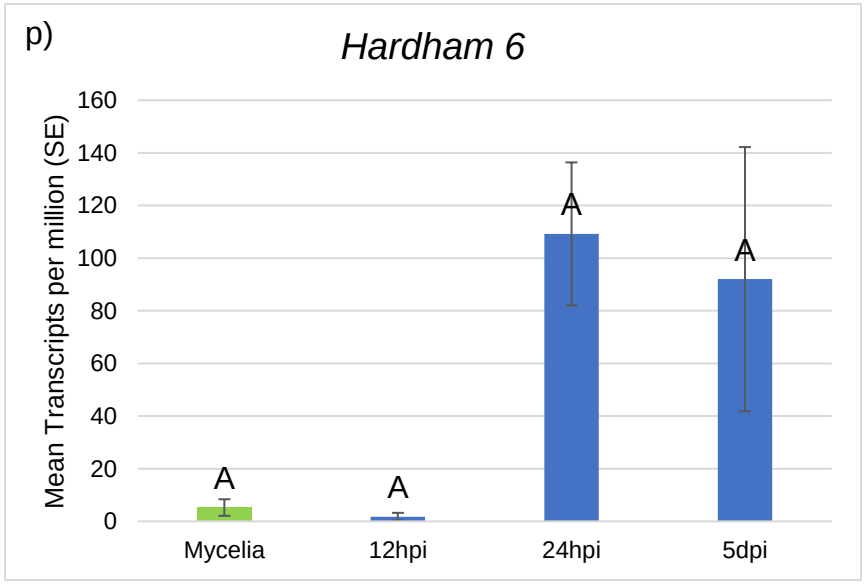

p)
Hardham 6
Mean Transcripts per million (SE)
160
140
120
100
80
60
40
20
0
A
A
A
A
Mycelia
12hpi
24hpi
5dpi

Figure 7. Charts showing the mean transcripts per million (TPM) for each candidate *Phytopthora cinnamomi RxLR* expressed in mycelia and over three timepoints during infection of the susceptible avocado rootstock R0.12. Statistical testing for the time-course was performed by a one-way ANOVA, followed by a post hoc Fisher's Least Significant Difference (LSD) test. Statistical significance is indicated above bars for mean TPM counts, and standard error bars are shown for each timepoint. Bars represented by different letters are significantly different at $p<0.05$.

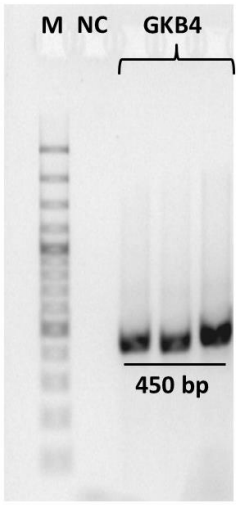


Figure 8: LPV3 PCR amplification to confirm the identity of three replicates of isolate GKB4 as *Phytophthora cinnamomi*. Lane 1 (M) contains a 100 bp molecular marker, lane 2 (NC) contains the negative PCR control, followed by three lanes for three replicates of GKB4. All three replicates showed successful amplification of the expected 450 bp length DNA region.

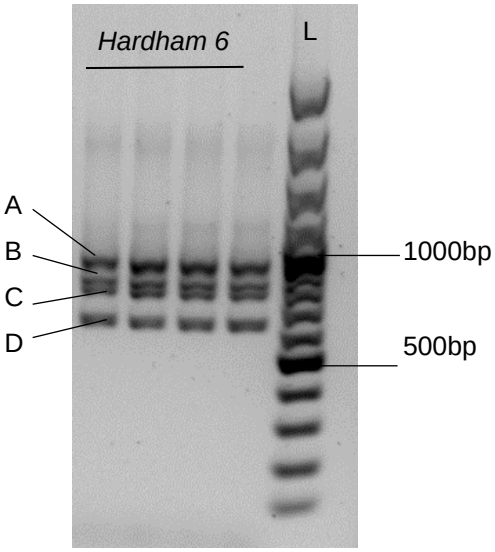


Figure 9. Amplification of *Hardham 6* for cloning and sequencing. Multiple PCR products were produced even at optimal PCR conditions (labelled **A-D**) across multiple replicates of the PCR. The two sequences closest to the desired size of 887 bp (product **B** and **C**) were extracted from the gel. The lane **L** contains 100 bp plus DNA ladder.

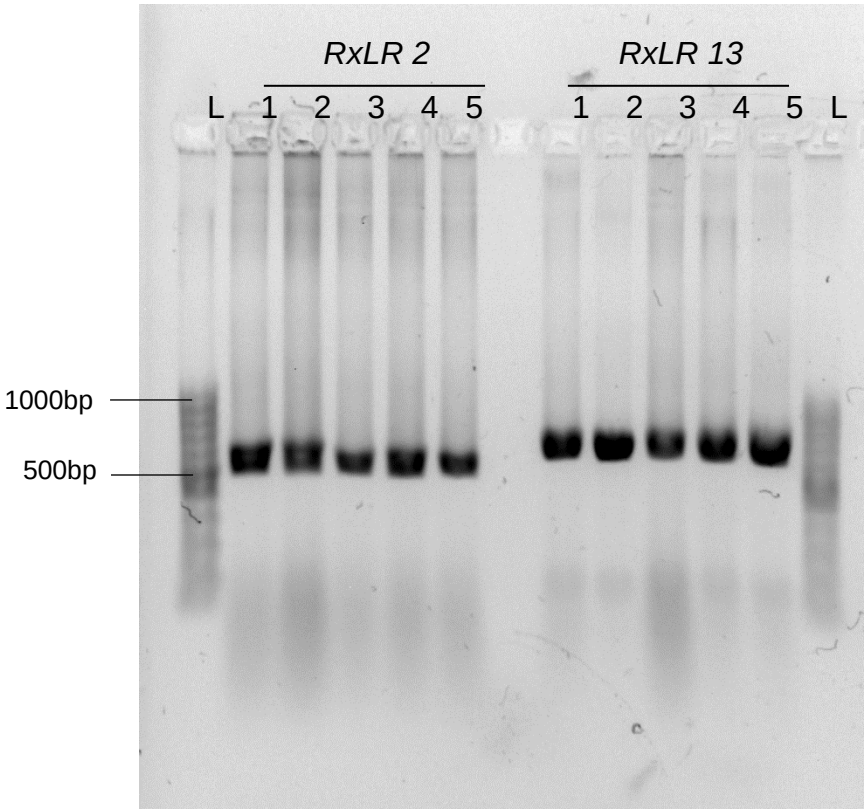


Figure 10. Amplification of cloned *RxLR* genes from successfully transformed *Escherichia coli* colonies by PCR screening using insert-specific primers. Shown in this image are the results for PCR screening of *RxLR 2* and *RxLR 13* from putative recombinant clones. Lanes 1-5 represent individual colonies screened by colony PCR. The lane L contains 100 bp DNA ladder.

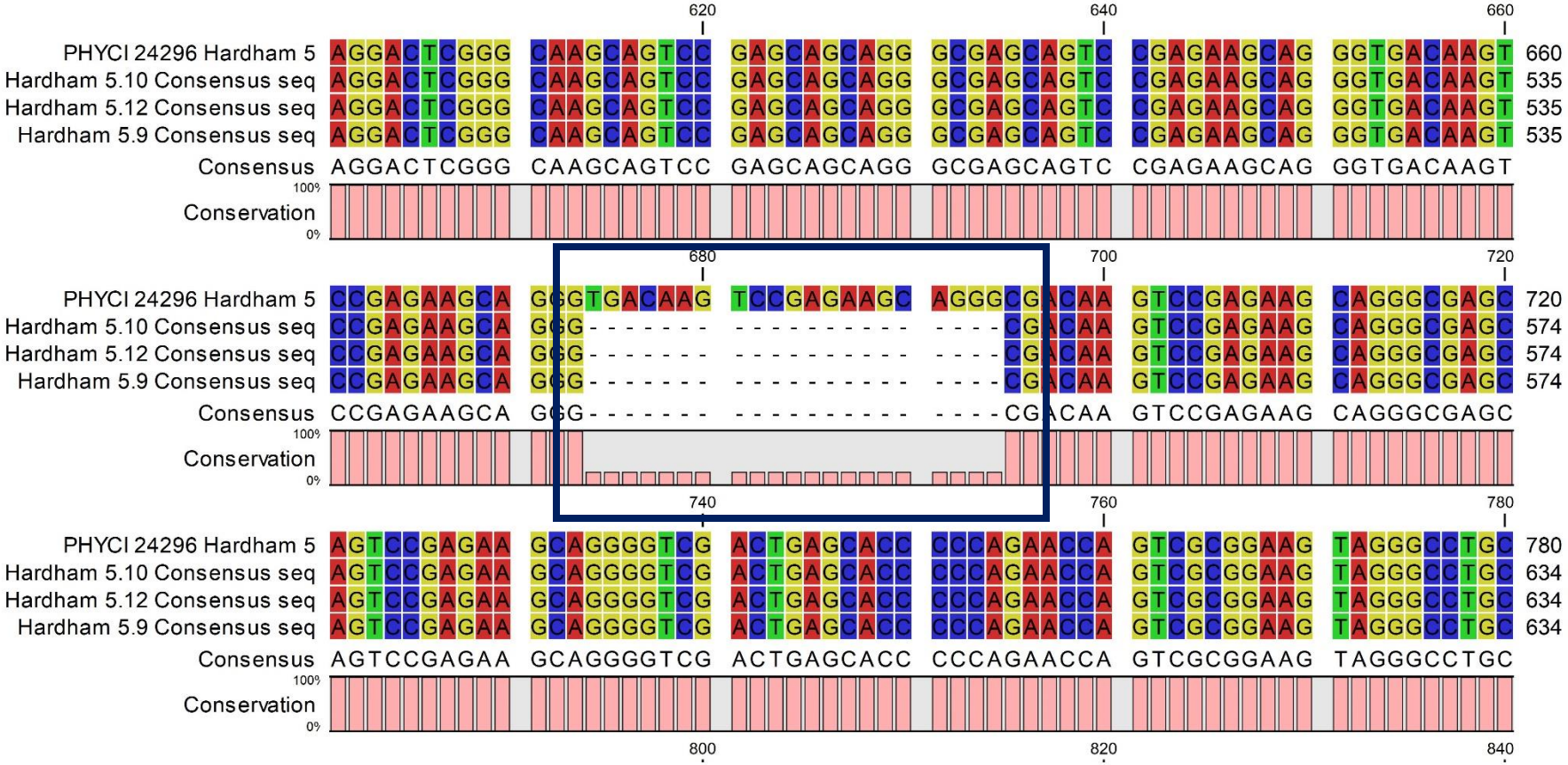


Figure 11. Snapshot of the alignment of *Hardham 5* in CLC Main Workbench 8.0.1. The alignment includes sequences obtained from three *Hardham 5* clones (*Hardham 5.9*, *5.10* and *5.12*) and the original *Hardham 5* sequence obtained from JGI (*Phyci_24296*). The image shows the high overall similarity between sequences, with the exception of 21 bases that are absent in the sequenced clones compared to the original sequence – this discrepancy is highlighted by the blue block in the figure above.

