Department of Environment and Agriculture

School of Science and Engineering

Genomic Analysis of Fungal Species Causing Ascochyta Blight in Field Pea

Chala Jefuka Turo

This thesis is presented for the Degree of

Doctor of Philosophy

of

Curtin University

February 2016

School of Science and Engineering

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature: _____Date: _____

Acknowledgements

I would like to express my appreciation to my initial principal supervisor Dr Judith Lichtenzveig for providing me the opportunity to study at Curtin University. Thank you for remarkable mentorship, encouragement and patience during my study. Special thanks go to my principal supervisor, Dr James Hane, for his tremendous mentorship and invaluable inputs. I am thankful for teaching me bioinformatics from the scratch to advanced computational biology, guidance and advice throughout writing my thesis. I have no words to explain, without you, it would have been impossible to produce this thesis. I am thankful to my co-supervisor, Dr Bernadette Henares, for your supervision, teaching quantitative PCR techniques and commenting the thesis. I would also like to thank my chair of supervisory committee, Professor Mark Gibberd, for evaluating my research progress, encouragement and serving committee members, particularly in the event of hardship situations. My special thanks extends to Dr Kate Trinajstic for serving as the chair of my supervisory committee toward the end of my studies.

I am grateful to Julie Lawrence for teaching me molecular laboratory methods and for provision of technical assistances. I extend my special thanks to Lydia Kupsky, who taught me how to perform florescent microscopy. I'm indebted to invaluable technical and moral support obtained from Lina Farfan Caceres from start to end of my thesis. I am also grateful for bioinformatics support obtained from Robert Syme and Angela Williams. I would like to thank Francis Kessie, Robert Lee, Winnie Liu, Johannes Debler, Christy Grime, Bahram Sharifnabi and Fiona Kamphuis for various inputs and assistances provided. I would like to extend my gratitude to all CCDM staff members for their kind collaboration. I'm grateful to support from Julie Craig, from the International Sponsored Student Unit, for logistics arrangements and support.

I am thankful to Australian Government for providing me a scholarship opportunity. I am also grateful to the financial assistance obtained from the Department of Environment and Agriculture, Curtin University. Finally, I would like express my appreciation to my beloved wife Aklil Knifu Kelifa, who always supports me in every moment and in all aspects throughout my studies.

Abstract

The fungal pathogen *Peyronellaea pinodes, Peyronellaea pinodella, Ascochyta pisi* and *Phoma koolunga* are the most destructive fungal pathogens of field pea. Whole genome sequencing, comparative and *in planta* transcriptome studies were pursued to test the hypothesis that these pathogens secrete effectors as part of their pathogenicity arsenal. The genome of isolates of *P. pinodes* (2), *P. pinodella* (2), *A. pisi* (3) and *Ph. koolunga* (2) were completely sequenced and evaluated. *In planta* transcriptomes of *P. pinodes*, *P. pinodella* and *Ph. koolunga* were obtained as part of effector prediction. The genome assemblies varied between 29 and 33 Mbp with equilibrated GC content (51- 53%). The synteny relationship between *P. pinodes* and other fungal pathogens is congruent to their established evolutionary relationship. Genome analysis uncovered plant cell wall de-polymerization strategies evolved by ascochyta blight pathogens to establish themselves in their host. Heterologous expression confirmed that necrosis-inducing proteins (NLP) from Didymellaceae elicited cell-death only to dicots. This cytotoxic NLP is absent from cereal pathogens analysed suggesting niche adaptive evolution of Didymellaceae pathogens. About 223, 226 and 123 effector candidates were predicted from *P. pinodes*, *P. pinodella* and *Ph. koolunga,* respectively. *In planta* expressed effectors might play crucial role during infection. Further functional analysis will lead to development of novel pea breeding techniques. The availability of these genome data have contributed to scare molecular data available for studying host pathogen interaction. It is perceived that public availability of these data will increase public research engagement, hasten integration of available host and pathogen genome resources and boost pea improvement programs.

List of Tables

List of Figures

Abbreviations

ABPFP	Ascochyta blight pathogens of field pea
AGRF	Australian Genome Research Facility
BLAST	Basic local alignment search tool
CAZyme	Carbohydrate Active Enzyme
CCDM	Centre for Crop and Disease Management
CEGMA	Core eukaryotic Gene Mapping Approach
CWDE	Cell Wall Degrading Enzymes
ETI	Effector Triggered Immunity
IPTG	Isopropyl-1-thio b- d-galactopyranoside
kDa	kilo Dalton
LJD	Long Jumping Distance
M	Mole
MAPK	Mitogen-Activated Protein Kinases
Mbp	Mega base pair
ml	Millilitre
NCBI	National Centre for Biotechnology Information
NGS	Next Generation Sequencing
NRPS	Non ribosomal peptide synthase
PAML	Phylogenetic analysis by maximum likelihood
PAMP	Pathogen Associated Molecular Pattern
PCR	Polymerase chain reaction
PDA	Pisatin demethylase
PHI	Pathogen Host Interaction
PKS	Polyketide synthase
PTI	AMP-triggered immunity
RBBH	Reciprocal Best Blast Hits
RIP	Repeat Induced Point Mutation
SMRT	Single molecule real-time sequencing technology
SNP	Single nucleotide polymorphism
µl	Micro litre
µm	Micro meter
µM	Micro Molar

1. General Introduction

The production of cereal crops continue to be the most important food source to feed the growing world population [81]. Increasing crop yield and crop diversifications are required to support projected population growth by 2050 [317]. Gan *et al.* [101] indicated that cropping systems diversification with pulses improves soil water conservation, ameliorate soil nitrogen availability, and promote overall system productivity. Field pea (*Pisum sativum* L.) is a pulse crop valued for its ability to aid in cereal production in crop rotation systems [215]. The productivity of field pea is constrained by a number of biotic stresses among which ascochyta blight (syn: Blackspot) is the most detrimental diseases of field pea worldwide [28, 84].

Ascochyta blight of field pea is caused by at least four fungal pathogens, including *Peyronellaea pinodes, Peyronellaea pinodella, Ascochyta pisi* and *Phoma koolunga* [31, 57, 240]. The disease is responsible for losses of AUD$19.6 million per annum to the Australian pea industry [223]. Littile is known about the molecular basis of the interaction between the four pathogens and field pea at a molecular level. Advances in molecular studies of these pathogens have been restricted largely by a scarcity of molecular data. Despite a plethora of fungal genomes having been sequenced, there is no published genome sequence available for the Didymellaceae family – an important fungal taxonomic group containing these four important pea pathogen species. This research project was designed to understand the molecular interaction between *P. pinodes, P. pinodella, Ph. koolunga* and their host *Pisum sativum* through genome and *in planta* transcriptome sequencing and profiling of each pathogen during pea infection. The objectives are: sequencing and comparative analysis of *P. pinodes, P. pinodes, P. pinodella, and Ph. koolunga* genomes, evaluation and characterization of key pathogenicity gene candidates, and *in planta* transcriptome profiling of *P. pinodes, P. pinodella*, and *Ph. koolunga* at various pea infection stages.

The project's outputs are organized into six main chapters as shown in Figure 1. The first chapter describes the profile of *P. pinodes, P. pinodella, Ph. koolunga* and *A. pisi* within the context of taxonomic relatedness, economic relevance, and host range and disease epidemiology. The overall summary of current research project is highlighted in chapter 1. Chapter 2 covers pathogen profile, population structure and molecular mechanisms of pathogenicity including phytoalexin detoxification, secretion of molecules that suppress host defence responses and production of toxic metabolites.

Chapter 3 describes genome assembly and gene annotation of the Didymellaceae pathogens, with a major focus on improving the assembly of the reference species (*P. pinodes* isolate M074) using multiple libraries of different insert sizes. It presents a summary of genome assembly statistics, repeat contents of each genome, number of predicted open reading frames and comparative analysis between fungal genomes within the taxonomic order Pleosporales. Analyses include investigation of compartmentalized genome structure and the location of isolate specific genomic regions in closely related species (*P. pinodes* and *P. pinodella*), enrichment of morphogenesis and transcription related domains in *P. pinodes*, and gene family expansion of carbohydrate active enzymes (CAZymes).

Chapter 4 describes findings associated with pathogenicity related genes that have been reliably identified and partially characterized in Ascochyta Blight Pathogens of Field Pea (ABPFP). It looks into the functional divergence between necrosis- and ethylene inducing - peptide like proteins (NLPs) of ABPFP based on cytotoxicity assay of purified proteins upon infiltration on pea stipules and other dicot plants. Also presented are comparative and phylogenetic analysis followed by a functional assay implemented to delineate the level of sequence conservation and functional relevance of NLPs in ABPFP are also presented.

Chapter 5 includes detailed accounts of conserved fungal effector candidate genes derived from the analyses presented in earlier chapters. This chapter comprises the prediction and identification of lysine motif (LysM) and Cyanovirin-N (CVNH)

protein families which are compared against well-characterized LysM effectors in other fungal pathogens to elucidate their possible involvement in fungal infection process. This chapter also covers prediction, comparative and phylogenetic analysis of putative pisatin demethylase CYP450 (PDA), a potent antifungal compound produced by pea plant. Finally, this chapter includes speculation that the LysM and pisatin demethylation processes may work together toward protecting the fungus and promoting survival in hostile host-associated environments during infection.

Chapter 6 provides secretome and *in planta* transcriptome analysis, combined with data derived from *in-silico* and experimental sources to predict putative effector candidate genes in *P. pinodes*, *P. pinodella* and *Ph. koolunga*. The level of congruence between fungal effector prediction using overall ranking of various evidences and recently developed machine learning based effector prediction tools is illustrated. The chapter also encompass catalogues of secreted protein sequences, *in planta* expression patterns of predicted effectors and shortlisted effector-like candidate genes in *P. pinodes*, *P. pinodella* and *Ph. koolunga*. It outlines the possible role of predicted effector like candidate genes, present evidence of selective constraints and indicates conservation of effector-like candidate genes exhibited across the three fungal pathogens.

The final chapter (Chapter 7) summarizes the purpose of the research project and the major findings of each previous chapters, with further discussion of the implications and benefits to both the scientific community and to industry.

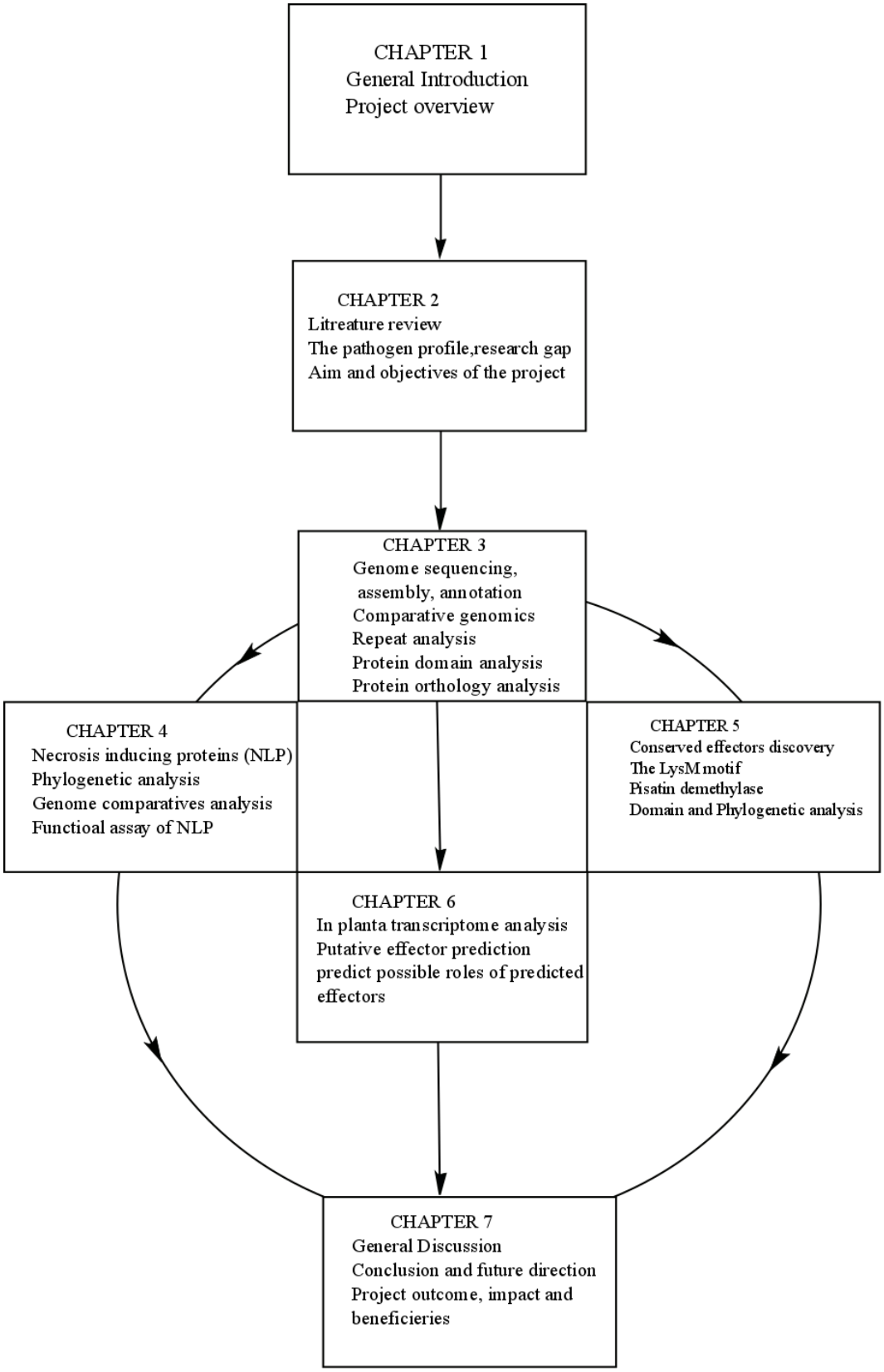


Figure 1 Flowchart of the research projects implemented in this studies.