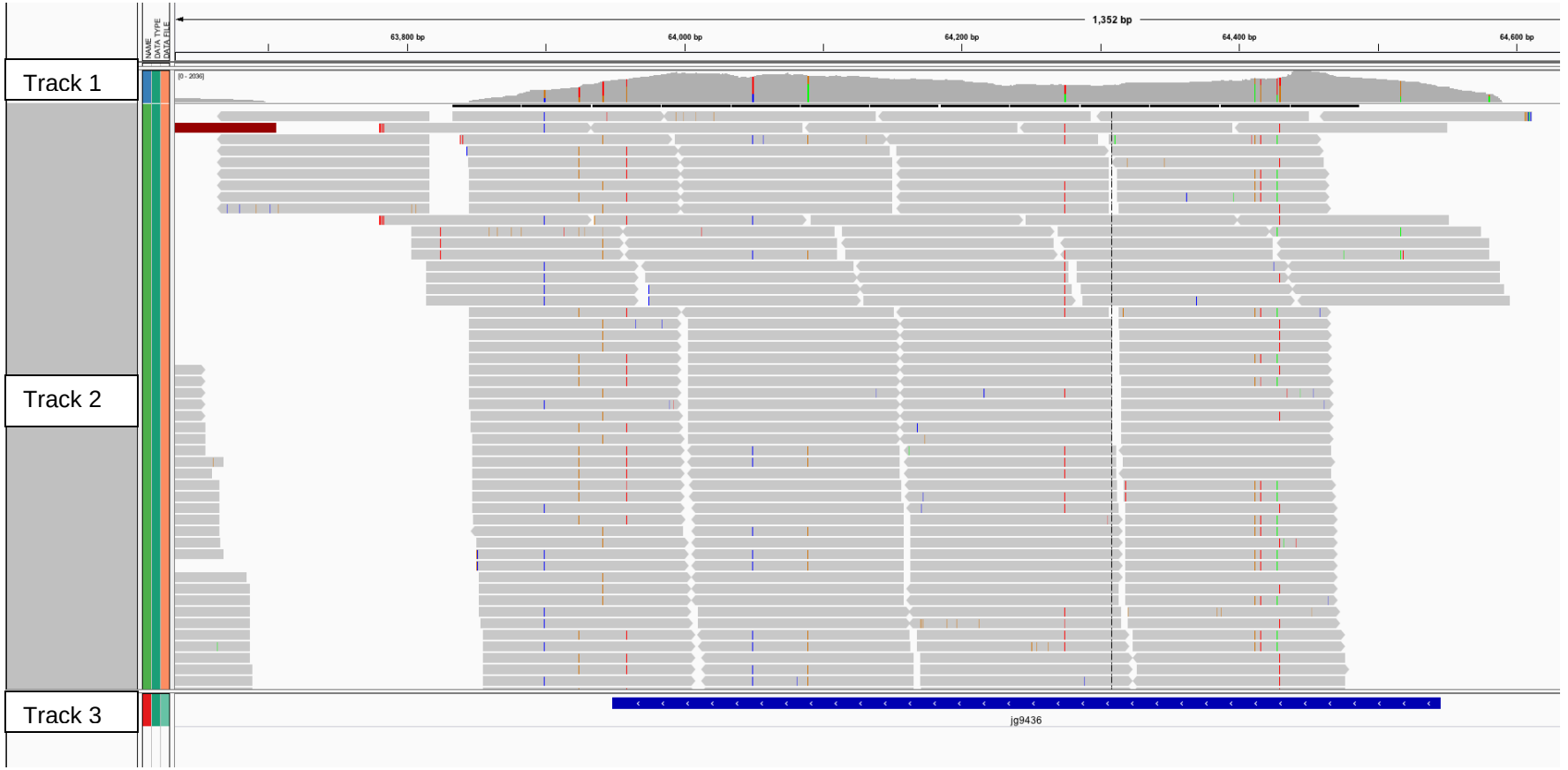


Figure 12. Snapshot of IGV results showing mapping of RNA-seq reads to the genomic location of *RxLR 1* (contig 14; 64,623-63,882). Track 1 shows the total coverage of the genomic coordinates based on the counts of RNA-seq reads mapping to the region. Track 2 illustrates visually which RNA-seq reads in the BAM file mapped to the genomic region. Track 3 shows the gene prediction for the genomic region produced by BRAKER2 – the solid bar represents the prediction of one exon for the predicted transcript – named jg9436.t1 for this region.

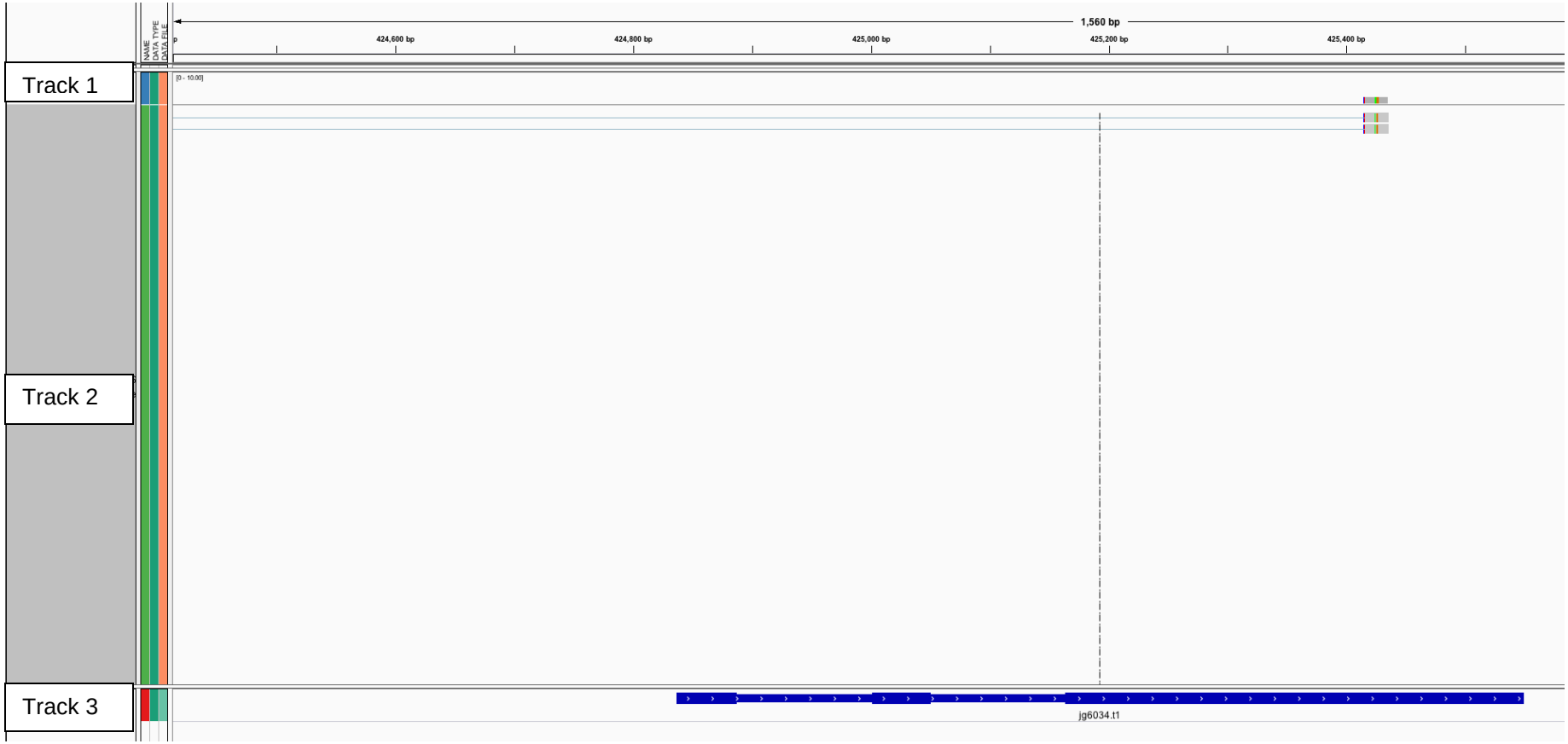


Figure 13. Snapshot of IGV results showing mapping of RNA-seq reads to the second predicted genomic location of *RxLR 2* (contig 38; 425,046-425,567). Track 1 shows the total coverage of the genomic coordinates based on the counts of RNA-seq reads mapping to the region – in this case the coverage was zero. Track 2 illustrates visually which RNA-seq reads in the BAM file mapped to the genomic region -there are no reads present in this case. Track 3 shows the gene prediction for the genomic region produced by BRAKER2 – named jg6034.t1 for this region. The absence of reads mapping to the genomic coordinates resulted in this prediction being disregarded as the genomic location for the *RxLR 2* in this study.

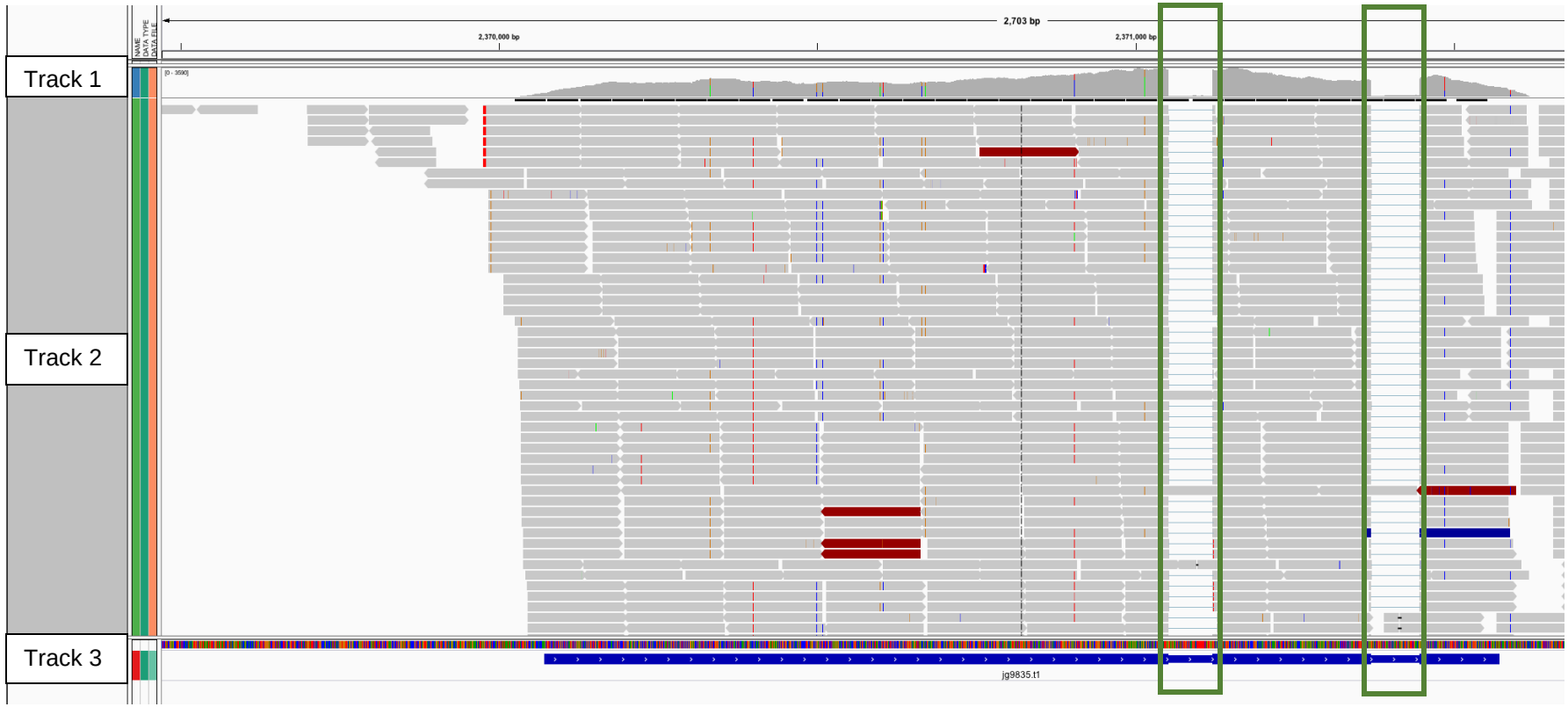


Figure 14. Snapshot of IGV results showing mapping of RNA-seq reads to the genomic location of *RxLR 7* (contig 14; 2,369,911-2,371,821). Track 1 shows the total coverage of the genomic coordinates based on the counts of RNA-seq reads mapping to the region. Track 2 illustrates visually which RNA-seq reads in the BAM file mapped to the genomic region. Track 3 shows the gene prediction for the genomic region produced by BRAKER2 – thick solid bars represent the prediction of exons and thinner bars represent introns predicted in this transcript – named jg9835.t1 for this region. Introns are highlighted in green blocks in this figure and can be observed as regions with reduced coverage of RNA-seq reads to genomic coordinates – shown as gaps in Track 1 and Track 2.

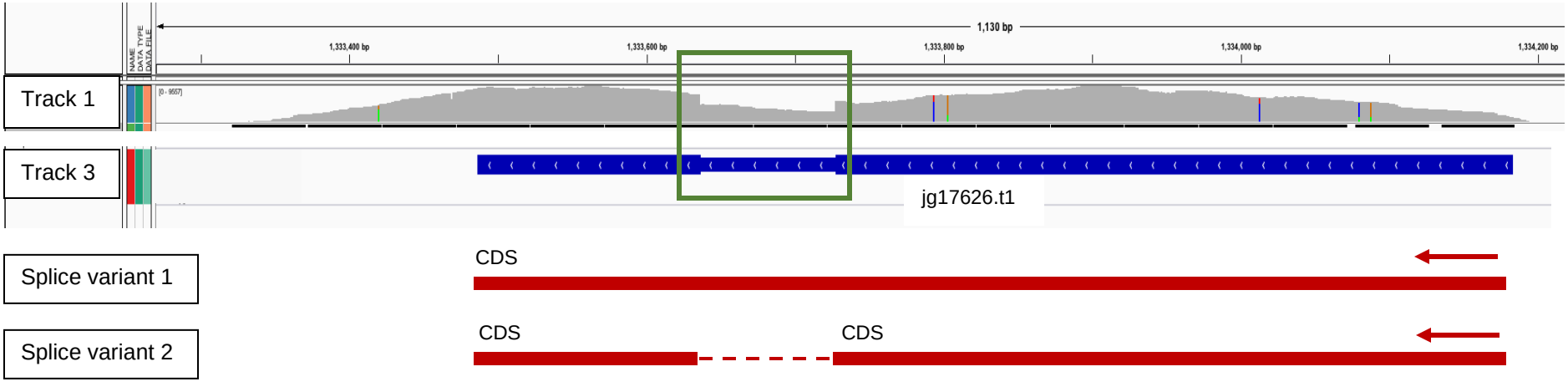


Figure 15. Snapshot of IGV results showing mapping of RNA-seq reads to the genomic location of *RxLR 4* (contig 1; 1,334,417-1,333,358). Track 1 shows the total coverage of the genomic coordinates based on the counts of RNA-seq reads mapping to the region. Track 3 shows the gene prediction for the genomic region produced by BRAKER2 – thick solid bars represent the prediction of exons and thinner bars represent the intron predicted in this transcript – named jg17626.t1 for this region. Track 2 has been removed for ease of visualisation. The intron is highlighted in the green block in this figure and can be observed as a region with reduced coverage of RNA-seq reads to genomic coordinates – shown as interruptions in Track 1, where the number of reads is decreased by approximately 50%. Interpretation of alternative RNA transcripts based on manual annotation are illustrated in red beneath the IGV snapshot, where solid bars represent the coding region for each splice variant. Directionality is indicated by red arrows above the bars.

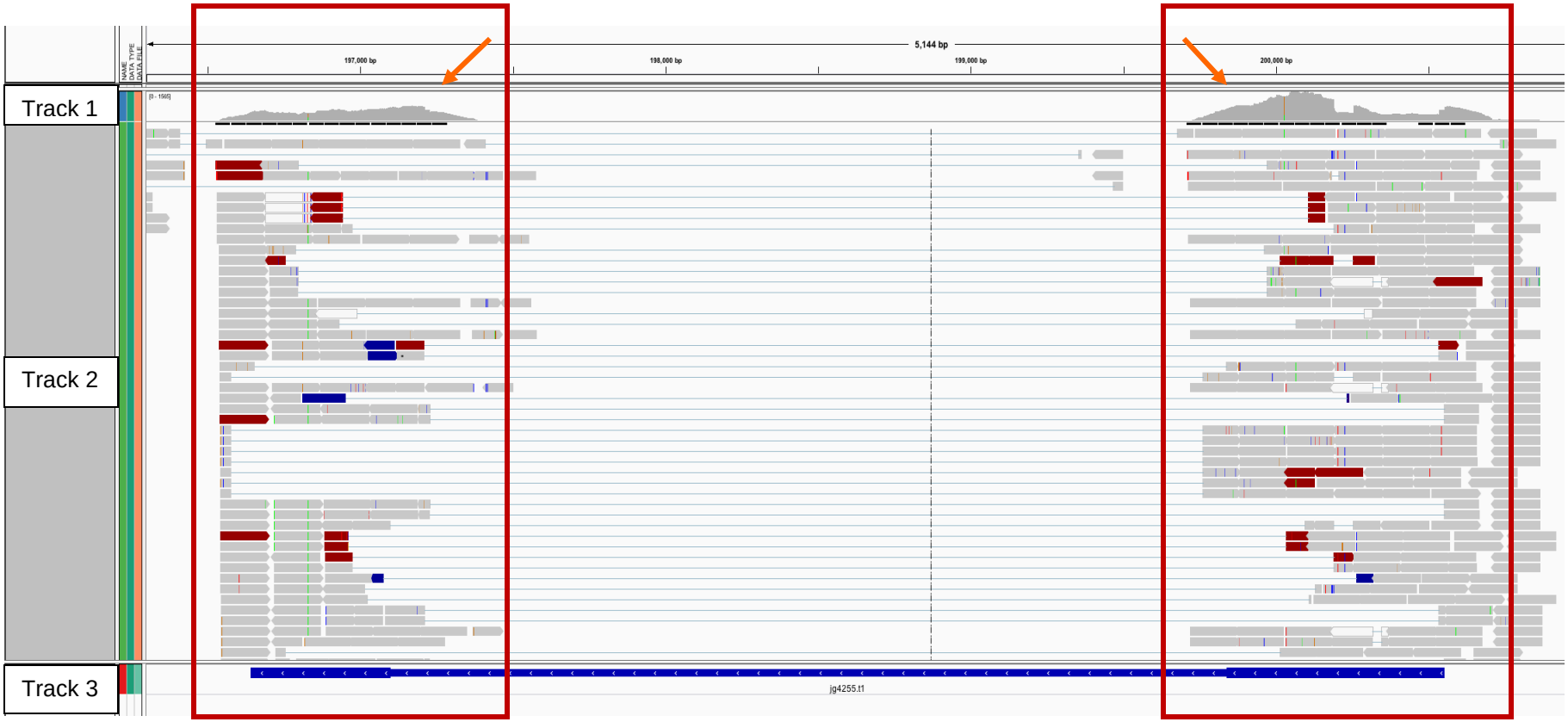


Figure 16. Snapshot of IGV results showing mapping of RNA-seq reads to the genomic location of *Hardham 6* (contig 50; 200,552 – 196,642). Track 1 shows the total coverage of the genomic coordinates based on the counts of RNA-seq reads mapping to the region. Track 2 illustrates visually which RNA-seq reads in the BAM file mapped to the genomic region. Track 3 shows the gene prediction for the genomic region produced by BRAKER2 – thick solid bars represent the prediction of exons and the thinner bar represents introns predicted in this transcript – named jg4255.t1 for this region. For this genomic region, the section predicted as exons, shown in red blocks, were manually annotated as two distinct coding regions, due to the gradually reduced read numbers at the ends of transcribed regions (indicated by orange arrows), rather than sudden interruptions that represent introns, as shown in green blocks in Figure 14 and 15.