2. Molecular Basis of Ascochyta Blight Pathogens of Field Pea in the genomic Era

2.1 Introduction

Ascochyta blight (syn: Blackspot) disease is among the most detrimental diseases of field pea worldwide [28, 84]. The disease is caused by at least four pathogen species including *Peyronellaea pinodes, Peyronellaea pinodella*, *Ascochyta pisi* and *Phoma koolunga* [28, 84]. It inflicts an estimated yield loss of between 10% and 100% under various conditions around the globe [31, 112, 270, 339].

Exploration of genomics of agriculture related microbe significantly contributed to the improvement of crop for human welfare through development of novel techniques to fight phytopathogens. The bulk of knowledge and information accumulated on contribution of genomics to host – pathogen interaction studies of major fungal pathogens of dicot and monocot fungal pathogens were summarized in recent publication [68, 69]. These resources continue to assist in the development of novel solutions to new problems that may arise in the agricultural development. However, there is limited information available for fungal pathogens of pulse crops in general and field pea in particular. So far, there is no published genome information available for the fungal pathogens in the Didymellaceae family that attack field pea.

2.2 Molecular basis of interaction between plants and pathogens

Comprehensive understanding of the interaction between the pathogen and host is crucial for understanding the compatibility that leads to disease symptom development. The discovery of effectors involved in host-pathogen interaction is one of the major advances in the field of molecular plant pathology. Effectors are proteins or small metabolites that alter host cells structure and function or alter immune responses [137, 154]. Various pathogenic fungi utilize diverse effectors to override

intrinsic host defence mechanisms regardless of pathogen lifestyle. Necrotrophic, biotrophic, mutualistic as well as endophytic microbes produce effector molecules to manipulate their host [197, 254, 267]. These indicate that fungi with diverse lifestyles secrete/produce effector repertoires to modulate their host to survive and flourish while restricting themselves from recognition by their respective host.

Current knowledge of necrotrophic fungal effectors indicate two types of effectors employed to debilitate host immune responses. The first group of necrotrophic pathogens such as *Cochliobolus victoriae*, *Parastagnospora nodurum* (SnTox1, SnTox3, SnToxA, SnTox6) and *Pyrenophora tritici-repentis* (PtrToxA) secrete effectors that triger host susceptibility (effector – trigered susceptibility) [90, 93, 102, 196, 198]. On the other hand, necrotrophic pathogens such as *Rhynchosporium commune* and *Botrytis cinerea* secrete non-specific effectors that induce plant cell death [165, 276]. Biotrophic and hemibiotrophic pathogens like *Cladosporium fulvum*, *Melampsora lini*, *Ustilago maydis*, *Leptosphaeria maculans*, *Magnaporthe oryzae* and *Colletotrichum higginsianum* secrete avirulence effectors that trigger hypersensitive response [100]. Some of these fungal pathogens have evolved effectors that prevent the activation of pathogen associated molecular pattern (PAMP triggered immunity). The *Ecp6* gene in *C. fulvum* and *Slp1* gene in *M. grisea* are examples of such effectors [63, 216]. Therefore, both generalized and specialized effector candidates could be expected from pathogens of necrotrophic lifestyle. Moreover, both necrotrophic and hemi/biotrophic pathogens maneuver their host and their distinction remains elusive [231].

Effector discovery in ABPFP of field pea is at its infant stage. As in other filamentous ascomycete fungi, ABPFP may secrete a number of effectors to establish compatibility with their host. *Peyronellaea pinodes* is reported to secrete suppressors that are able to inhibit host plant defence responses [284, 285]. This review will attempt to document recent developments in the fields of ascochyta blight pathogen and field pea interaction and highlight research areas that could be addressed in the future to enhance the development of pre-breeding tools in field pea improvement programs.

2.3 Taxonomy and relatedness of ascochyta blight pathogens of field pea

Ascochyta blight of field pea is caused by at least four fungal pathogens including *Peyronellaea pinodes, Peyronellaea pinodella, Ascochyta pisi* and *Phoma koolunga.* In addition, *Boeremia exigua* var. *exigua* and *Phoma herbarum* were reported as part of the ascochyta blight of field pea [192, 193]. Based on phylogenetic analysis of multiple loci, *P. pinodes, P. pinodella* and *A. pisi* are closely related species that belong to specific clade of Didymellaceae family within the order Pleosporales [11, 57, 60, 240].

Peyronellaea pinodes and *P. pinodella* are very closely related species. Phylogenetic analysis of sequence data from Internal Transcribed Spacer regions (ITS) are unable to differentiate the two specie [240, 335]. They can, however, be distinguished based on phylogenetic analysis of glyceraldehyde-3-phosphate dehydrogenase (G3PD) sequence [240]. Evolutionary analysis using multiple loci including 28S rDNA (Large Subunit - LSU), 18S rDNA (Small Subunit - SSU), the ITS 1 and 2 and 5.8S rDNA (ITS) and part of the β-tubulin (TUB) gene region refined their polytomy distribution and confirmed their placement in the *Peyronellaea* (Berk. & A. Bloxam) Aveskamp, Gruyter & Verkley [11].

Molecular analysis based on mating loci can also be used as a marker to distinguish the *P. pinodella* from *P. pinodes* as opposed to ITS [335] . In fact *P. pinodes* and *P. pinodella* are different in their mode of reproduction. *Peyronellaea pinodes* is a homothallic where both reproductive structures are on same thallus whereas *P. pinodella* is a heterothallic requiring two opposite partners for sexual reproduction [29, 38, 171]. Similarly, *A. pisi* is heterothallic mating type confirmed by both molecular investigation and *in vitro* crossing [38].

The prevalence and economic importance of *A. pisi* in Australia is currently limited. A recent report by Murray and Brennan [223] showed that *A. pisi* is present in major pea growing regions with negligible impact on crop yield, which is consistent with

previous reports from southern Australia [5]. The availability of good sources of resistance to *A. pisi* might have controlled the pathogen in Australia and elsewhere [5], but the disease is recently expanding in global distribution. It has been recently reported from Spain [151] and South Dakota [212] and Idaho [123] in the United States of America. The pathogen could be a potential threat to the pea industry in the event of its resurgence under the present climate scenario.

Phoma koolunga has been reported as an important causative agent of ascochyta blight of field pea in Southern Australia [57, 223]. Recent report indicated the existence of this pathogen in Western Australia (Li *et al.,* 2014). The species is becoming more frequent in samples from field pea sites with similar distribution to *P. pinodes* [58]. The fungus can be clearly distinguished from the other ABPFP in culture [57]. According to phylogenetic analysis based on DNA-dependent RNA polymerase II nucleotide sequences (RPB2) and G3PD *Ph. koolunga* is more closely related to *A. pisi* than *P. pinodes* and *P. pinodella* [38]. To date, there is no information on the species' teleomorph nor its mating types. The reproductive mode has significant epidemiological implications and is one of the factors influencing the relative prevalence of ascochyta blight complexes in the field.

2.4 Infection and disease development in ascochyta blight pathogens of field pea

The nature of infection and establishment strategies of the ABPFP are not fully understood. Only Clulow *et al.* [44] indicated occurrence of biotrophic and necrotrophic stage of development in *P. pinodes* during epicotyle infection which was not observed during leaf infection. As opposed to this, Roger and his colleagues [262] indicated that cuticle penetration takes place by germ-tube formed underneath appressoria followed by intercellular mycelial growth before inciting cell death and suggested as hemibiotrophic pathogen. Toyoda *et al.* [318] regarded *P. pinodes* as hemibiotrophic while Le May *et al.* [183] and Tivoli *et al.* [316] considered as necrotrophic pathogen. Based on recent fungal developments observed during pea leaf infection, consideration of ABPFP as a necrotrophic pathogen where penetration takes place via natural openings or direct penetration through formation of appresoria.

Leaf infection by *P. pinodes*, *P. pinodella* and *Ph. koolunga* induces indistinguishable disease symptoms under field conditions. Early infection of plants results in small purple spots that under favourable conditions enlarge and turn brown to black with sometimes-definite margins and zonate appearance [31, 315]. Lesion expansion is followed by aggressive mycelium extension through diffusion of toxins, enzymes and suppressors that allow removal of physical barriers and delay host responses [315]. The lesions developed in response to infection by *A. pisi* were easily distinguished from those symptoms induced by *P. pinodes* and *P. pinodella* by tan-coloured lesions with discrete dark margins [38, 118, 315].

The ABPFP disease cycle is very similar to those described for *Ascochyta rabiei* by Kaiser and his co-workers [150]. During wet weather ascospores of *P. pinodes* can be transported a considerable distance by wind while asexual pycnidiospores will be distributed over short distances by rain splash [261]. Early infection coupled with periodic wet weather events lead to rapid spread of the disease and can result in complete crop failure. Affected leaves remain attached to the plant. Stem lesions are similar in colour, elongate vertically or girdle the stem killing the entire plant with a blue-black appearance. Under conditions unfavourable for the pathogen, the lesions remain brownish on the stem. Infection can also appear on flowers and seed infection may show no symptoms or may show shrinkage and dark-brown discolouration [31]. Severe infection aggravates the senescence of plants at early maturity leading to extensive quality deterioration.

Peyronellaea pinodes and *P. pinodella* can survive on stubble, seeds and pea trash and in the soil as sclerotia and chlamydospores that serve as a primary source of inoculum (Figure 2)In *A. pisi*, chlamydospores are rare or absent and the pathogen has low saprophytic ability to survive during non-cropping season and over-summer on pea remnants [31, 249]. Therefore, seed infection is the most important aspect of *A. pisi* transmission [31]. Tivoli and Banniza [315] also showed that seeds were the main source of introduction and dissemination of various ascochyta species in many countries. Although epidemiological studies on *Ph. koolunga* is limited, emerging reports indicated that *Ph. koolunga* can be transmitted through seed infection [160].

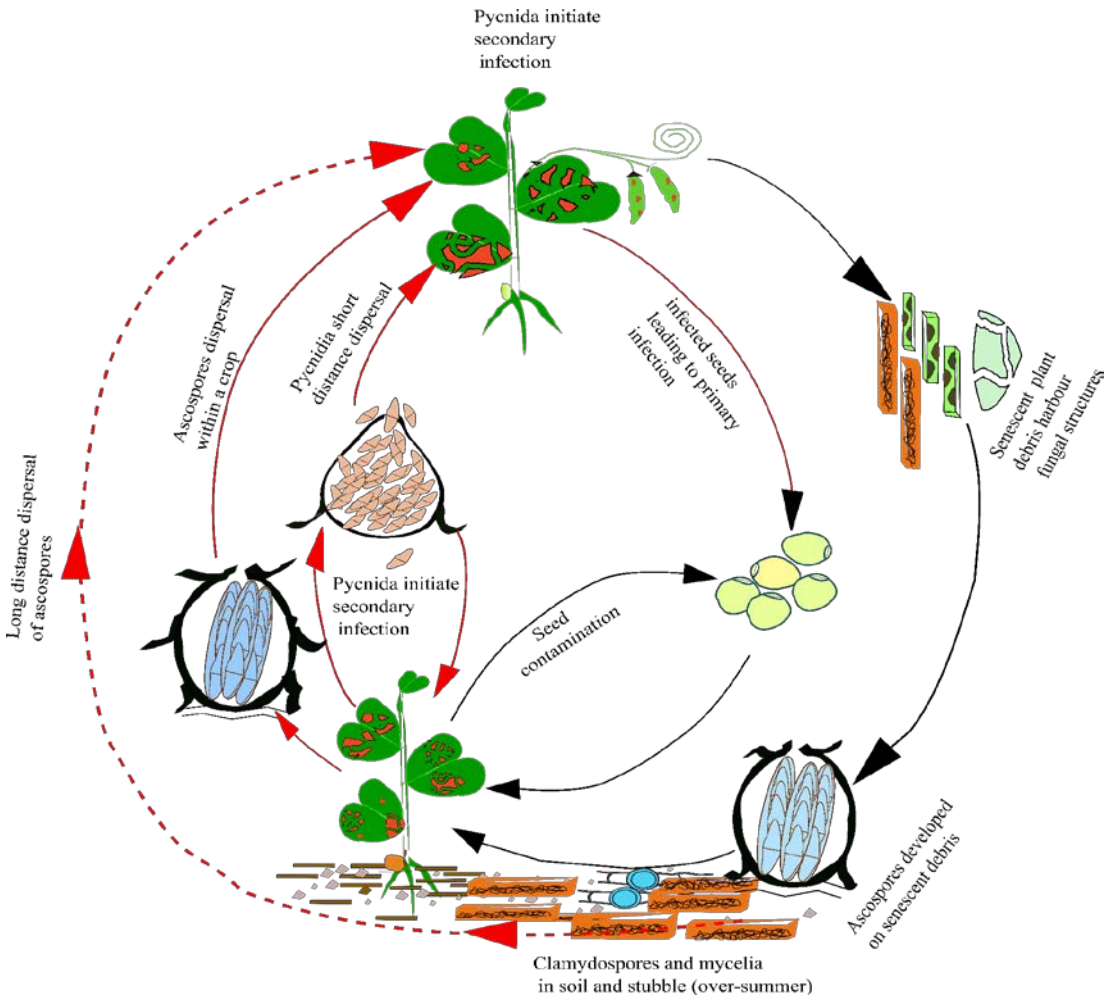


Figure 2 Peyronellaea pinodes disease cycle.

Red and black arrows indicate processes that take place during winter and summer, respectively. Red arrows with broken line indicates long distance dispersal of ascocpores released from pseudothesia. The figure is collated based on [31, 261, 315].