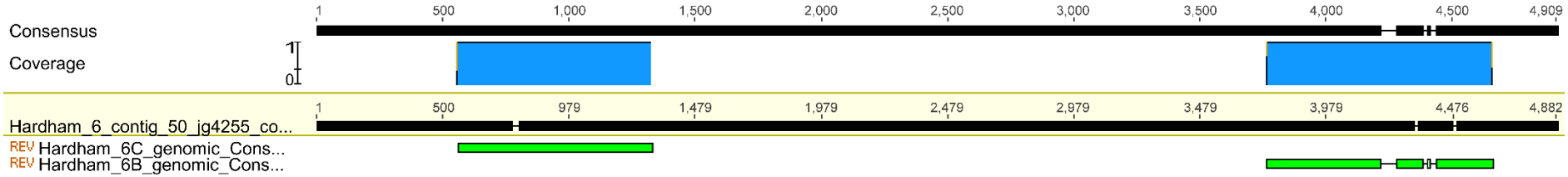


Figure 17. Assembly of genomic consensus sequences for *Hardham 6B* and *Hardham 6C* to the full consensus sequence generated for the predicted genomic coordinates for the coding region jg4255.1, illustrated for visual observation in Geneious v7.0.6. *Hardham 6C* assembled to the upstream region of the genomic consensus, while *Hardham 6B* assembled to the downstream region of the consensus sequence. Mapping of these two candidate *RxLR*s to the different sections of this region proved that they formed two distinct coding regions on the same genomic contig.

Figure 18. Visual examples of Custom BLAST hits where new peptide sequence predictions were between 95 and 100% similar to the original protein sequences predicted. Shown in this image is a snapshot of part of the BLAST hit for RxLR 2 (a) and RxLR 3 (b). Both of the new sequences are very similar to the original predictions, shown by green regions of the identity bar. In both cases, difference between the new sequences and the original predictions are limited to single amino acid differences between the proteins – indicated by the white gaps in the identity bar.

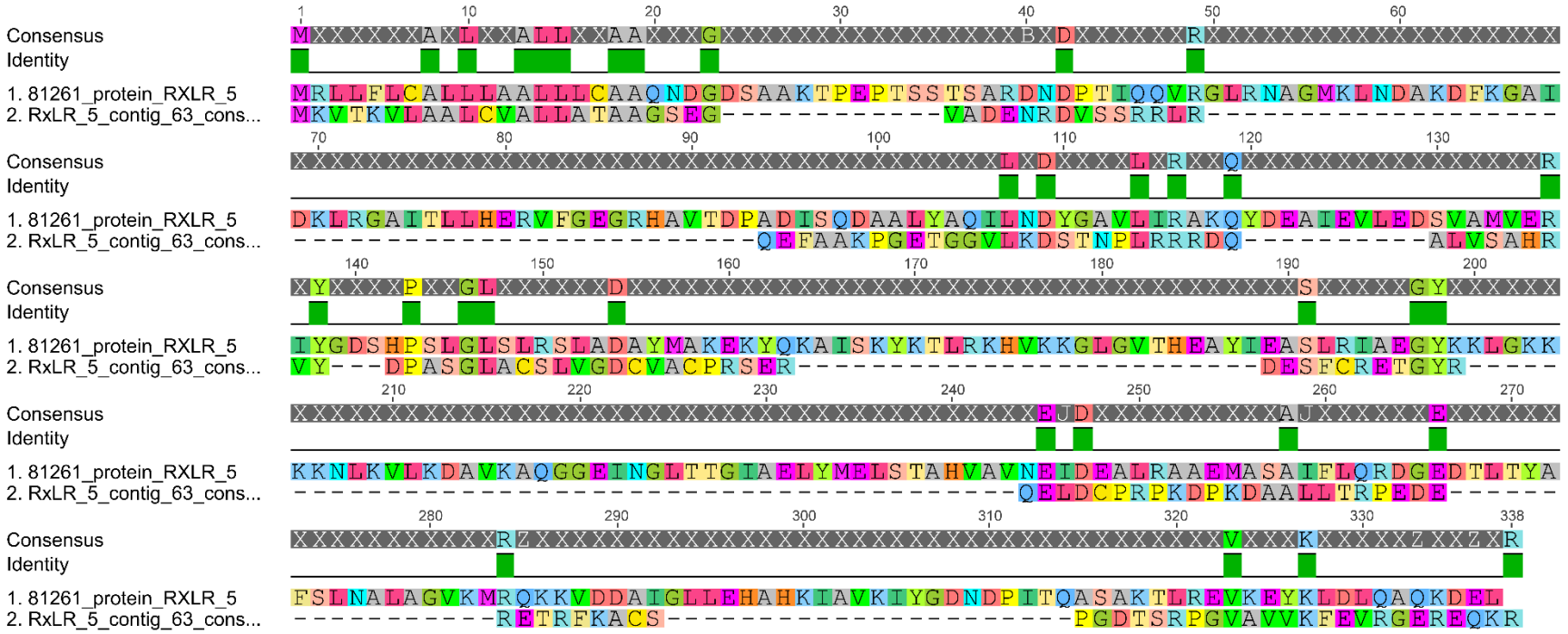
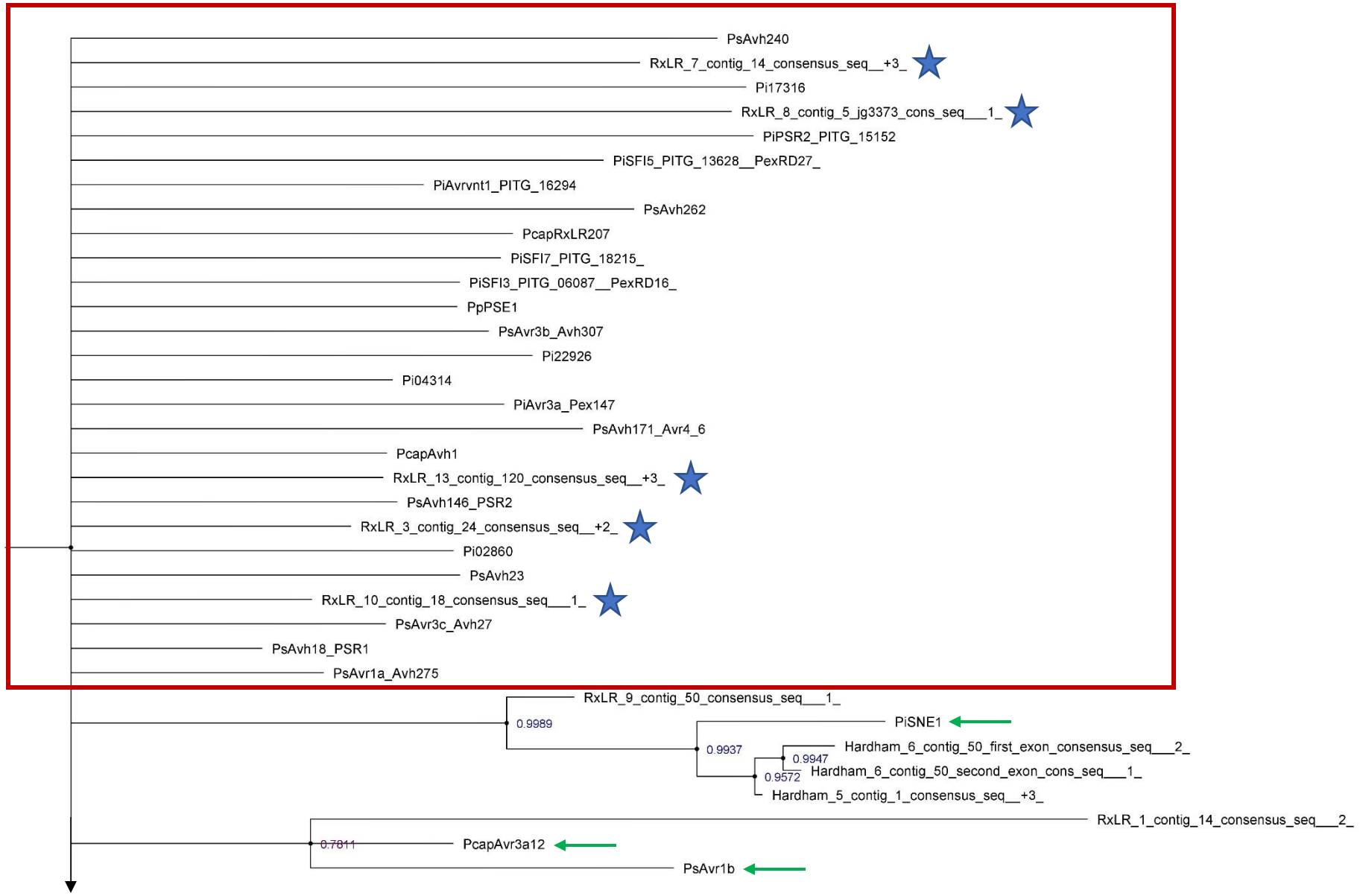


Figure 19. ClustalW alignment of the original and new protein sequences predicted for RxLR 5, as viewed in Geneious Prime 2020.0.3. Singular identical residues are visible in the alignment as blocks of green in the identity bar and residues in the consensus. No overall consensus could be inferred from the alignment, and it is clear from visual observation of the alignment that two sequences are not similar to each other.

Figure 20. ClustalW alignment of all the new predicted protein sequences that were similar to the original prediction for RxLR 4, as viewed in Geneious Prime 2020.0.3. The peptide sequences predicted for RxLR 5 (sequence 3) and the first splice variant of RxLR 4_contig 1 (sequence 2) were identical to each other, and only differed from the original RxLR 4 prediction (sequence 1) by three residues at their C-terminal ends. The peptide sequences predicted for RxLR 4_contig 28 (sequence 4) and the second splice variant of RxLR 4_contig 1 (sequence 5) were identical to each other and to the first 152 residues of the original RxLR 4 prediction (sequence 1), but were nearly unique compared to the original from residues 153-201.

Figure 21. ClustalW alignment of the original and new protein sequences predicted for Hardham 6 as viewed in Geneious Prime 2020.0.3. The alignment included the original Hardham 6 protein prediction from the Joint Genome Institute (sequence 1), the new protein sequence predicted by the first coding region of Hardham 6 (sequence 2), and the first splice variant (sequence 3) and second splice variant (sequence 4) of the second coding region found for Hardham 6. All of the sequences in the alignment were similar to one another, allowing the formulation of a reliable consensus sequence for the aligned sequences, but each sequence was on its own unique from all the others included in the alignment. None of the new predictions was an exact match to the original predicted protein sequence (sequence 1).


PsAvh240
RxLR_7_contig_14_consensus_seq__+3_
Pi17316
RxLR_8_contig_5_jg3373_cons_seq___1_
PiPSR2_PITG_15152
PiSFI5_PITG_13628__PexRD27_
PiAvrvnt1_PITG_16294
PsAvh262
PcapRxLR207
PiSFI7_PITG_18215_
PiSFI3_PITG_06087__PexRD16_
PpPSE1
PsAvr3b_Avh307
Pi22926
Pi04314
PiAvr3a_Pex147
PsAvh171_Avr4_6
PcapAvh1
RxLR_13_contig_120_consensus_seq__+3_
PsAvh146_PSR2
RxLR_3_contig_24_consensus_seq__+2_
Pi02860
PsAvh23
RxLR_10_contig_18_consensus_seq___1_
PsAvr3c_Avh27
PsAvh18_PSR1
PsAvr1a_Avh275
RxLR_9_contig_50_consensus_seq___1_
0.9989
PiSNE1
0.9937
Hardham_6_contig_50_first_exon_consensus_seq___2_
0.9947
0.9572
Hardham_6_contig_50_second_exon_cons_seq___1_
Hardham_5_contig_1_consensus_seq__+3_
RxLR_1_contig_14_consensus_seq___2_
0.7811
PcapAvr3a12
PsAvr1b

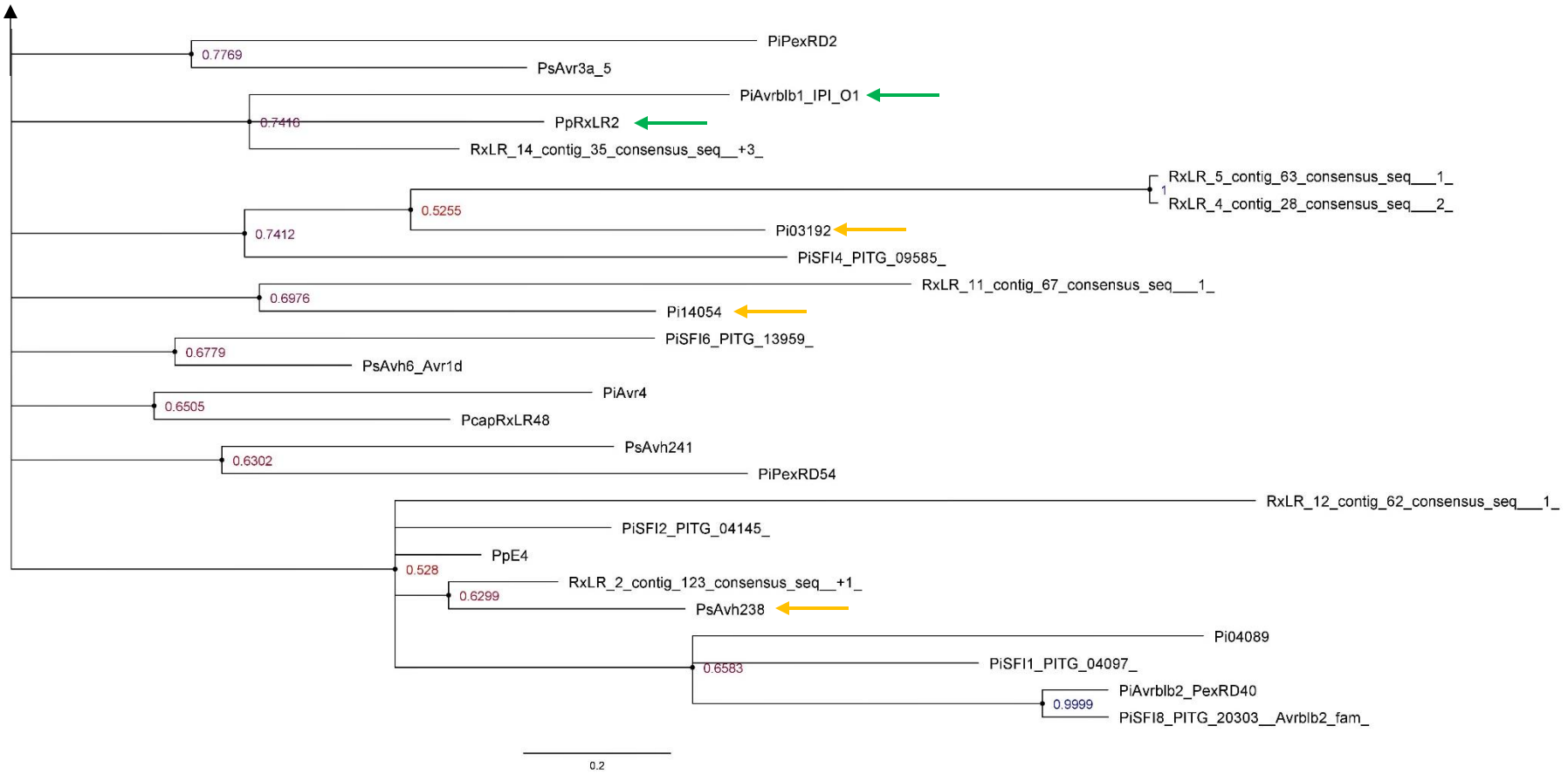


Figure 22. Phylogenetic tree resulting from Bayesian inference analysis of the predicted N-terminal regions of the *Phytophthora cinnamomi* candidate RxLR effectors aligned with the N-terminal regions of functionally characterised RxLRs in other species. Branch tips of the tree are labelled with the protein names given to *P. cinnamomi* candidate effectors in this study, or the names of confirmed RxLRs in other species, as obtained from literature. Posterior probability values are shown up to the fourth significant digit for each node, coloured according to the probability value. Probability value colours fall on a spectrum ranging from red (lowest probability values) to blue (highest probability values). The sections indicated

by red blocks represent sequences that were not inferred to be evolutionarily related to any other sequences and can all be perceived as outgroups. Proteins indicated by green arrows are characterised RxLRs from other species which grouped with candidate *P. cinnamomi* RxLRs with a posterior probability above 0.7. Effector names indicated by yellow arrows are characterised RxLRs from other species which were shown to be more distantly related to *P. cinnamomi* RxLRs, with a probability of less than 0.7 for these groupings.

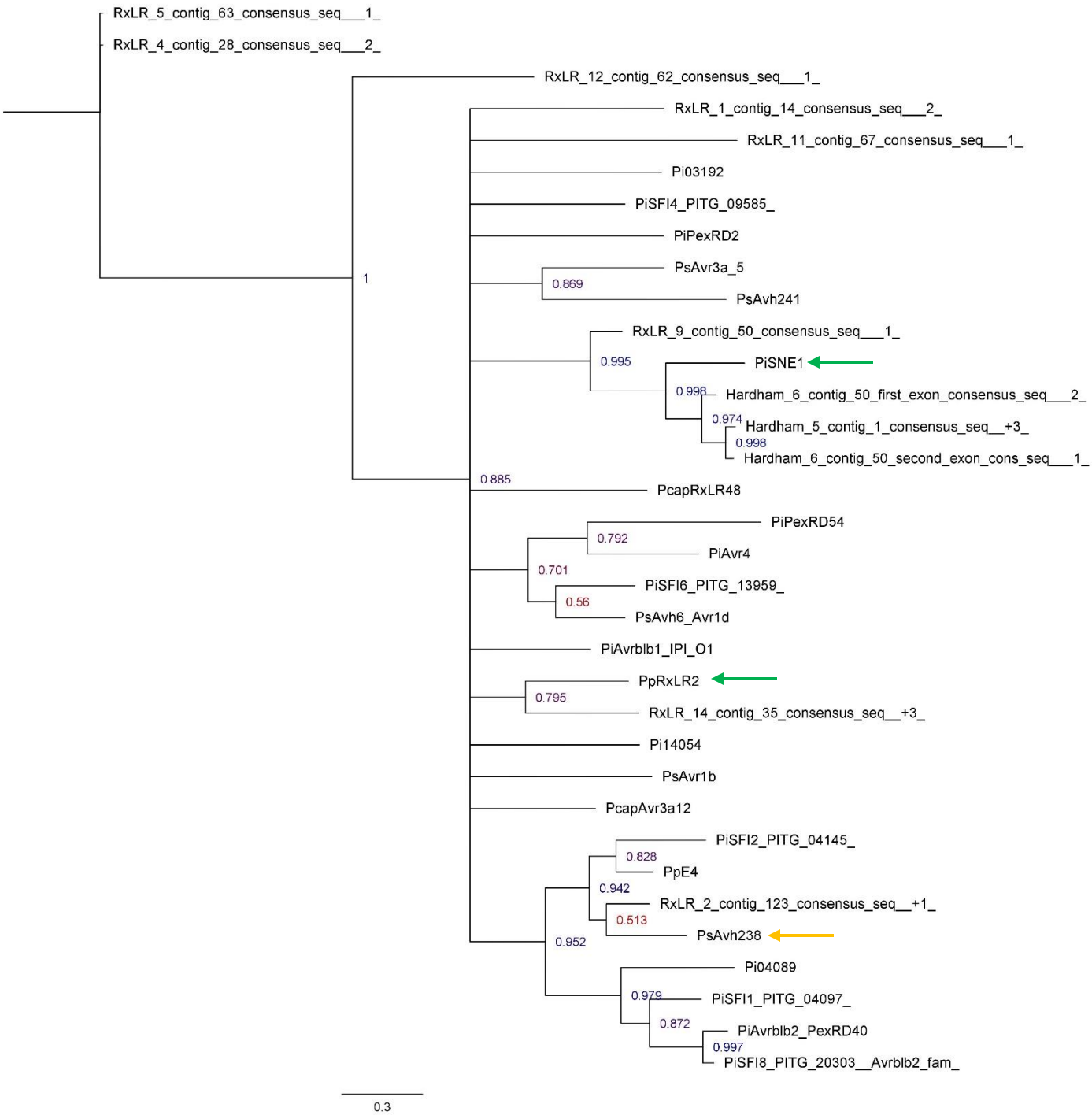
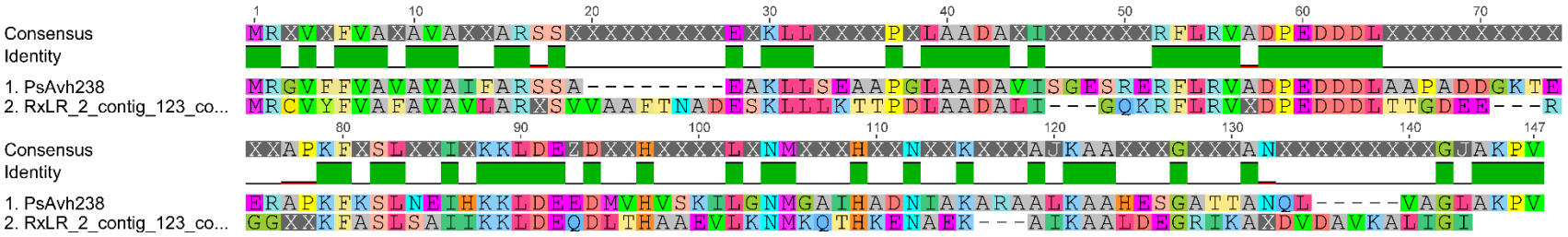