2.5 Population structure of ascochyta blight pathogens of pea (ABPFP)

There are two conflicting views with regard to *P. pinodes* pathotype variation. The first posits that there is no or only minor host genotype specialization in the *P. pinodes* pathogen population [232, 336]. The other view is the existence of 6 – 22 pathotypes among *P. pinodes* population [280, 300, 338, 350]. Each study considered different with respect to host genotypes, isolate population, culture age, seedling stage, and experimental conditions used to categorize the pathogen population. The lack of concordant findings is related to the quantitative nature of resistance in cultivated *Pisum* [87, 248, 313, 314, 351] and high level of genotype by environment interaction of such traits and differences in experimental settings.

Variation in the level of aggressiveness among isolates of *A. pisi* has been reported. Ali *et al.* [4] identified fifteen groups of *A. pisi* isolates based on 58 differential lines in Australia. Similarly, Darby *et al.* [56] identified five pathotype groups among 57 isolates of *A. pisi* tested on 15 host lines in Britain. As opposed to *P. pinodes* and *A. pisi*, there was no variation in aggressiveness of *P. pinodella* that could group into pathotype [4]. The level of isolate variation in *Ph. koolunga* population has not been reported. Generally, it seems that there is no compelling evidence of host genotype specificity in cultivated *Pisum* species.

It is essential to consider isolates from both cultivated and wild *Pisum* species in pathotype variation studies. Human selection for agronomically desirable trait decreases the level of genetic variation in the domesticated crop relative to their wild type ancestors [107]. This will compromise the resistance level in the domesticated crop and predispose these cultivars to susceptibility [298]. The spatial and temporal selection pressure imposed by the host can lead to emergence of homogeneous pathogen population structure. Therefore, it is not a surprise to see more or less uniform ABPFP population from cultivated *Pisu*m species which are selected for desirable agronomic traits, but lack high level of resistance.

2.6 Host specificity of ascochyta blight pathogens of field pea

Fungal pathogens in the Didymellaceae exhibited different levels of host specificity. Previous studies indicate that *P. pinodes* was able to infect *Pisum, Lathyrus, Phaseolous, Vicia species, Medicago, Vicia* and *Trifolim* [249]. Because of this, Le May *et al.* [182] concluded that *P. pinodes* is not a specialized pathogen. However, when taxonomic relatedness of the hosts exploited by *P. pinodes* are considered, all the enlisted hosts belong to the Fabaceae family indicating specialization to this specific host group. Similarly, *A. pisi* is able to induce visible disease symptom only on specific hosts within Fabaceae [135, 240]. [240]. In contrast to the narrow host ranges of *P. pinodes* and *A. pisi*, *P. pinodella* was able to infect hosts that belong to at least ten families including Fabaceae (pea, glycine), Amaranthaceae (Beta), Rubiaceae (coffee), Malvaceae (Gossypium), Poaceae (Hordium, Casuarinaceae (Casuarina), Polemoniaceae (Phlox), Asteraceae (Lactuca), Amaryllidaceae (Galanthus), and Appiaceae (petroselinum)[164]. Thus, consideration of taxonomic relatedness is important to capture the level of host exploitation over evolutionary time rather than classical cross infection tests alone [246]. Critical analysis of such concept using host and pathogen phylogenetic approach would answer why and how very closely related pathogen species evolved different degree of host specificity in ascochyta blight pathogens of field pea.

Interests has been growing in discovering determinants and mechanisms of pathogens host specificity. Determinants of host specificity are impacted by changes in genetic constituents of the pathogen, ranging from single nucleotide changes to genomic islands or mobile elements [17, 166, 299]. Some of this host specificity determinants are under rapid evolution [245]. Most of the necrotrophic host specific determinants are secreted molecules that confer pathogen compatibility to a particular host [93]. Interestingly, the generalist necrotrophic fungal pathogen *B. cinerea* also contains a host specific virulence factor encoding endo-arabinanase (*BcAra1*) [224]. It will be interesting to investigate the molecular basis of host specificity exhibited by the closely related pathogens of ABPFP. The research achievements made toward understanding

the ABPFP – pea interaction at molecular level were described in the following sections.

2.7 Molecular mechanisms of pathogenicity (effectors in ABPFP)

Most necrotrophic plant pathogen have evolved similar mechanisms to penetrate their host surface. They kill their host cell ahead of colonization using toxic molecules and lytic enzymes to support their life [231, 322]. The phytotoxins and lytic enzymes are secreted into the host tissue both prior to and during colonization leading to appearance and development of necrotic lesions [175, 322].

The early infection mechanisms, disease development and induced symptoms by ABPFP are typical characteristics of necrotrophic lifestyle. Microscopic *in planta P. pinodes* development showed that host cell death are evident ahead of colonization within the first 21 hours post infection (Figure 3). In addition, previous and emerging research results indicated that the pathogens are able to produce a number of molecules including toxic metabolites and hydrolytic enzymes during their interaction. Cell wall degrading enzyme production and possible contribution to lesion formation during infection by *P. pinodes, P. pinodella* and *A. pisi* was reported in 1970s [10, 134] but no further progress made to understand their role as virulence factors.

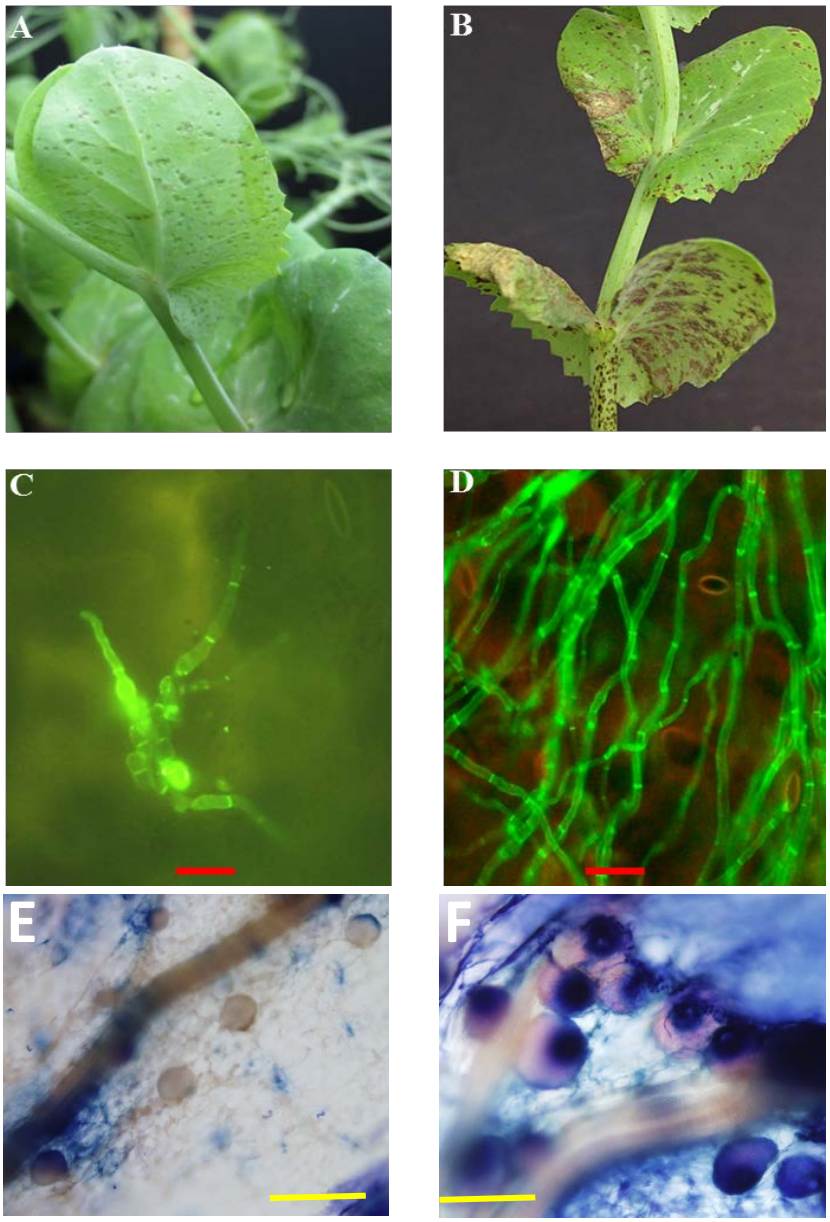


Figure 3 Disease symptom and in planta development of *P. pinodes* during pea infection.

A) Conspicuous sunken lesions at 21 hours post inoculation on youngest stipules.
B) Coalescence of lesions and infection advanced toward stem 96 hours post inoculation.
C) Mycelium development inside plant tissue observed under microscope after Alexa flour staining.
D) Proliferation of the fungus inside death tissue at 96 hours post infection.
E) Formation of pychnidia
F) Pychnidia maturation and spore libration

Red and yellow scale bar indicate 50 µm and 100 µm, respectively.

2.8 Phytoalexin detoxification is associated to virulence

Phytoalexins are toxic molecules synthesized *de novo* by plant tissues in response to microbial infection or stress conditions [304]. A number of reports indicate that fungal pathogens are capable of detoxifying phytoalexins. Studies have shown that *Alternaria brassicola, Leptosphaeria maculans* and *B. cinerea* are able to detoxify camalexin and brassinin, potent phytoalexins produced by cruciferous species, to non-toxic compounds [236, 237]. Moreover, Joubert *et al.* [148] suggested that camalexin and brassinin might activate fungal mitogen-activated protein kinases (MAPKs) which can help in shielding the fungus from host-induced stresses. In line with this, brassinin detoxification by *A. brassicola* leads to the suppression of various crucial plant defence responses [238].

A well characterized phytoalexin of pea is pisatin. It has been demonstrated that many pea pathogens (*P. pinodes, P. pinodella, A. pisi, Nectria haematococca, Fusarium oxysporum f.sp. pisi)* and other pathogens such as *F. oxysporum* f.sp. *phaseoli* detoxify pisatin via demethylation [71, 105, 120, 213]. The rate of demethylation positively correlates with virulence [71]. The gene responsible for the detoxification of pisatin was identified as a cytochrome P450 monooxygenase (known as pisatin demethylase, PDA)[213]. The PDA encoding cytochrome P450 monooxygenase are natural inducer of pisatin demetylation [159] and shown to have strong substrate specificity toward pisatin [104], are highly divergent from other known cytochrome P450 monooxygenases [203] and are expressed during infection of pea [195]. The existence of a similar cytochrome P450 in *F. oxysporum f.sp. pisi* was determined via southern blot analyses [71] and later by comparative genomics. P450 also exists in *F. oxysporum* f.sp. *phaseoli,* a pathogen of common bean [47].

Previous reports suggest that pisatin detoxification contributes to ascochyta blight disease development in field pea. Studies using a biochemical approach have shown that cytochrome P450 monooxygenase induce pisatin demethylation in *P. pinodes, P.*

pinodella and *A. pisi* [71, 105]. Very high concentration of pisatin was observed at early stage of pea (48 hours post infection) epicotyl infection by *A. pisi* [221].

Available evidence suggests that ABPFP carry unique form of PDA different from the wilt pathogen *N. haematococca*. There was no significant hybridization observed when DNA specific to the pisatin demethylase gene from *N. haematococca* was used to probe genomic DNA from *P. pinodes, P. pinodella and A. pisi* [71]. In contrast, cytochrome P450 genes with high similarity to *N. haematococca* MPVI pisatin demethylase (*PDA*) genes have been identified from *F. oxysporum* and *Neocosmospora* species [71, 309]. The authors hypothesized that pisatin demethylase from ABPFP may have a divergent evolutionary path. Further sequence analysis using phylogenetic and comparative genomics will allow us to better understand the evolution and relatedness of pisatin demethylases within ABPFP and across distantly-related fungal pathogens.

Plant pathogens have also evolved non-degradative phytoalexin tolerance mechanisms. The *N. haematococca* ABC transporter gene (*NhABC1*) was able to confer tolerance to a phytoalexin produced by potato (rishitin) and confer virulence (Coleman, *et. al*., 2011). Similarly, in *B. cinerea*, an ABC transporter (*BcatrB*,) is a virulence factor that increases pathogen tolerance towards camalexin in *A. thaliana* [294]. Stergiopoulos *et al.* [297] also indicated that an ABC transporter (*MgAtr4*) is a virulence factor of *M. graminicola* during pathogenesis on wheat. Because both the ABC transporter and P450 contribute to pathogenicity, it was proposed that *NhABC1* and the cytochrome P450 act sequentially toward pisatin tolerance [47]. It will be interesting to see whether contribution of both ABC transporter and cytochrome P450 are functionally active in ABPFP. Whole genome sequencing and comparative analysis of aggressive pathogens of field pea and their *in planta* transcriptomic analysis will help to identify the expression pattern of P450-related genes.

2.9 Suppressors interfere with host defence via the jasmonic acid signalling pathway

The intricate product of host and pathogen interaction determine their compatibility. Plants can sense their environment and are able to recognize invading pathogens through pathogen-associated molecular patterns (PAMPs), a process called PAMP-triggered immunity (PTI). At the same time, fungal pathogens may produce a diverse arsenal of pathogenicity effector molecules that can manipulate host-defence mechanisms or suppress host-immune responses. Although effectors are evolved by plant pathogens to establish basic compatibility, some effectors could be recognised by specific intracellular proteins leading to effector triggered immunity (ETI) which contribute to basal host defence [349].