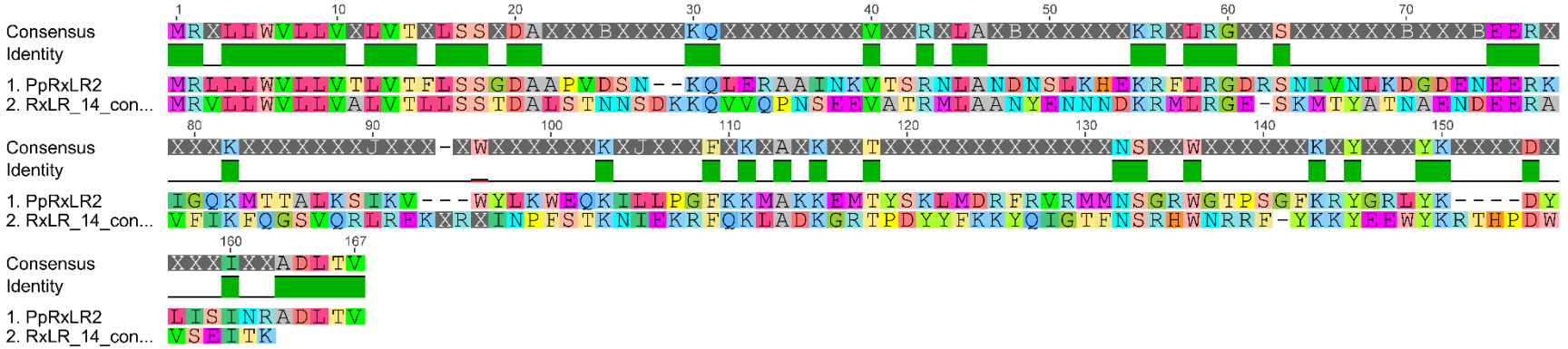
Figure 23. Phylogenetic tree resulting from Bayesian inference analysis of the predicted N-terminal regions of the *Phytophthora cinnamomi* candidate RxLR effectors aligned with the N-terminal regions of functionally characterised RxLRs in other species, with uninformative sequences from previous phylogenetic analysis removed. Branch tips of the tree are labelled with the protein names given to *P. cinnamomi* candidate effectors in

this study, or the names of confirmed RxLRs in other species, as obtained from literature. Posterior probability values are shown up to the third significant digit for each node, coloured according to the probability value. Probability value colours fall on a spectrum ranging from red (lowest probability values) to blue (highest probability values). Proteins indicated by green arrows are characterised RxLRs from other species which grouped with candidate *P. cinnamomi* RxLRs with a posterior probability above 0.7. The effector indicated by the yellow arrow is a characterised RxLR from another species which was shown to be more distantly related to a *P. cinnamomi* RxLR, with a probability of less than 0.7 for this grouping.

a)


Consensus
Identity
1. PsAvh238
2. RxLR_2_contig_123_co...

b)


Consensus
Identity
1. PpRxLR2
2. RxLR_14_con...

c)

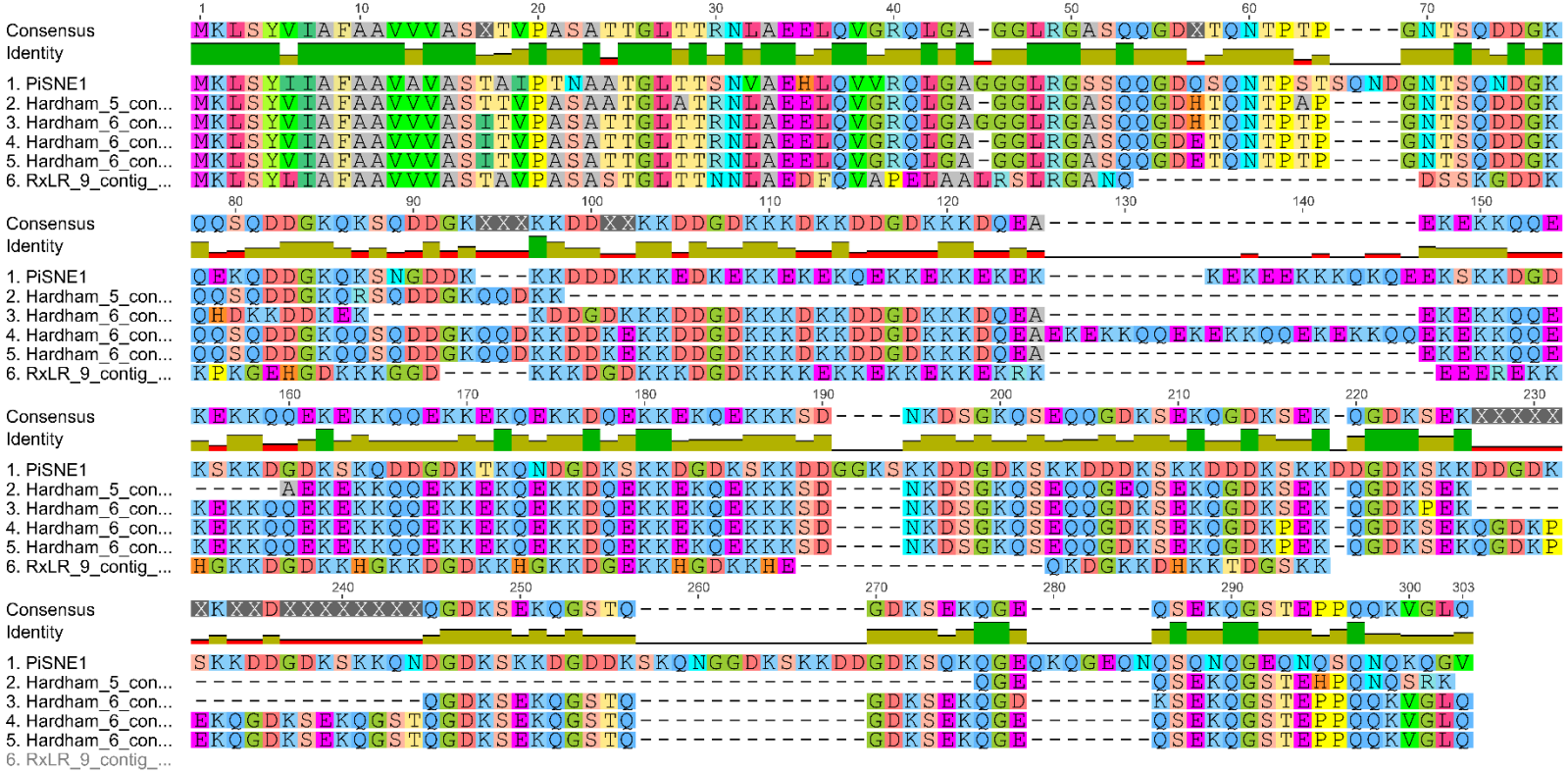


Figure 24. MUSCLE alignment of the protein sequences which grouped within their own clades in phylogenetic analysis. Alignments were viewed in Geneious v7.06. a) The alignment of *Phytophthora cinnamomi* RxLR 2 with *Phytophthora sojae* Avh238 (**a**) and the alignment of *P. cinnamomi* RxLR 14 and *Phytophthora parasitica* RxLR 2 (**b**) showed some overall similarity, with several short regions of consensus (green regions of the identity bar) separated by short regions of divergent residues (white gaps in the identity bar). The alignment of RxLR 9, Hardham 5 and Hardham 6 protein sequences with *Phytophthora infestans* PiSNE1 (**c**) had the highest overall similarity, with several regions in their C-terminal domains being conserved across the sequences allowing the formulation of a reliable consensus sequence for the a large portion of the aligned sequences.

References

Ali SS, Shao J, Lary DJ, Kronmiller BA, Shen D, Strem MD, Amoako-Attah I, Akrofi AY, Begoude BD & ten Hoopen GM (2017) *Phytophthora megakarya* and *Phytophthora palmivora*, closely related causal agents of Cacao Black Pod Rot, underwent increases in genome sizes and gene numbers by different mechanisms. *Genome Biology and Evolution* **9**: 536-557.

Armstrong MR, Whisson SC, Pritchard L*, et al.* (2005) An ancestral oomycete locus contains late blight avirulence gene *Avr3a,* encoding a protein that is recognized in the host cytoplasm. *Proceedings of the National Academy of Sciences* **102**: 7766-7771.

Bailey K, Çevik V, Holton N, Byrne-Richardson J, Sohn KH, Coates M, Woods-Tör A, Aksoy HM, Hughes L & Baxter L (2011) Molecular cloning of ATR5^{Emoy2} from *Hyaloperonospora arabidopsidis*, an avirulence determinant that triggers RPP5-mediated defense in *Arabidopsis*. *Molecular Plant-Microbe Interactions* **24**: 827-838.

Belisle RJ, McKee B, Hao W, Crowley M, Arpaia ML, Miles TD, Adaskaveg JE & Manosalva P (2019) Phenotypic Characterization of Genetically Distinct *Phytophthora cinnamomi* Isolates from Avocado. *Phytopathology* **109**: 384-394.

Boevink PC, Wang X, McLellan H*, et al.* (2016) A *Phytophthora infestans* RXLR effector targets plant PP1c isoforms that promote late blight disease. *Nature Communications* **7**: 10311.

Bolger AM, Lohse M & Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.

Bos JI, Armstrong MR, Gilroy EM, Boevink PC, Hein I, Taylor RM, Zhendong T, Engelhardt S, Vetukuri RR & Harrower B (2010) *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proceedings of the National Academy of Sciences* **107**: 9909-9914.

Bouwmeester K, De Sain M, Weide R, Gouget A, Klamer S, Canut H & Govers F (2011) The lectin receptor kinase LecRK-I. 9 is a novel *Phytophthora* resistance component and a potential host target for a RXLR effector. *PLoS Pathogens* **7**: e1001327.

Bozkurt TO, Schornack S, Banfield MJ & Kamoun S (2012) Oomycetes, effectors, and all that jazz. *Current Opinion in Plant Biology* **15**: 483-492.

Bozkurt TO, Schornack S, Win J, Shindo T, Ilyas M, Oliva R, Cano LM, Jones AM, Huitema E & van der Hoorn RA (2011) *Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. *Proceedings of the National Academy of Sciences* **108**: 20832-20837.

Bray NL, Pimentel H, Melsted P & Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**: 525-527.

Chang S, Puryear J & Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* **11**: 113-116.

Chen X-R, Zhang Y, Li H, Zhang Z-H, Sheng G-L, Li Y-P, Xing Y-P, Huang S-X, Tao H & Kuan T (2019) The RXLR effector PcAvh1 is required for full virulence of *Phytophthora capsici*. *Molecular Plant-Microbe Interactions* **32**: 986-1000.

Chen Y, Liu Z & Halterman DA (2012) Molecular determinants of resistance activation and suppression by *Phytophthora infestans* effector IPI-O. *PLoS Pathogens* **8**: e1002595.

Cheng B, Yu X, Ma Z, Dong S, Dou D, Wang Y & Zheng X (2012) *Phytophthora sojae* effector Avh331 suppresses the plant defence response by disturbing the MAPK signalling pathway. *Physiological and Molecular Plant Pathology* **77**: 1-9.

Chou S, Krasileva KV, Holton JM, Steinbrenner AD, Alber T & Staskawicz BJ (2011) *Hyaloperonospora arabidopsidis* ATR1 effector is a repeat protein with distributed recognition surfaces. *Proceedings of the National Academy of Sciences* **108**: 13323-13328.

Coffey DM (1987) Phytophthora root rot of avocado - an integrated approach to control in California. *California Avocado Society Yearbook* **71**: 121-137.

Consortium U (2014) UniProt: a hub for protein information. *Nucleic Acids Research* **43**: D204-D212.

Cook DE, Mesarich CH & Thomma BP (2015) Understanding plant immunity as a surveillance system to detect invasion. *Annual Reviews in Phytopathology* **53**: 541-563.

Dagdas YF, Belhaj K, Maqbool A*, et al.* (2016) An effector of the Irish potato famine pathogen antagonizes a host autophagy cargo receptor. *eLife* **5**.

Dalio R, Maximo H, Oliveira T, Dias R, Breton M, Felizatti H & Machado M (2018) *Phytophthora parasitica* effector PpRxLR2 suppresses *Nicotiana benthamiana* immunity. *Molecular Plant-Microbe Interactions* **31**: 481-493.

de Vries S, von Dahlen JK, Uhlmann C, Schnake A, Kloesges T & Rose LE (2017) Signatures of selection and host-adapted gene expression of the *Phytophthora infestans* RNA silencing suppressor PSR2. *Molecular Plant Pathology* **18**: 110-124.

Dong S, Qutob D, Tedman-Jones J, Kuflu K, Wang Y, Tyler BM & Gijzen M (2009) The *Phytophthora sojae* avirulence locus *Avr3c* encodes a multi-copy RXLR effector with sequence polymorphisms among pathogen strains. *PLoS One* **4**: e5556.

Dong S, Yu D, Cui L, Qutob D, Tedman-Jones J, Kale SD, Tyler BM, Wang Y & Gijzen M (2011a) Sequence variants of the *Phytophthora sojae* RXLR effector Avr3a/5 are differentially recognized by Rps3a and Rps5 in soybean. *PloS One* **6**: 1-8.

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

Dou D & Zhou JM (2012) Phytopathogen effectors subverting host immunity: different foes, similar battleground. *Cell Host & Microbe* **12**: 484-495.

Dou D, Kale SD, Wang X, Jiang RH, Bruce NA, Arredondo FD, Zhang X & Tyler BM (2008a) RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. *Plant Cell* **20**: 1930-1947.

Dou D, Kale SD, Liu T, Tang Q, Wang X, Arredondo FD, Basnayake S, Whisson S, Drenth A & Maclean D (2010) Different domains of *Phytophthora sojae* effector Avr4/6 are recognized by soybean resistance genes *Rps 4* and *Rps 6*. *Molecular Plant-Microbe Interactions* **23**: 425-435.

Dou D, Kale SD, Wang X*, et al.* (2008b) Conserved C-terminal motifs required for avirulence and suppression of cell death by *Phytophthora sojae* effector Avr1b. *Plant Cell* **20**: 1118-1133.

Du Y, Mpina MH, Birch PR, Bouwmeester K & Govers F (2015) *Phytophthora infestans* RXLR effector AVR1 interacts with exocyst component Sec5 to manipulate plant immunity. *Plant Physiology* **169**: 1975-1990.

Duarte JM, Cui L, Wall PK, Zhang Q, Zhang X, Leebens-Mack J, Ma H, Altman N & DePamphilis CW (2005) Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis*. *Molecular Biology and Evolution* **23**: 469-478.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792-1797.

Evangelisti E, Govetto B, Minet-Kebdani N, Kuhn ML, Attard A, Ponchet M, Panabières F & Gourgues M (2013) The *Phytophthora parasitica* RXLR effector Penetration-Specific Effector 1 favours *Arabidopsis thaliana* infection by interfering with auxin physiology. *New Phytologist* **199**: 476-489.

Fabro G, Steinbrenner J, Coates M, Ishaque N, Baxter L, Studholme DJ, Körner E, Allen RL, Piquerez SJ & Rougon-Cardoso A (2011) Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host plant immunity. *PLoS Pathogens* **7**: e1002348.

Fan G, Yang Y, Li T, Lu W, Du Y, Qiang X, Wen Q & Shan W (2018) A *Phytophthora capsici* RXLR effector targets and inhibits a plant PPIase to suppress endoplasmic reticulum-mediated immunity. *Molecular Plant* **11**: 1067-1083.

Fang Y & Tyler BM (2016) Efficient disruption and replacement of an effector gene in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Molecular Plant Pathology* **17**: 127-139.

Gilroy EM, Taylor RM, Hein I, Boevink P, Sadanandom A & Birch PR (2011b) CMPG1-dependent cell death follows perception of diverse pathogen elicitors at the host plasma membrane and is suppressed by *Phytophthora infestans* RXLR effector AVR3a. *New Phytologist* **190**: 653-666.

Goritschnig S, Krasileva KV, Dahlbeck D & Staskawicz BJ (2012) Computational prediction and molecular characterization of an oomycete effector and the cognate Arabidopsis resistance gene. *PLOS Genetics* **8**: e1002502.

Goss EM, Press CM & Grünwald NJ (2013) Evolution of RXLR-class effectors in the oomycete plant pathogen *Phytophthora ramorum*. *PLoS One* **8**: e79347.

Guo B, Wang H, Yang B, Jiang W, Jing M, Li H, Xia Y, Xu Y, Hu Q & Wang F (2019) *Phytophthora sojae* effector PsAvh240 inhibits host aspartic protease secretion to promote infection. *Molecular Plant* **12**: 552-564.

Haas BJ, Kamoun S, Zody MC*, et al.* (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* **461**: 393-398.

Hahn MW (2009) Distinguishing among evolutionary models for the maintenance of gene duplicates. *Journal of Heredity* **100**: 605-617.

Hardham AR (2005) *Phytophthora cinnamomi*. *Molecular Plant Pathology* **6**: 589-604.

Hardham AR & Blackman LM (2017) *Phytophthora cinnamomi*. *Molecular Plant Pathology* **19**: 260-285.

Hardy GSJ, Colquhoun IJ, Shearer BL & Tommerup I (2001) The impact and control of *Phytophthora cinnamomi* in native and rehabilitated forest ecosystems in Western Australia. *Forest Snow and Landscape Research* **76**: 337-343.

He Q, McLellan H, Hughes RK, Boevink PC, Armstrong M, Lu Y, Banfield MJ, Tian Z & Birch PR (2019) *Phytophthora infestans* effector SFI 3 targets potato UBK to suppress early immune transcriptional responses. *New Phytologist* **222**: 438-454.

He X & Zhang J (2005) Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics* **169**: 1157-1164.

Hein I, Gilroy EM, Armstrong MR & Birch PR (2009) The zig-zag-zig in oomycete-plant interactions. *Molecular Plant Pathology* **10**: 547-562.

Hoff KJ, Lomsadze A, Borodovsky M & Stanke M (2019) Whole-Genome Annotation with BRAKER. *Gene Prediction,*p.^pp. 65-95. Springer.

Huang G, Liu Z, Gu B, Zhao H, Jia J, Fan G, Meng Y, Du Y & Shan W (2019) An RXLR effector secreted by *Phytophthora parasitica* is a virulence factor and triggers cell death in various plants. *Molecular Plant Pathology* **20**: 356-371.

Huang J, Gu L, Zhang Y, Yan T, Kong G, Kong L, Guo B, Qiu M, Wang Y & Jing M (2017) An oomycete plant pathogen reprograms host pre-mRNA splicing to subvert immunity. *Nature Communications* **8**: 2051.

Jiang RH & Tyler BM (2012) Mechanisms and evolution of virulence in oomycetes. *Annual Reviews in Phytopathology* **50**: 295-318.

Jiang RH, Tripathy S, Govers F & Tyler BM (2008) RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the Natural Academy of Sciences* **105**: 4874-4879.

Jing M, Guo B, Li H*, et al.* (2016) A *Phytophthora sojae* effector suppresses endoplasmic reticulum stress-mediated immunity by stabilizing plant Binding immunoglobulin Proteins. *Nature Communications* **7**: 11685.

Jones JDG & Dangl JL (2006) The plant immune system. *Nature* **444**: 323-329.

Kale SD & Tyler BM (2011) Entry of oomycete and fungal effectors into plant and animal host cells. *Cellular Microbiology* **13**: 1839-1848.

Kale SD, Gu B, Capelluto DG, Dou D, Feldman E, Rumore A, Arredondo FD, Hanlon R, Fudal I & Rouxel T (2010) External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell* **142**: 284-295.

Kamoun S (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Phytopathology* **44**: 41.

Kelley BS, Lee SJ, Damasceno CM, Chakravarthy S, Kim BD, Martin GB & Rose JK (2010) A secreted effector protein (SNE1) from *Phytophthora infestans* is a broadly acting suppressor of programmed cell death. *The Plant Journal* **62**: 357-366.

King SR, McLellan H, Boevink PC, Armstrong MR, Bukharova T, Sukarta O, Win J, Kamoun S, Birch PR & Banfield MJ (2014) *Phytophthora infestans* RXLR effector PexRD2 interacts with host MAPKKKε to suppress plant immune signaling. *The Plant Cell* **26**: 1345-1359.

Kong L, Qiu X, Kang J*, et al.* (2017) A *Phytophthora* Effector Manipulates Host Histone Acetylation and Reprograms Defense Gene Expression to Promote Infection. *Current Biology* **27**: 981-991.