Some plant pathogens secret molecules that incite both PTI and ETI simultaneously to deceive their host. For instance, two nonspecific glycoproteins are known to be secreted from pycnidiospores of *P. pinodes* and recognized by *P. sativum*. The glycoproteins are reported to act as elicitors and induce a number of active defence responses such as phytoalexins, superoxide generation, and formation of infection inhibitors and induction of pathogenesis related proteins (PR-proteins) [284, 318]. The induced defence in response to infection of pea plants by *P. pinodes* is blocked by two molecules released from the fungal spores known as supprescin A and B. These suppressors are small molecules secreted by *P. pinodes* that help the pathogen to establish inside plant tissue [285]. Over three decades of research devoted to understanding the mechanism of suppression of defence responses at the plant cell walls has been reviewed recently [284]. Accordingly, suppressors condition pea plants to be susceptible to a number of otherwise avirulent pathogens including: *A. alternata, Mycosphaerella ligulicola, Mycosphaerella melonis* and *Stemphylium sarcinaeforme* [284].

Recent studies on the molecular responses of pea to *P. pinodes* suppressors have used *Medicago truncatula* as a model infection system to dissect *P. pinodes* - pea interaction [318]. Toyoda *et al.* [318] showed that application of suppressors from *P. pinodes*

significantly reduced the accumulation of medicarpin, the phytoalexin produced by *M. truncatula*. Rao Uppalapati *et al.* [255] showed that the suppressor acts through inhibition of mitogen-activated protein kinase (MAPK) pathway, likely via the phosphoinositide (PI) signalling pathway, which leads to establishment of infection in pea. Recent suppression subtractive hybridization (SSH) analysis using a *M. truncatula* detached leaf disc assay indicated that *P. pinodes* was able to induces disease susceptibility via interference with jasmonic acid regulated cellular processes [318]. Such leaf disc inoculation analysis in *M. truncatula* has provided invaluable information for pea improvement, but may not reflect actual induced responses in pea.

2.10 Phytotoxic metabolites in ascochyta blight pathogens of field pea

Peyronellaea pinodes produces potent phytotoxic metabolites enable to induce necrosis on pea. Evidente *et al.* [79] identified three structurally related compounds in *P. pinodes* that belong to pinolidoxin nonenolides, which have strong phytotoxic activity on pea and beans. Later, [42] isolated and identified, pinolide herbarumin II and 2-epi-herbarumin II, compounds closely related to nonenolides from the aggressive *P. pinodes* isolate C0-99 liquid culture filtrates. Phytotoxicity assay of the pure metabolites on the host plant, some legumes and weeds indicated that only pinolidoxin induced non-specific necrosis across multiple host species. As mentioned earlier, *P. pinodes* has been suggested to have a broader host range than previously reported [182]. This may be related to the *P. pinodes'* ability to produce various phytotoxic compounds that might assist in establishment and disease development within Fabaceae.

Ascochtya pisi is also known to produce a phytotoxic secondary metabolite known as ascochytine [88]. Among the four species that form the ascochyta blight pathogen complex, ascochitine production is restricted to *A. pisi* [208]. A non-host specific phytotoxic chemical known as ascosalitoxin has also been reported to be produced by *A. pisi* [79]. There is no directly reported evidence for the contribution of ascochytine to *A. pisi* virulence. However, *in vitro* toxin production positively correlated to isolate virulence [189]. Conversely, Marcinkowska *et al.* [208] found that equally pathogenic

isolates of *A. pisi* produced different amounts of ascochytine under *in vitro* conditions. Furthermore, no relationship between pathogen virulence and *in vitro* production of ascochytine was observed in *Ascochyta fabae* [19], the ascochyta blight causing fungal pathogen in faba bean.

Chapter Three Attribution Statement

Title: Genome Sequencing and Comparative Analysis of *Peyronellaea pinodes*, *P. pinodella*, *Ph. koolunga* and *Ascochyta pisi*

Chala Jefuka Turo, Francis Kessie, Julie Lawrence, James Hane and Judith Lichtenzveig

These chapter includes genome sequencing and analysis of ascochyta blight pathogens of field pea, a project started before arrival of the PhD candidate. As a result, not all research activities performed by the candidate. The main contributors to this chapters are:

Julie Lawrence

- Performed fungal culturing and DNA/RNA extraction of all isolates except *A. pisi* (AP1).

Francis Kessie

- Performed initial 75 bp sequencing of *P. pinodes* (M074)

Judith Lichtenzveig

- Performed library preparation, sequencing and lead the project

Chala Jefuka Turo (PhD candidate)

- Performed culturing, DNA extraction and sequencing *Ascochyta pisi* (AP1).
- Quality control and whole genome assembling of all species.
- Gene prediction and functional annotation of all species
- Performed comparative genome analysis
- Wrote the chapter

I, Chala Jefuka Turo, certify that this attribution statement is accurate record of my contribution to this chapter.

_____ _____

Chala Jefuka Turo (PhD candidate) Date

I, James Hane, certify that this attribution statement is accurate record of Chala Jefuka Turo's contribution to results presented in this chapter 3.

_____ _____

James Hane (Principal Supervisor) Date

Note that the order of the authors will be determined at publication stage.

3. Genome Sequencing and Comparative Analysis of *Peyronellaea pinodes, P. pinodella, Ph. koolunga* and *Ascochyta pisi*

3.1 Introduction

Sequencing of fungal genomes started with the sequencing of the commercially important eukayotic fungus, *Saccharomyces cerevisiae*, first published in 1996 [109]. This was followed by genome sequences for *Neurospora crassa*, the first model organism representing filamentous ascomycetes [96]. Since then, the whole-genome sequencing of a number of economically-important fungal phytopathogens was driven by Sanger-based whole-genome shotgun (WGS) sequencing approaches [99]. The arrival of next-generation sequencing (NGS) technologies, single molecule real-time sequencing technology (SMRT), and advances in bioinformatics techniques for handling and processing large volumes of data have revolutionised whole-genome analysis. These advancements have each made it increasingly possible to study non-model organisms and perform comparative genomics at relatively low cost.

The revolution in phytopathohen sequencing and genome availability expanded our understanding of fungal evolutionary dynamics. Comparative genome analysis has enabled the identification and tracking of the origin of horizontal gene transfer between fungi and other organisms [85, 92], [207, 259]. Horizontal transfer of the necrotrophic effector gene *ToxA* from *Parastagonospora nodorum* to *Pyrenophora tritici-repentis* was first inferred from comparisons of both whole genomes [92, 271]. Recent reports suggest that *Pyrenophora teres* and *P. tritici-repentis* contain 14 horizontally transferred genes that have originated from bacteria and plants [301]. In addition, the acquisition of a hybrid NPPS/PKS gene cluster by ascomycete fungi from bacteria [181], transfer of cellulase genes from microbes to nematodes [214], and gain of a number of small secreted proteins from fungi by oomycetes [259] are only a few examples of HGT events that have been uncovered via whole genome sequence analysis.

Progress in comparative genomics has also expanded our understanding of mode of chromosomal evolution in fungi. Hane *et al.* [129] identified a unique pattern of chromosome sequence conservation specific to the filamentous ascomycetes, known as '*mesosynteny'*. The authors indicated that in mesosynteny, genes within homologous chromosomes are generally conserved, however gene order and orientation becomes rearranged with increasing phylogenetic distance. Subsequent simulation of different types of chromosomal dynamics in whole-genomes sequences from the class Dothideomycetes suggested that mesosyntenic may arise through the accumulation of multiple inversions [229].

The existence of repeat-induced point mutation (RIP) in fungi was first observed in in *Neurospora crassa* [279]. As more fungal whole-genome sequences became available, it became clear that signatures of RIP are common in filamentous Ascomycetes [8, 30, 45, 126, 229]. Previous reports have indicated that RIP recognizes duplicated DNA sequences and is likely to prevent genome invasion by transposable element through their inactivation, primarily via nonsense mutations [97]. While RIP is generally accepted to be restricted to the sub-phylum Pezizomycotina, a few studies also report RIP-like mutation in some species of the Basidiomycetes [127, 138]. Whole-genome analysis of the multi-nuclear fungus *Rhizoctonia solani* AG8, may suggest that where multiple nuclei are present RIP may not be restricted solely to repetitive elements, but could also broadly affect coding regions as well [127]. Furthermore, a number of avirulence genes have been observed within RIP affected genomic regions (typically long stretches of AT-rich sequences) in repeat-rich genomes such as that of *Leptosphaeria maculans* [94, 266, 321], where it may also be rapidly mutated by 'leaked' RIP mutations. It has been suggested that RIP mechanism may contribute to sequence diversification and adaptation to new hosts or new resistance genes within a host [116]. Evidence of RIP and inferences about its influence on pathogenicity adaptation learned from numerous genome studies, has informed monitoring and management strategies of *L. maculans* in canola [140, 266].

Whole genome sequencing followed by comparative analysis indicated the movement of pathogenicity genes in the genus *Fusarium* [201]. Genes in fungal pathogens with

dispensable chromosomes are enriched with duplicated genes and repetitive elements as observed in *Nectria haematococca* [46].

Whole-genome sequencing has exposed concealed lifestyle of necrotrophic fungal pathogens. Necrotrophic pathogens produce a number of effector genes to manipulate their host and establish compatibility [116, 168, 227, 231, 266, 308, 322]. Molecular studies increased our understanding of pathogen speciation and host specialization [299] and has enabled large scale analysis of plant cell wall degrading enzymes in fungi [163]. Moreover, widespread occurrence of necrosis and ethylene-inducing peptide-like proteins across three major taxa (fungi, bacteria & oomycetes) and their evolutionary histories were recognized from bioinformatics analysis of sequences sourced from various public database genome resources [106, 233, 234, 293].

Taking advantage of whole-genome sequencing, a number of economically important fungal plant pathogens from the class Dothideomycetes have been sequenced. Among these, *P. nodorum* [128], *Cochliobolus heterostrophus* [229], *Alternaria brassicicola* [55], *Pyrenophora teres fsp. teres* [78], *Pyrenophora tritici-repentis* [205], *Mycosphaerella graminicola* [111] and *Leptosphaeria maculans* [266] are few of the well-studied pathogens species from Dothideomycetes. The fungal pathogen *Peyronellaea pinodes, Peyronellaea pinodella, Phoma koolunga* and *A. pisi* also belong to the class Dothideomycetes within the family Didymellaceae. The genome sizes of *P. pinodes*, *P. pinodella* and *A. pisi* have been estimated to be between 24.0 - 28.8, 21.6 - 29.9 and 4.3 - 32.1 Mbp with 11 - 12, 10 - 15 and 12 - 16 chromosomes, respectively [2]. Only the *P. pinodes* genome has been sequenced using Illumina's Solexa paired-end (75 bp) sequencing technology, giving a fragmented assembly of 32.8 Mbp regardless of fragmented assemblies [158].

Sequencing the closely related pathogens from Didymellaceae would enable us to better understand the molecular basis of their pathogenicity shared among these complex pathogens, allow for comparative evaluation of the genome landscape between and within a species, and to design molecular markers for rapid species-

specific identification suitable for disease surveillance applications. One approach would be to build a good quality reference genome and sequence multiple isolates sequence at lower depth for comparison. The main objectives of this chapter are to assemble, annotate, and comparatively evaluate the *P. pinodes, P. pinodella, Ph. koolunga* and *A. pisi* genome sequences and improve the assemblies of the reference species (*P. pinodes* M074). This chapter covers comparative analysis of these genome sequences, describes their gene content including their expansion of carbohydrate active enzymes (CAZymes), and assesses the potential of ascochyta blight pathogens of field pea (ABPFP) to produce secondary metabolites.

3.2 Materials and Methods

3.2.1 Genome data acquisition and genome assembling

Aggressive isolate of *P. pinodes* M074 was collected from Medina, the Western Australia[158]. Additional isolates of *P. pinodes* and isolates of *P. pinodella, Ph. koolunga* and *A. pisi* were obtained from various sources as indicated in Table 1. The DNA from each pathogen species was isolated from single spored mycelium following cetyltrimethyl ammonium bromide (CTAB) genomic DNA extraction protocol [333].

The genome of *Peyronellaea pinodes* reference isolate M074 was initially sequenced using 454 life science sequencing (454 GS-FLX Titanium, Roche, Basel, Switzerland) with insert size 2000 bp, at University of Western Australia (UWA). Additional sequencing was performed using Illumina 75-100 bp paired-end reads with 200 – 500 bp insert sizes (Illumina, San Diego, USA). Additional long jumping distance (LJD) libraries (insert sizes 3kb, 8kb and 20kb) were applied to the reference isolate M074 (Eurofins, Luxembourg, Germany). All the remaining isolates of *P. pinodes, P. pinodella, Ph. koolunga* and *A. pisi* were sequenced using Illumina libraries with insert sizes between 200 and 500 with read lengths of 100 bp.

Table 1. Sequenced fungal pathogen species and generated genome data in genome assembling

Pathogen species	Isolates	Number of libraries	Number of reads after quality control (in Million)						Region of Origin
			75 bp	100 bp	3kb (LJD)	8kb (LJD)	20kb (LJD)	2kb (454)	
P. pinodes	M074	5	18.20*	-	30.42	16.02	23.8	0.23	Western Australia
P. pinodes	MP1	1	-	16.12	-	-	-	-	USA
P. pinodella	410/95	1	-	19.22	-	-	-	-	South Australia
P. pinodella	AWPP4B1I0	2	-	72.16	-	-	-	-	USA
Ph. koolunga	FT04040	2	-	74.91	-	-	-	-	South Australia
Ph. koolunga	FT0713	1	-	30.02	-	-	-	-	South Australia
A. pisi	AP1	3	-	74.89	-	-	-	-	Bulgaria
A. pisi	Georgia-12	1	-	8.73	-	-	-	-	Georgia
A. pisi	Georgia-7	1	-	8.71	-	-	-	-	Georgia

*Data available from prevous studies [158]. LJD long jumping distance; bp base pair; kb killobase

Low quality (≤ 28 phred score) sequences, adapter sequences and reads with less than 25 bp after trimming were removed from the reads using cutadapt version 1.1 (parameters -q 28 and -m 25) [211]. Unpaired reads resulted from quality clean up were separated from paired data into singleton fastq file formats. Paired-end reads with detectable overlap at their 3' ends were combined together to obtain longer single reads using FLASH version 1.2.2 (–x set to 0.2) [202]. The data quality process for LJD libraries was pre-processed by Eurofins as part of their services. The genome sequences were *de novo* assembled using SOAPdenovo2 version 2 [199]. GapCloser, (a module of SOAPdenovo) was also used to fill 'N' gaps created during scaffolding [199]. The general *P. pinodella, Ph. koolunga* and *A. pisi* genome assembly workflow was depicted in figure 4 B.