Leonelli L, Pelton J, Schoeffler A, Dahlbeck D, Berger J, Wemmer DE & Staskawicz B (2011) Structural elucidation and functional characterization of the *Hyaloperonospora arabidopsidis* effector protein ATR13. *PLoS Pathogens* **7**: e1002428

.

Li Q, Chen Y, Wang J, Zou F, Jia Y, Shen D, Zhang Q, Jing M, Dou D & Zhang M (2019a) A *Phytophthora capsici* virulence effector associates with NPR1 and suppresses plant immune responses. *Phytopathology Research* **1**.

Li Q, Ai G, Shen D, Zou F, Wang J, Bai T, Chen Y, Li S, Zhang M & Jing M (2019b) A *Phytophthora capsici* effector targets ACD11 binding partners that regulate ROS-mediated defense response in *Arabidopsis*. *Molecular Plant* **12**: 565-581.

McGowan J & Fitzpatrick DA (2017) Genomic, network, and phylogenetic analysis of the oomycete effector arsenal. *MSphere* **2**: e00408-00417.

McLellan H, Boevink PC, Armstrong MR, Pritchard L, Gomez S, Morales J, Whisson SC, Beynon JL & Birch PR (2013) An RxLR effector from *Phytophthora infestans* prevents re-localisation of two plant NAC transcription factors from the endoplasmic reticulum to the nucleus. *PLoS Pathogens* **9**: e1003670.

Mengiste T (2012) Plant immunity to necrotrophs. *Annual Review of Phytopathology* **50**: 267-294.

Morgan W & Kamoun S (2007) RXLR effectors of plant pathogenic oomycetes. *Current Opinion in Microbiology* **10**: 332-338.

Murphy F, He Q, Armstrong M, Giuliani LM, Boevink PC, Zhang W, Tian Z, Birch PR & Gilroy EM (2018) The potato MAP3K StVIK is required for the *Phytophthora infestans* RXLR effector Pi17316 to promote disease. *Plant Physiology* **177**: 398-410.

Na R, Yu D, Qutob D, Zhao J & Gijzen M (2013) Deletion of the *Phytophthora sojae* avirulence gene *Avr1d* causes gain of virulence on *Rps*1d. *Molecular Plant-Microbe Interactions* **26**: 969-976.

Na R, Yu D, Chapman BP, Zhang Y, Kuflu K, Austin R, Qutob D, Zhao J, Wang Y & Gijzen M (2014) Genome re-sequencing and functional analysis places the *Phytophthora sojae* avirulence genes *Avr1c* and *Avr1a* in a tandem repeat at a single locus. *PloS ONE* **9**: e89738.

Pimentel H, Bray NL, Puente S, Melsted P & Pachter L (2017) Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods* **14**: 687-690.

Qiao X, Yin H, Li L, Wang R, Wu J, Wu J & Zhang S (2018) Different modes of gene duplication show divergent evolutionary patterns and contribute differently to the expansion of gene families involved in important fruit traits in pear (*Pyrus bretschneideri*). *Frontiers in Plant Science* **9**: doi: 10.3389/fpls.2018.00161.

Qiao Y, Shi J, Zhai Y, Hou Y & Ma W (2015) *Phytophthora* effector targets a novel component of small RNA pathway in plants to promote infection. *Proceedings of the National Academy of Sciences* **112**: 5850-5855.

Qiao Y, Liu L, Xiong Q, Flores C, Wong J, Shi J, Wang X, Liu X, Xiang Q & Jiang S (2013) Oomycete pathogens encode RNA silencing suppressors. *Nature Genetics* **45**: 330-333.
Qutob D, Tedman-Jones J, Dong S, Kuflu K, Pham H, Wang Y, Dou D, Kale SD, Arredondo FD & Tyler BM (2009) Copy number variation and transcriptional polymorphisms of *Phytophthora sojae* RXLR effector genes *Avr1a* and *Avr3a*. *PloS One* **4**: e5066.

Reeksting BJ, Olivier NA & Van den Berg N (2016) Transcriptome responses of an ungrafted Phytophthora root rot tolerant avocado (*Persea americana*) rootstock to flooding and *Phytophthora cinnamomi*. *BMC Plant Biology* **16**: DOI 10.1186/s12870-12016-10893-12872.

Reitmann A, Berger DK & Van den Berg N (2017) Putative pathogenicity genes of *Phytophthora cinnamomi* identified via RNA-Seq analysis of pre-infection structures. *European Journal of Plant Pathology* **147**: 211-228.

Ren Y, Armstrong M, Qi Y, McLellan H, Zhong C, Du B, Birch PR & Tian Z (2019) *Phytophthora infestans* RXLR effectors target parallel steps in an immune signal transduction pathway. *Plant Physiology* **180**: 2227-2239.

Renny-Byfield S, Gallagher JP, Grover CE, Szadkowski E, Page JT, Udall JA, Wang X, Paterson AH & Wendel JF (2014) Ancient gene duplicates in *Gossypium* (cotton) exhibit near-complete expression divergence. *Genome Biology and Evolution* **6**: 559-571.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA & Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**: 539-542.

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**: 502-509.

Shan W, Cao M, Leung D & Tyler BM (2004) The *Avr1b l*ocus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene *Rps*1b. *Molecular Plant-Microbe Interactions* **17**: 394-403.

Stassen JH & Van den Ackerveken G (2011) How do oomycete effectors interfere with plant life? *Current Opinion in Plant Biology* **14**: 407-414.

Stassen JH, Boer Ed, Vergeer PW, Andel A, Ellendorff U, Pelgrom K, Pel M, Schut J, Zonneveld O & Jeuken MJ (2013) Specific *in planta* recognition of two GKLR proteins of the downy mildew *Bremia lactucae* revealed in a large effector screen in lettuce. *Molecular Plant-Microbe Interactions* **26**: 1259-1270.

Stefańczyk E, Sobkowiak S, Brylińska M & Śliwka J (2017) Expression of the potato late blight resistance gene *Rpi-phu1* and *Phytophthora infestans* effectors in the compatible and incompatible interactions in potato. *Phytopathology* **107**: 740-748.

Tian M, Win J, Savory E, Burkhardt A, Held M, Brandizzi F & Day B (2011) 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Molecular Plant-Microbe Interactions* **24**: 543-553.

Tyler B (2011) Entry of oomycete and fungal effectors into host cells. *Effectors in Plant-Microbe Interactions,*(Martin F & Kamoun S, eds.), p.^pp. 243-278. Wiley-Blackwell, Oxford.

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D & Beynon JL (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* **313**: 1261-1266.

van Damme M, Cano LM, Oliva R, Schornack S, Segretin ME, Kamoun S & Raffaele S (2011) Evolutionary and functional dynamics of oomycete effector genes. *Effectors in Plant–Microbe Interactions,*(Martin F & Kamoun S, eds.), p.^pp. 101-102. Wiley-Blackwell, Oxford, UK.

van Poppel PM, Guo J, van de Vondervoort PJ, Jung MW, Birch PR, Whisson SC & Govers F (2008) The *Phytophthora infestan*s avirulence gene *Avr4* encodes an RXLR-dEER effector. *Molecular Plant-Microbe Interactions* **21**: 1460-1470.

Vetukuri RR, Whisson SC & Grenville-Briggs LJ (2017) *Phytophthora infestans* effector Pi14054 is a novel candidate suppressor of host silencing mechanisms. *European Journal of Plant Pathology* **149**: 771-777.

Wang Q, Han C, Ferreira AO*, et al.* (2011) Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. *Plant Cell* **23**: 2064-2086.

Wang X, Boevink P, McLellan H, Armstrong M, Bukharova T, Qin Z & Birch PR (2015) A host KH RNA-binding protein is a susceptibility factor targeted by an RXLR effector to promote late blight disease. *Molecular Plant* **8**: 1385-1395.

Whisson SC, Boevink PC, Moleleki L*, et al.* (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* **450**: 115-118.

Win J, Morgan W, Bos J, Krasileva KV, Cano LM, Chaparro-Garcia A, Ammar R, Staskawicz BJ & Kamoun S (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* **19**: 2349-2369.

Xiong Q, Ye W, Choi D, Wong J, Qiao Y, Tao K, Wang Y & Ma W (2014) *Phytophthora* suppressor of RNA silencing 2 is a conserved RxLR effector that promotes infection in soybean and *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions* **27**: 1379-1389.

Yang B, Wang Q, Jing M*, et al.* (2017) Distinct regions of the *Phytophthora* essential effector Avh238 determine its function in cell death activation and plant immunity suppression. *New Phytologist* **214**: 361-375.

Yang B, Wang Y, Guo B*, et al.* (2019) The *Phytophthora sojae* RXLR effector Avh238 destabilizes soybean Type2 GmACSs to suppress ethylene biosynthesis and promote infection. *New Phytologist* **222**: 425-437.

Yang L, McLellan H, Naqvi S, He Q, Boevink PC, Armstrong M, Giuliani LM, Zhang W, Tian Z & Zhan J (2016) Potato NPH3/RPT2-like protein StNRL1, targeted by a *Phytophthora infestans* RXLR effector, is a susceptibility factor. *Plant Physiology* **171**: 645-657.

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.

Yin W, Dong S, Zhai L, Lin Y, Zheng X & Wang Y (2013) The *Phytophthora sojae Avr1d* gene encodes an RxLR-dEER effector with presence and absence polymorphisms among pathogen strains. *Molecular Plant-Microbe Interactions* **26**: 958-968.

Yu X, Tang J, Wang Q*, et al.* (2012) The RxLR effector Avh241 from *Phytophthora sojae* requires plasma membrane localization to induce plant cell death. *New Phytologist* **196**: 247-260.

Zeilinger S, Gupta VK, Dahms TE, Silva RN, Singh HB, Upadhyay RS, Gomes EV, Tsui CK & Nayak SC (2016) Friends or foes? Emerging insights from fungal interactions with plants. *FEMS Microbiology Reviews* **40**: 182-207.

Zentmyer GA (1984) Avocado diseases. *International Journal of Pest Management* **30**: 388-400.

Zhang J (2003) Evolution by gene duplication: an update. *Trends in Ecology & Evolution* **18**: 292-298.

Zheng X, Wagener N, McLellan H, Boevink PC, Hua C, Birch PR & Brunner F (2018) *Phytophthora infestans* RXLR effector SFI 5 requires association with calmodulin for PTI/MTI suppressing activity. *New Phytologist* **219**: 1433-1446.

Zheng X, McLellan H, Fraiture M, Liu X, Boevink PC, Gilroy EM, Chen Y, Kandel K, Sessa G & Birch PR (2014) Functionally redundant RXLR effectors from *Phytophthora infestans* act at different steps to suppress early flg22-triggered immunity. *PLoS Pathogens* **10**: e1004057.