In hybrid assembling of the *P. pinodes* reference isolate M074, the LJD paired - end reads were sequentially used to scaffold the pre-assembled contigs using SSPACE (parameters: -x 1 –m 32 –o 20 –t 0 –k 5 –a 0.70 –n 15 -z 200 –p 0 –v 0 –g 0 –T 1) [22]. Gaps introduced between each scaffolding step were closed by GapFiller using paired-end reads between each steps [23]. In GapFiller, options were customised (-m 30 -o 2 -r 0.7 -n 10 -d 50 -t 10 -g 0 -T 1 -i 1) to a stringent gap overlap. Final assembly was achieved by a single round application GapCloser during initial assembling of 75 bp paired - end reads followed by four round of SSPACE and four rounds of gap filling steps. The overall *P. pinodes* reference isolate (M074) genome assembly strategies employed is indicated in Figure 4 A.

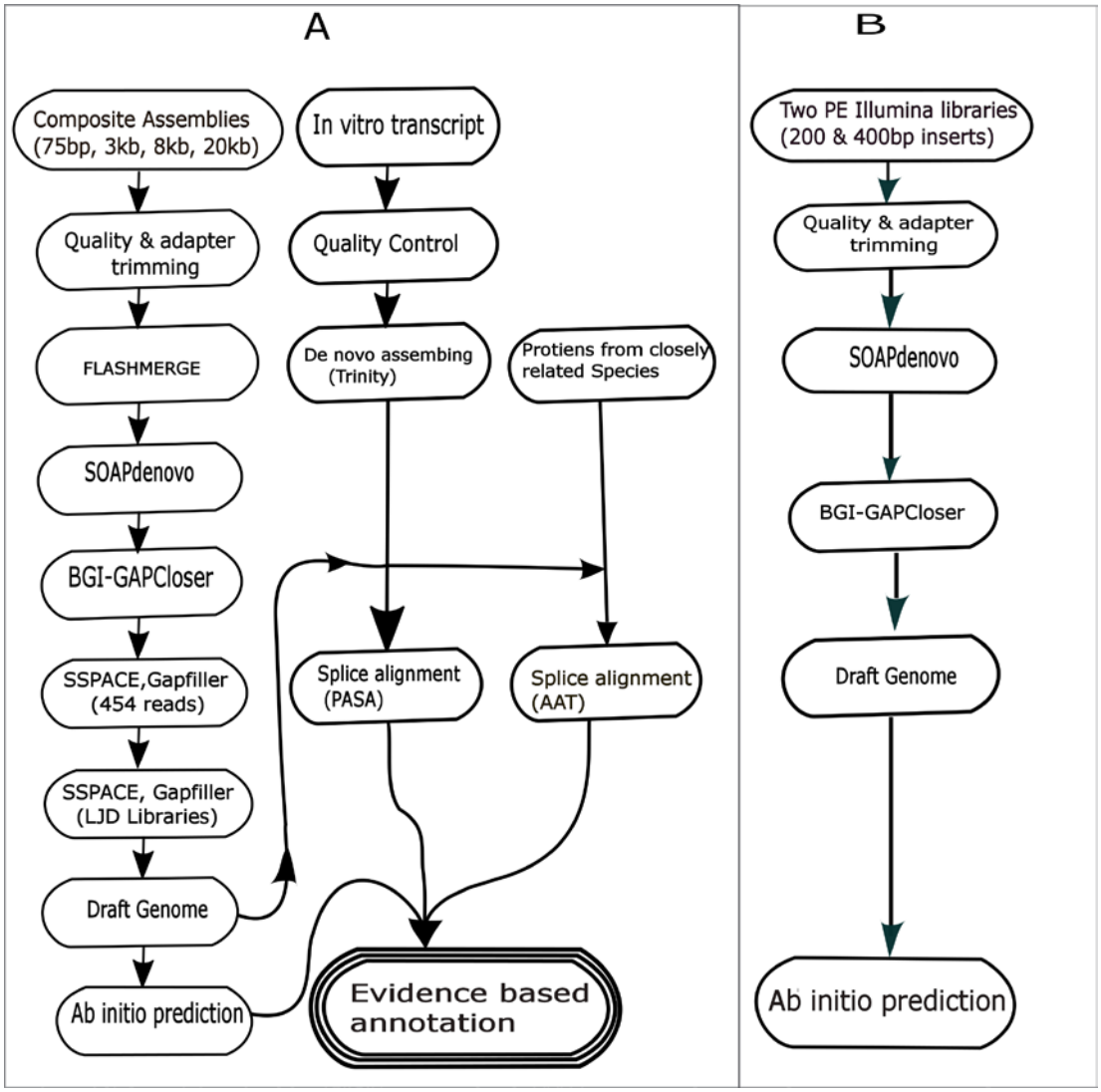


Figure 4 Whole genome assembly and annotation pipeline workflow for *P. pinodes* reference isolate (M074) (A) and *P. pinodella, Phoma, koolunga* and *A. pisi* (B).

3.2.2 Repeat identification and repeat-induced point mutation (RIP) analysis

The repeats in the genome sequences of ABPFP were predicted using RepeatMasker version 4.0.3 [16]. The distribution of repeats in the genome sequences were determined by aligning the genome sequences against RepBase version 20130422 with Cross Match version 1.090518 in sensitive mode [16]. The GFF output file from the RepeatMasker and respective genome sequences were supplied to RIPCAL to detect RIP-like activities following majority consensus model [125]. RIP index scanning were performed on each aligned repeat families using default RIP-index scanning parameters.

3.2.3 Gene finding

The assembled sequence of each fungal pathogen species was supplied to *ab inito* gene finding tool, GeneMark-ES version 2, to identify potential protein-coding genes in the fungal genome [27, 311] (Figure 4 A and B). Except for the reference isolate (where default parameters were used), minimum scaffold length was set to 15,000 nucleotides at initial self-training stage. The resulting proper output was translated into protein coding and nucleotide sequences with a minimum threshold length of ten amino acids.

To improve accuracy of genome annotation in the *P. pinodes* reference isolate (M074), *in vitro* transcripts from four conditions were *de novo* assembled using Trinity (2013-02-25) [114, 122]. The assembled transcript from each library was aligned to reference isolate genome through PASA alignment assembly pipeline using genomic mapping and alignment program (GMAP) as aligner [121]. Protein sequences from closely-related fungal pathogens were obtained from sources described in Supplementary Table 1. The proteomes were aligned to the genome assembly of *P. pinodes* reference isolate M074 using Analysis and Annotation Tool (AAT) [141] to obtain spliced-alignment of protein homologs. The PASA transcript assemblies, *ab initio* gene

prediction and proteome spliced alignments were supplied to EVidenceModeler [121] to generate a final set of gene predictions in GFF3 format.The reference isolate gene prediction workflow was indicated in figure 4 (A).

3.2.4 Functional annotation

The potential predicted genes from genomes of each ABPFP were functionally annotated using gene similarity search at National Centre for Biotechnology Information (NCBI) non-redundant dataset using BLASTp alignments tool [6]. InterProScan (version 4.8 - with signalp version 4.1) is used to predict the functional domains present in the protein sequences. Putative Gene Ontology (GO) terms were assigned based on InterproScan and BLASTp hits analysed through BLAST2GO [50]. Protein domains were identified by searching Pfam (Pfam-A) database [250] through HMMER3 using default parameters (gathering thresholds). Proteins destined for secretion were predicted using a combination of SignalP4.1 [243], Phobius 1.01 [153] and localized using WOLFPsort 0.2 [139]. Proteins with predicted transmembrane domains were identified using TMHMM 2.0 [172] . Protein sequences were classified as secreted if supported by either of SignalP, Phobius or WOLFPsort and contained $\leq$ 1 transmembrane domain.

3.2.5 Comparative genomic analysis

3.2.5.1 Synteny and read mapping analysis

Global alignment of whole genome were made using PROmer via MUMmer 3.0 [174] using max-match option. The MUMmer program delta-filter (-l 100 -i 70 -g) was also use to filter repetitive matches from the PROmer alignment. The mummerplots and coordinates of genomic regions were generated from the filtered output using mummerplot programs and show-coords, respectively. Estimates of the proportion of aligned genomic regions were computed for each reference sequence, from the filtered

coordinate's files. In addition, genome sequences from each *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* isolates were aligned to their respective reference using NUCmer (MUMmer 3.0) to determine genome coverage and single nucleotide polymorphisms (SNP), computed via MUMmer 3.0 show-snps [174]. Sequence length of less than 5kb were excluded from the synteny analysis. The genome sequences of *P. pinodes* isolate M074, which is considered as a reference species for the ABPFP, was aligned against genome from Dothideomycetes (*Parastagonospora nodorum, Pyrenophora tritici-repentis, Leptosphaeria maculans,* and *Cochliobolus sativus*), Sordariomycetes (*Fusarium graminearum, Magnaporthe oryzae, Fusarium verticillioides* and *Neurospora crassa)* to determine synteny relationship. The sources of the genome data were indicated in supplementary Table 1

Because *P. pinodes* and *P. pinodella* are closely related species, genome comparisons could be further generated employing a read mapping approach. The sequence reads from two isolates of *P. pinodella* were mapped to the assembled genome of *P. pinodes* reference isolate M074 using Bowtie 2 [177] in a sensitive mode. The mapped reads were processed and converted to BAM file using SAMtools [190]. Genome coverage was computed from the read mapping across non-overlapping intervals of 2 kb using BedTools coverageBed [251]. Intervals were considered absent from the reference if mapped reads covered (breadth) was less than 5% of the interval length [48]. Similar analyses were performed by mapping the genome sequences of the reference isolate of *P. pinodes* (M074) to each of the *P. pinodella* isolates. The genomic regions (Scaffolds) unique to *P. pinodella* relative to *P. pinodes* were then extracted and aligned against *P. nodorum, L. maculans* and *M. oryzae* using PROmer. Absence of matching sequence in *P. nodorum, L. maculans* and *M. oryzae* confirmed species or isolate specificity.

3.2.5.2 Analysis of secondary metabolite gene clusters

Plant pathogenic fungi produce a number of putative secondary metabolites during their life cycle. Polyketides, non-ribosomal peptides, terpenes, and alkaloids are the

major natural secondary metabolites (SM) produced by fungi [157]. The fungal toxin such as HC-toxin, AM-toxin and the extracellular coprogen siderophore produced by *C. carbonum, A. alternata* and *C. heterostrophus,* respectively are pathogenicity determinant products of NRPSs [36, 49, 230]. Similarly, the pathogenicity determinant T toxin in *C. heterostrophus* [142], the botcinic acid in *B. cinerea* [54], the mycotoxin alternariol in *A. alternata* (strawberry pathotype) [268] are the products of iterative type I PKS.

Recent comparative genomics has shown that genes encoding secondary metabolites are high in Pleosporales and there is high variation across species [229]. There was little information available for pathogens that belong to the Didymellaceae family with regard to SM gene clusters. Since fungi mainly contain type I PKS, the predicted ORFs from each pathogen species were analysed for PKS I and NRPS gene clusters. The PKS-I and NRPS backbone genes and modular organizations were identified by searching Pfam database [250] through HMMER3. An automated online antibiotics and secondary metabolites analysis tool (antiSMASH) [21] was also used for detection of homologous protein sequences. Each potential sequence involved in SM gene clusters was reanalysed via InterProScan [220] webserver (http://www.ebi.ac.uk/interpro/) as a confirmatory signature to functional domain organization.

3.2.5.3 Carbohydrate active enzyme (CAZymes) annotation

The CAZymes content of *P. pinodes, P. pinodella, Ph. Ph. koolunga* and *A. pisi* were assigned using database for automated carbohydrate-active enzyme annotation (dbCAN). The analysis was performed on the dbCAN web server (http://csbl.bmb.uga.edu/dbCAN/index.php) and parsed outputs are used for comparative analysis [343]. The CAZymes encoding domains were also identified by searching Pfam database through HMMER3 [250] as supportive evidence. Multiple hits to the same CAZymes domain on a single gene are only counted once, and the

existence of some CBM families associated to other CAZymes families (such as GH18) were ignored from absolute counts.

To establish if there are similarities or differences in pectin degradation between ABPFP and other pathogens infecting dicot plants in pectin degradation, we predicted and compared genes for their potential to encode pectin-degrading enzymes. All predicted pectin lyase (PL) family 3 encoding protein sequences were extracted from proteome database and aligned using CLUSTAL W. A phylogenetic tree was constructed using Molecular Evolutionary Genetics Analysis version 6 [307] to make inference about the relationship of PL3 across ABPFP and other plant pathogens.

3.3 Results and Discussion

3.3.1 Phylogenetic relatedness of ascochyta blight pathogens of field pea

The phylogenetic relationship among the ABPFP of field pea were analysed using concatenated sequences of beta-tubulin, RNA polymerase II subunit (RPB2) and lyceraldehyde-3-phosphate dehydrogenase (G3PD) predicted from the genome assemblies. As indicated in (Figure 5). *P. pinodes* and *P. pinodella* are closely related species with 100% bootstrap support and in agreement with previous report [11]. Similarly, *Ph. koolunga* and *A. pisi* are more closely related to each other with 100% bootstrap support as observed in [38]. Further comparison to Dothideomycetes and Sordariomycetes indicate the pathogens from the Didymellaceae are closely related to other pathogens in the Dothideomycetes, such as *P. nodorum and L. maculans* [11].The overall phylogenetic analysis corroborates previous findings [11, 38, 60].

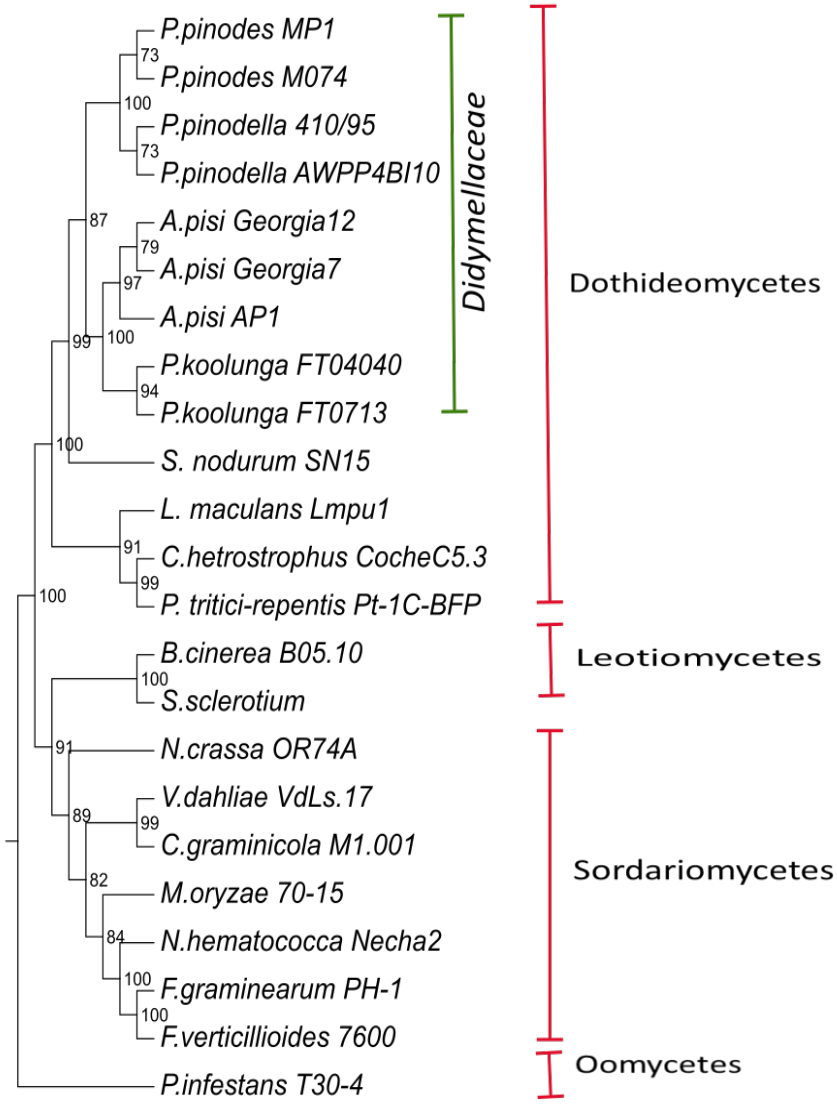


Figure 5 Species phylogenetic tree constructed using Neighbour-Joining method.

Analysis was conducted based on Poisson distribution of concatenated protein sequences of beta-tubulin, RNA polymerase II subunit (RPB2) and glyceraldehyde-3-phosphate dehydrogenase (G3PD). Values at the node indicate percent bootstrap obtained from 1,000 replicates. All ambiguous positions were removed for each sequence pair. There were a total of 895 positions in the final dataset. The tree is rooted to *Phytophthora infestans* as outgroup.

3.3.2 Genome sequence assembly statistics

Full genome sequences of various fungal plant pathogens in the order Pleosporales were published. Here, the first comprehensive genome sequences for *P. pinodes* and draft genome sequences for closely related species and their evaluation are presented. Genome sequences of the *P. pinodes* reference isolate was assembled in to 33.3 Mbp. The first 15 largest scaffolds (N50) varying from 0.82 Mbp to 1.53 Mbp in the genome assembly of *P. pinodes* is comparable size to chromosome estimates (11–12) [2]. Equilibrated GC (53.2%) content was observed in the largest scaffolds of the reference genome (*P. pinodes* M074), as opposed to regions of AT rich sequence such as previously observed throughout the *L. maculans* genome [266]. An overview of the genome sequences of *P. pinodes* reference isolate (M074) is shown in (Figure 6). The daft genome assemblies of *P. pinodes* isolate MP1 was assembled into equivalent size of 31.7 Mbp with average GC content of 53.4%.

Similarly, the first draft genome sequence assembly of isolates of *P. pinodella* (2), *Ph. koolunga* (2) *and Ascochyta pisi* (3) were assembled into comparable genome sizes spanning 32 – 33.9, 29.4 – 30.9 and 30.3 – 31.7 Mbp, respectively (Table 2). The average GC content is 53.3% in *P. pinodella* and varies from 51.9 – 52.2% in *Ph. koolunga* and 52.7 – 53% in *A. pisi* (Table 2).

The genome size of ABPFP pathogens are smaller than *M. graminicola* (39.7 Mbp) [111], *Botrytis cinerea* (39 – 42.3 Mbp) [7], *P. nodorum* (37.2 Mb) [128], *F. graminearum* (36.2 Mbp) and *Pyrenophora tritici*-repentis 37.84 Mbp [205] but equivalent to *Dothistroma septosporum* [66] and higher than the postharvest pathogen *Penicillium digitatum* [206]. The overall genome assemblies are estimated to represent 87 – 97% of core conserved genes based on the core eukaryotic genome mapping approach (CEGMA, [235]) indicating good capture of the gene contents in the genomes of these ABPFP pathogens. Nevertheless, genome assemblies of *P. pinodella*, *Ph. koolunga* and *A. pisi* are relatively fragmented and would require further improvement for bioinformatics analyses dependent on long sequences.

The average GC content in ABPFP is higher than observed in Leotiomycetes [7] and equivalent to *Magnoporthe oryzae* [340]. The relatively high and uniform GC contents across the majority of the ABPFP pathogen genomes may indicate a stable genomic landscape in the Didymellaceae that is free from the influence of repeat-induced point mutation (RIP). What is more likely however, is that these regions comprise the shorter length sequences, due to difficulties in the *de novo* assembly of repetitive and AT-rich sequences, as RIP is widespread across the filamentous Ascomycota [45].

3.3.3 Repeat content of the genome assembly

The distribution of repeat class (types) across *P. pinodes, P. pinodella, Ph. koolunga* and *Ascochyta pisi* genome sequences are similar but vary in the amount of repeats present in the genome. *Peyronellaea pinodes* isolate M074 and *P. pinodella* isolate AWPP4B1I0 contained 4.5% and 5% of repetitive elements, respectively Table 3). Other isolates in each species contain less than 3% repetitive elements. DNA transposons invasion is higher in *P. pinodes* and *Ph. koolunga* while *P. pinodella* isolate AWPP4BB1I0 and all isolates of *A. pisi* contain high retrotransposon. The amount of repetitive DNA identified in the ascochyta blight pathogens is generally comparable to *P. nodorum* (4.5%) but less than *S. sclerotiorum* (7.7%) and *Cladosporium fulvum* (47.2%) [7, 66, 128].

Table 2 Feature of genome sequence assemblies of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi*

Species	Isolate	Assembly length (Mbp)	Scaffold max. length (Kb)	Number of scaffolds			N50[ac]	L50[b]	Unknown N (%)	GC (%)
				>200bp	>5kb	>10kb				
*P. pinodes**	M074	33.3	1534.9	828	109	97	15	817,551	1.35	53.2
P. pinodes	MP1	31.7	143.9	2,795	1366	963	347	27,524	2.21	53.4
*P. pinodella***	410/95	32.0	600.3	1,158	512	427	100	102,821	0.08	53.3
P. pinodella	AWPP4b1I0	33.9	254.5	4,764	764	623	176	60,527	0.18	53.3
*Ph. koolunga***	FT04040	30.9	446.0	4,541	960	718	461	18,683	0.04	51.9
Ph. koolunga	FT0413	29.4	110.7	2,603	1505	944	201	43,333	0.69	52.2
A. pisi	Georgia-12	31.7	157.7	1,828	1036	800	234	41,289	0.06	53.0
A. pisi	Georgia-7	30.3	82.8	4,743	1001	1815	638	14,152	0.04	52.7
*A. pisi***	AP1	32.6	418.9	1953	552	461	106	95354	0.40	52.1

**P. pinodes* reference isolate. a N50: The number of scaffolds >=L50. b L50: weighted average scaffold length covering at least 50% of the scaffold length. c N50 indicated for total scaffolds greater than 200bp. ** Reference isolate for each species.

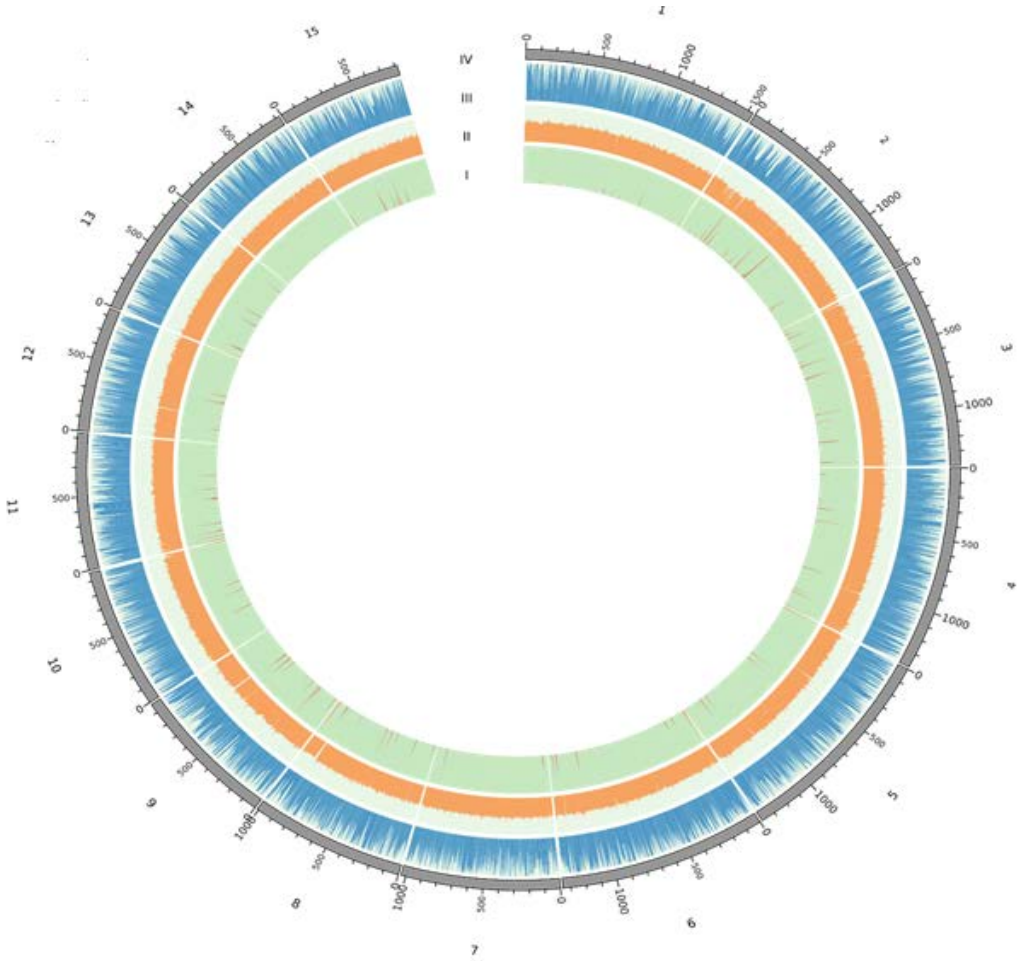


Figure 6 Overview of the largest 15 scaffolds of the *P. pinodes* (reference isolate M074).

Note that the distribution of repeats, GC and gene density across the N50 scaffolds are uniform across the genome. I, II, III and IV indicate distribution of repetitive elements, percent GC content, gene density, and the first N50 scaffolds, respectively.

Table 3 Repeat contents in genome sequences of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi*

Class/Family	Length Occupied (Kbp)								
	P. pinodes		*P. pinodella*		*Ph. koolunga*		*A. pisi*		
	M074	MP1	410/95	AWPP4B1I0	FT04040	FT0713	G-12	G-7	AP1
Retroelements	69.0	37.8	27.9	217.9	49.5	24.8	23.7	20.5	39.6
Penelope	0.1	0.0	0.0	0.05	0.1	0.07	0.0	0.07	0.4
LINEs	24.9	13.8	9.7	4.9	2.2	1.0	3.0	2.1	7.2
R1/LOA/Jockey	0.0	0.0	0.0	0.0	0.2	0.2	0.0	0.0	0.0
LTR elements	44.2	24	18.2	213	47.3	23.8	20.7	18.45	32.4
Ty1/Copia	8.5	5.8	10.3	67.8	19.2	8.8	7.7	8.06	11.3
Gypsy/DIRS1	35.6	18.1	7.8	145	28.0	15.0	13.0	10.4	21
DNA transposons	349.7	65.6	91.8	229.6	87.5	56.2	26.6	19.9	33.9
hobo-Activator	0.36	0.5	0.2	0.03	0.05	0.05	0.1	0.04	0.1
Tc1-IS630-Pogo	347.9	63.5	91	228.9	84.9	54	25.1	18	31.7
Tourist/Harbinger	0.2	0.37	0.1	0.18	1.4	0.95	0.4	0.53	0.5
Other	0.0	0.05	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Unclassified	0.07	0.2	0.1	0.06	0.6	0.77	0.5	0.0	0.5
Total interspersed repeat	418.8	103.6	119.8	447.6	137.7	81.8	50.8	40.5	74
Small RNA	11.8	4.6	4.9	14.3	7.2	3.8	6.6	4.47	7.1
Satellites	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.04	0.2
Simple repeats	160.4	128.6	160.6	209.5	366.7	305.6	341.9	334.1	453.5
Low complexity	20.1	15.9	18.1	21.4	25.8	23.3	29.3	267.7	31.8
Total	1491.6	482.4	560.5	1800.2	858.5	600.1	549.6	744.9	745.3
Percent	4.49	1.53	1.75	5.3	2.76	1.95	1.72	1.66	2.5

3.3.4 Gene content of genome assemblies

The number of predicted genes vary from 9,679 in *A. pisi* (Georgia_12) to 12,432 in *P. pinodes* (MP1) (Table 4). In the well-assembled *P. pinodes* reference isolate (M074), a total of 11,352 genes with a minimum translated length of 50 amino acids were predicted. About 75 – 83% of RNA-seq data derived from *in vitro* culture were mapped to the M074 genome, indicating that the assembly represents most of the transcribed genes. The number of predicted genes in *P. pinodes* M074 are larger than the 'finished' genome sequence of *Zymoseptoria tritici* (10, 933 genes) [111] but equivalent to the recent re-annotation of *P. nodurum* SN15 (12,383 [305]. Aside from the *P. pinodes* reference isolate M074, it is likely that the number of predicted genes are either under-estimated as in *A. pisi* (Georgia 12) or over-estimated as in *P. pinodes* MP1.

Table 4 Number of genes functionally annotated in ascochyta blight pathogens of field pea

Pathogen Species	Isolates	Total predicted orf*	Pfam	GO
P. pinodes	M074	11352	9318	7000
	MP1	12432	9613	6177
P. pinodella	410/95	11058	9083	6185
	AWPP4B1I0	11482	8585	6260
Ph. koolunga	FT04040	10084	8340	6488
	FT0713	10899	8746	5834
A. pisi	AP1	11260	9513	6018
	Georgia-7	10625	8412	5530
	Georgia-12	9679	7849	5715

*Only amino acids >=50 considered.

3.3.5 Comparative genomics

3.3.5.1 RIP like activities are unusual in ascochyta blight pathogens of field pea

RIP analysis of whole genome sequences indicated occurrences of RIP-like mutation with variable strength according to the species and repeat classes involved, which may be attributable to the relatively fragmented nature of the current assemblies. Both DNA transposons and long terminal repeat (LTR) retrotransposons were mutated. Directional mutation preferences of CT ↔ TT in both LRT and DNA transposon were noticed. Clear evidence is presented for LTR retrotransposon LMR1 in Figure 7.

Ascochyta blight pathogens of field pea showed different level of RIP strength. Strongest RIP-like activities in LTR retrotransposons were observed in *Ph. koolunga* isolate FT0713 and *P. pinodella* isolate AWPPB4I10 followed by *A. pisi* isolate AP1 and *P. pinodes* M074 Figure 8. *Phoma koolunga* and *A. pisi* showed relatively more RIP in DNA transposon than *P. pinodes* or *P. pinodella* while RIP in non–LTR retrotransposon is high in *Ph. koolunga* (FT04040) and *A. pisi (*AP1).

The strong RIP-like activities in *P. pinodella* and *Ph. koolunga* may reflect their better genome defence against invasion by repetitive elements. RIP is known as fungal defence strategy that recognizes and inactivates duplicated sequences [97, 98, 279]. The observed directional mutation in this study is also consistent with typical of RIP mutation [45, 97]. Previous research showed that RIP is common among ascomycete's fungi [45]. The RIP strength within a species varied with isolate origin. *P. pinodella* isolate from USA (WPPB4I10) had stronger RIP than isolate from Australia (410/95). On the other hand, *P. pinodes* isolate from USA (MP1) showed weak RIP than isolate from Australia (M074). This is partly associated to the high repeat content of *P. pinodella* isolates AWPPB4I10 than isolate 410/95. The result may also suggest geographically isolated pathogen species exposed to transposable element invasion evolved different strength of deactivation mechanism (RIP). The high RIP-like activities in DNA transposon of *P. pinodes* (MP1) *Ph. koolunga* and *A. pisi* (Georgia-

7 and Georgia-12) is likely related to the relatively high DNA transposons identified in the genome assemblies (Table 3).

Repeat induced point mutation index comparison between homothallic species *P. pinodes* and heterothallic species *P. pinodella* is similar and even higher in heterothallic species (Figure 8). *Peyronellaea pinodes* is expected to possess strong RIP index compared to *P. pinodella* or *A. pisi* because *P. pinodes* can undergo multiple sexual cycle in a season (See chapter 1) and RIP is active during sexual cycle [97]. This unexpected result may imply sexual cycles during disease epidemics within a season may have little impact on RIP occurrence in *P. pinodes. S*tudies in *Podospora anserina* indicated that RIP is detected in late maturing ascospores [115] suggesting early maturing ascospores within a season could escape RIP effect.

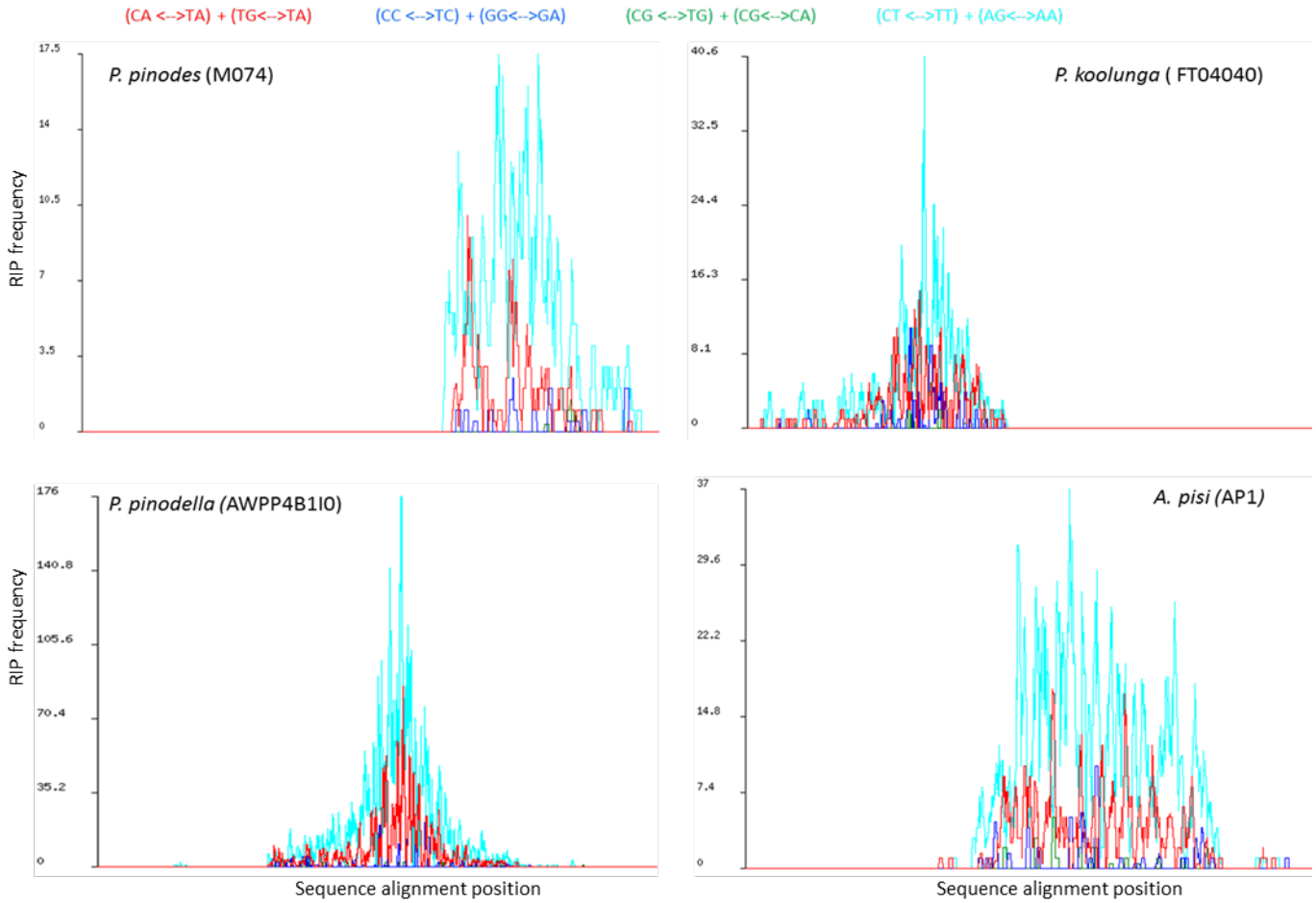


Figure 7 Evidence of RIP like mutations observed in ascochyta blight pathogens of field pea invaded with repeat family LMR1. Types of observed mutation in each species were indicated above the graph according to respective colour.

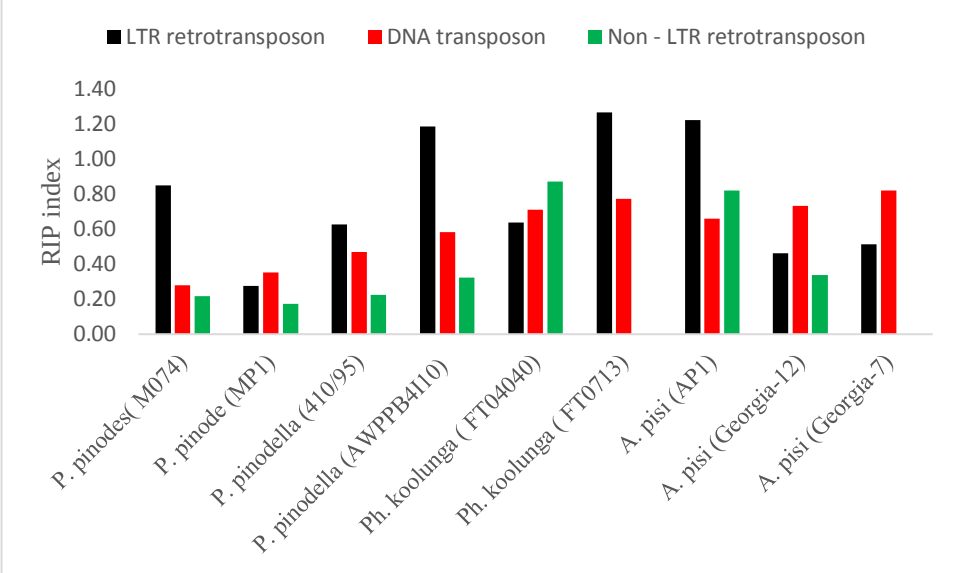


Figure 8 Dinucleotide mutation preferences observed in ascochyta blight pathogens of field pea.
The graph indicate CT ↔ TT transition index as computed by RIPCAL.

3.3.5.2 Synteny relationship between *P. pinodes* and other pathogens

The relationship between *P. pinodes* genomes and related fungal species were investigated through aligning their translated genome sequences via MUMmer (PROmer). Sequences of each species are laid out along the x- and y-axes and sequence matches are depicted as dotplots along diagonal lines. Dots appearing as a diagonal line and randomized 'dots within boxes' indicate chromosome-level synteny (shared gene-content in co-linear order between two sequences (syn. 'macrosynteny') and widespread intra-chromosomal rearrangements (syn. 'mesosynteny'), respectively [129]. The matching sequences between the genomes of *P. pinodes* reference isolate (M074) and both isolates of *P. pinodella* showed high level of matching sequences with up to 100% similarity. Bases on PROmeralignment, 92 – 94% of *P. pinodes* genome can be covered by sequences from *P. pinodella* isolates (Figure 9, Table 5). This showed that a large proportion of the genomic regions between *P. pinodes* and *P.*

pinodella are conserved and only about 6% of *P. pinodella* genomic regions contribute to its difference from *P. pinodes.*

Regardless of high sequence similarity, there are potential genome re-arrangements observed along assembled scaffolds. For instance, Scaffold_1 from *P. pinodella* isolate 410/95 is co-linear to Scaffold_5 and Scaffold_14 in *P. pinodes* (M074) with shuffled sequences (Figure 10; Table 1). It could be possible that genome re-arrangement at this particular locus was driven by transposable elements, as remnants of repeats such as Marriner_ABr, Gypsy and Molly were observed on *P. pinodes* M074 Scaffold_5 and Scaffold_14. As opposed to the *P. pinodes*, there were no repetitive elements detected on corresponding Scaffold_1 in *P. pinodella* (410/95).

Most sequence matches between *P. pinodes* and *Ph. koolunga* or *P. pinodes A. pisi* showed 80% sequence similarity forming continuous straight line across diagonals of the dot-plot matrix. The synteny relationship between *P. pinodes* and other pathogens from the Didymellaceae family, Dothideomycetes and Sordariomycetes is congruent to the established evolutionary relationship of the fungal pathogens considered in the analysis (Figure 9). Dot plot comparison of *P. pinodes* (M074) against pathogens from Dothideomycetes showed typical mesosynteny patterns. The level of similarity dropped to below 60% similarity with widespread inter-chromosomal rearrangements when compared to pathogens from class Dothideomycetes and Sordariomycetes. The observed mesosynteny within Dothideomycetes corroborate previous findings [129]. The synteny relationships between the genomes of *P. pinodes* and other fungal species follow expected phylogenetic pattern. Very closely related ABPFP showed macrosynteny, while more distantly related fungal species showed breakdown of macrosynteny (mesosynteny) (Figure 9).

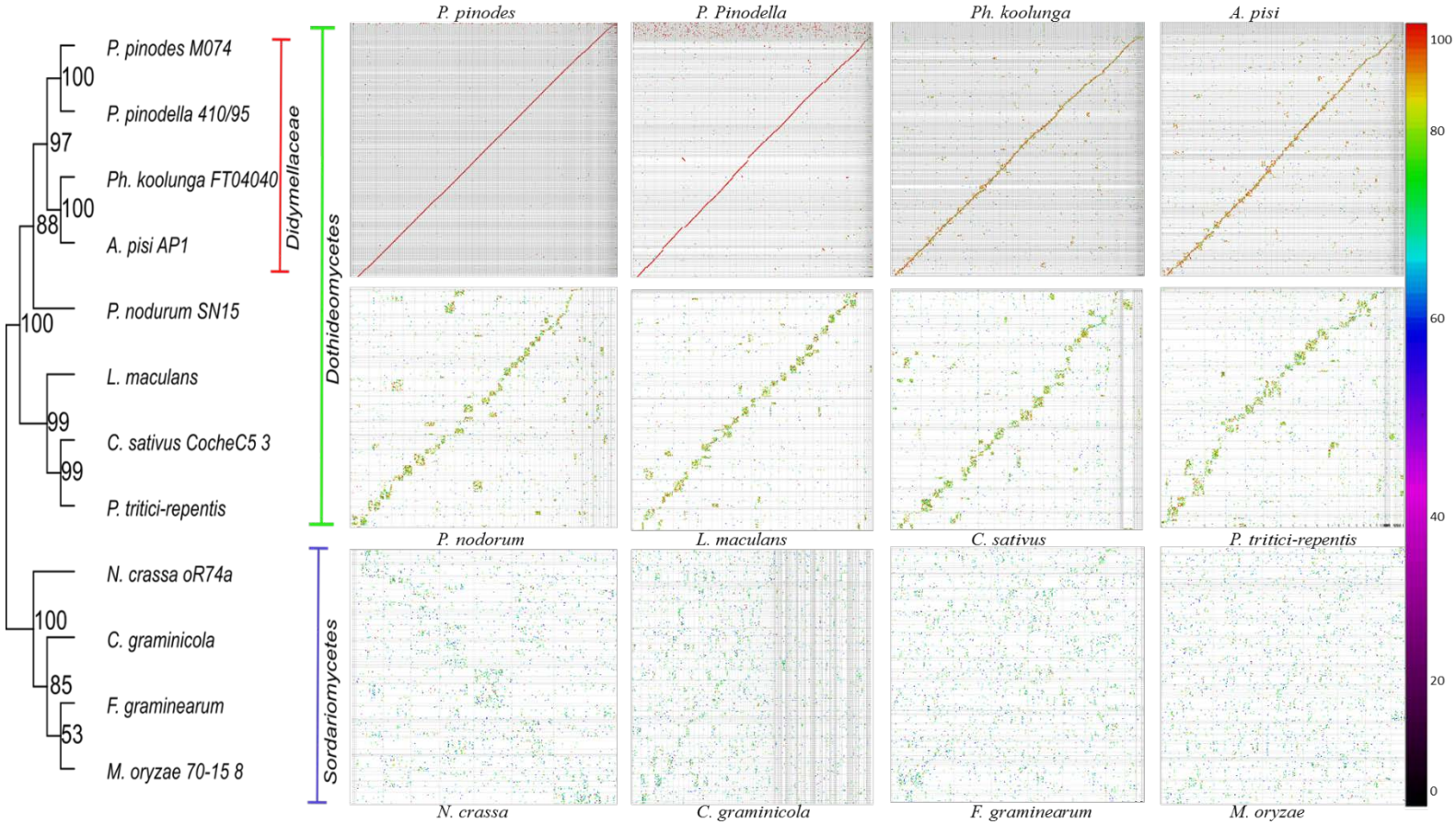


Figure 9 Six fram translation alignment comparison between *P. pinodes* reference isolate M074 genome to closely related fungal species in Didymellaceae and representatives from Dothideomycetes and Sordariomycetes.

Sequences of each species are laid out along the x- and y-axes and sequence matches are drown as colour coded dots according to their similarity level as indicated on right side of the graph. Dot plots appearing as a diagonal line and randomized dots within a boxes indicate chromosome-level synteny and widespread intra-chromosomal rearrangements, respectively.

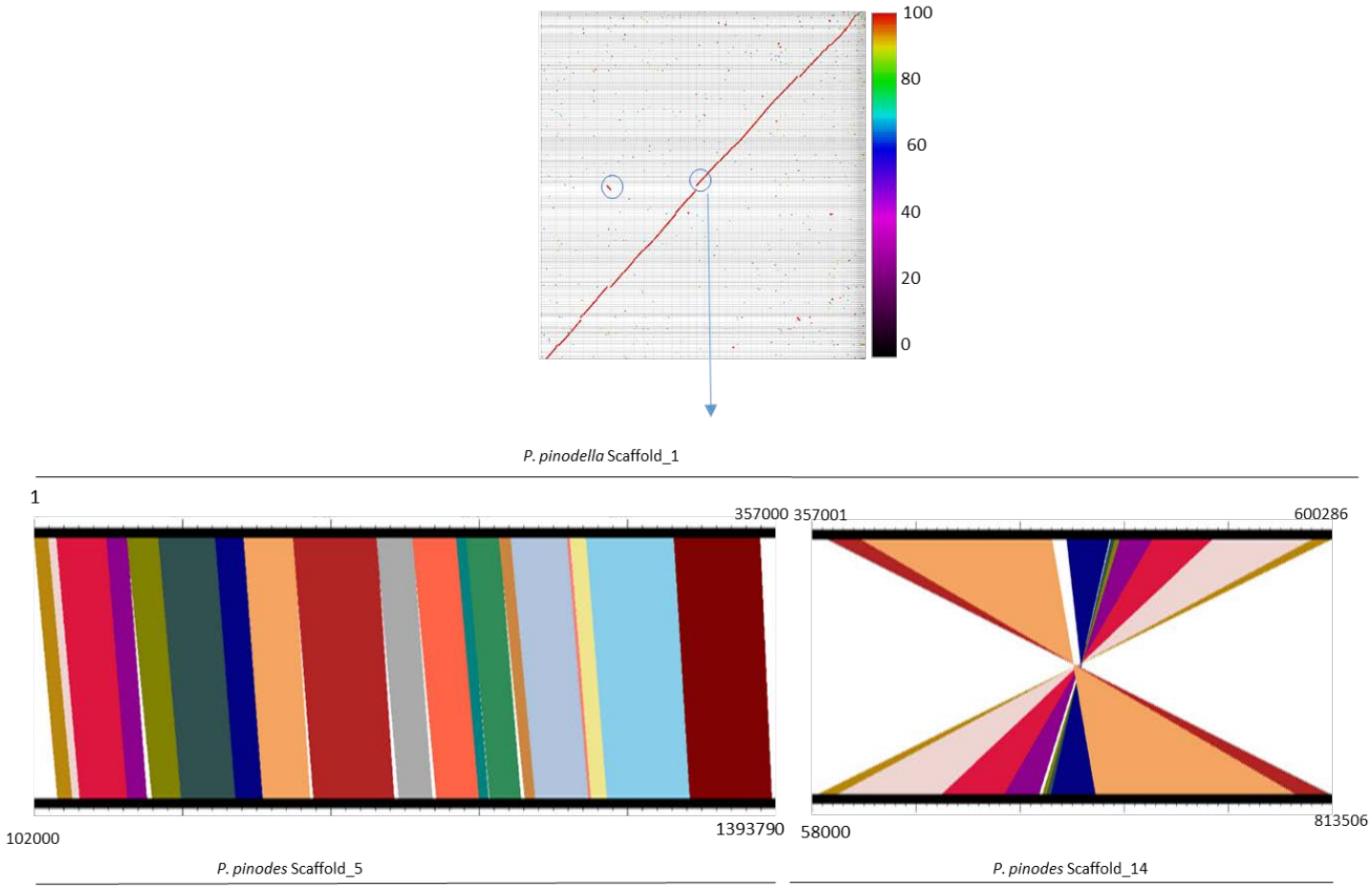


Figure 10 Genomic re-arrangements between sister species *P. pinodes* and *P. pinodella*. Upper line indicate Scaffold_1 from *P. pinodella* and lower line indicate Scaffold_5 and Scaffold_14 from *P. pinodes.*

Coloured block indicate sequence matches between the two species.

To investigate genome regions contributed to difference between *P. pinodes* and *P. pinodella*, unassembled reads from *P. pinodella* were mapped to *P. pinodes* genome and vice- versa. About 71 - 80% of the unassembled reads from *P. pinodella* can be mapped to *P. pinodes* reference isolate M074. This result was comparable to unassembled reads from *P. pinodes* isolate MP1 (83%) mapped to the same genome (Table 5). When *P. pinodes* unassembled reads were mapped to *P. pinodella* isolate 410/95, there about 1.1 – 1.4 Mbp were unique to *P. pinodella* (at 10% minimum read coverage cut-off) which is in agreement to the PROmeralignment analysis (Table 5, Table 6). The slight discrepancy between the PROmer alignment and the read mapping maybe related to exclusion of sequences less than 5 kb in PROmer analysis. Majority of scaffolds uncovered by read mapping are short scaffolds of less than 2kb (146 scaffolds) each consisting of one gene or no gene at all (Supplementary Table 2). The longest sequence region or scaffold carrying the highest number of genes not observed in *P. pinodes* contained fifteen predicted genes on Scaffold_236.

Similar mapping of reads from *P. pinodella* to *P. pinodes* isolate M074 showed absence of 41 – 43 genes on Scaffold_2. Interestingly, about 27% to 50% of predicted genes from *P. pinodella* genome specific regions are isolate/species specific (Table 6). Potential species/isolate specific regions identified in *P. pinodes* and *P. pinodella* were summarized in supplementary Table 3 and 4. Given the current assembly, the variation between *P. pinodes* and *P. pinodella* mainly relate to regions assembled into short scaffolds. These short scaffolds may explain the observed variation in genome size and chromosome number between *P. pinodes* (11 – 12 chromosomes) and *P. pinodella* (10 – 15 chromosomes)[2]. The result uncovered the existence of compartmentalized species specific loci in the genome of *P. pinodes* and *P. pinodella* that harbour species specific genomic regions.

Table 5 Comparison of *P. pinodes* reference isolate (M074) genome to the closely related species *P. pinodella*

	P. pinodella (410/95)	*P. pinodella* (AWPP4B1I0)	*P. pinodes* (MP1)	*P. pinodes* (M074)
PROmeralignment analysis				
Assembly length (Mbp)	31.98	33.88	31.72	33.3
PROmeralignment length	29.96	30.92	30.74	-
Percent alignment length	91.26	94.20	93.64	-
M074 read mapping analysis				-
Mapping (%)	71.25	79.80	83.20	-
Total regions not covered (Mbp)[a]	0.994	1.20	0.23	-
Total regions not covered (Mbp) [b]	1.10	1.40	0.23	-

a and b represent unassembled read mapping coverage at cut-off of 5% and 10%, respectively. Mapping coverage is computed in non-overlapping intervals of 2 kb using BEDTool suite “coverage”. Coverage is computed against M074 whole assembly.

Table 6 Species specific genome regions identified in *Peyronellaea* species using reciprocal unassembled read mapping approach

	Total number of Genes not covered*	Number of genes in the longest sequence region	The longest sequence region	#Isolate specific proteins	% Isolate specific
P. pinodella (410/95)[c]	508	15	Scaffold_236	252	49.6
*P. pinodella (*AWPP4B1I0) [c]	742	26	Scaffold_443	204	27.5
P. pinodes (M074)[d]	329 - 384	41- 43	Scaffold_2	142 - 161	41.9 - 43.2

c The reads from *P. pinodes* were mapped to *P. pinodella* (410/95 and AWPP4B1I0) to identify unique regions in *P. pinodella*

d The reads from *P. pinodella* were mapped to *P. pinodes* (M074) to identify unique regions in *P. pinodes*

Reads from M074 (3 libraries, 3, 8 & 20kb) mapped to both *P. pinodella* isolates.

3.3.5.3 Defining core protein sequences in Didymellaceae

Ascochyta blight pathogens core proteins were defined as protein sequences that are common to all isolates of *P. pinodes, P. pinodella, Ph. koolunga* and *A. pisi.* A total of 297, 042 proteome sequences from 24 fungal genomes were classified into 86,702 protein families using Proteinortho. Among these, a total of 4,555 protein families were found common to the four fungal pathogens complexes of field pea (Figure 11). These proteome common to ABPFP may determine their communalities of niche adaptation. *Peyronellaea pinodes* and *P. pinodella* shared the largest number of predicted protein sequences (>70%) while lowest number of proteins were shared between *P. pinodes* and *A. pisi* (Figure 11). This result is in accordance with the species phylogenetic tree where *P. pinodes* is more closely related to *P. pinodella* while *A. pisi* is distantly related to *P. pinodes* and *P. pinodella* [11, 38].

Further analysis indicated that about 109 protein families from 25 fungal pathogens were found specific to ABPFP. Functional assignment of these protein families based on *P. pinodes* reference species (M074) indicated that most of them are contain conserved Pfam domain associated to oxidoreductase, hydrolases, and transferases, peptidases, and kinases, transcription factors, production of toxin or protein domains of unknown function (Supporting Table 5). Most of these protein domain containing proteins are linked to fungal pathogenesis or domains of unknown function. For instance, glutathione S-transferase (GST; PF13417) are large family proteins involved in the degradation of xenobiotic compounds and oxidative stress responses [136, 209]. Recent research finding indicated that GST contribute to virulence of *A. brassicicola* to its host [35]. In general those protein sequences specific to the four species play a role in determining host specificity of the ABPFP pathogens determine host specificity. Furthermore, protein domains with unknown function would be a potential effector candidates for further evaluation.

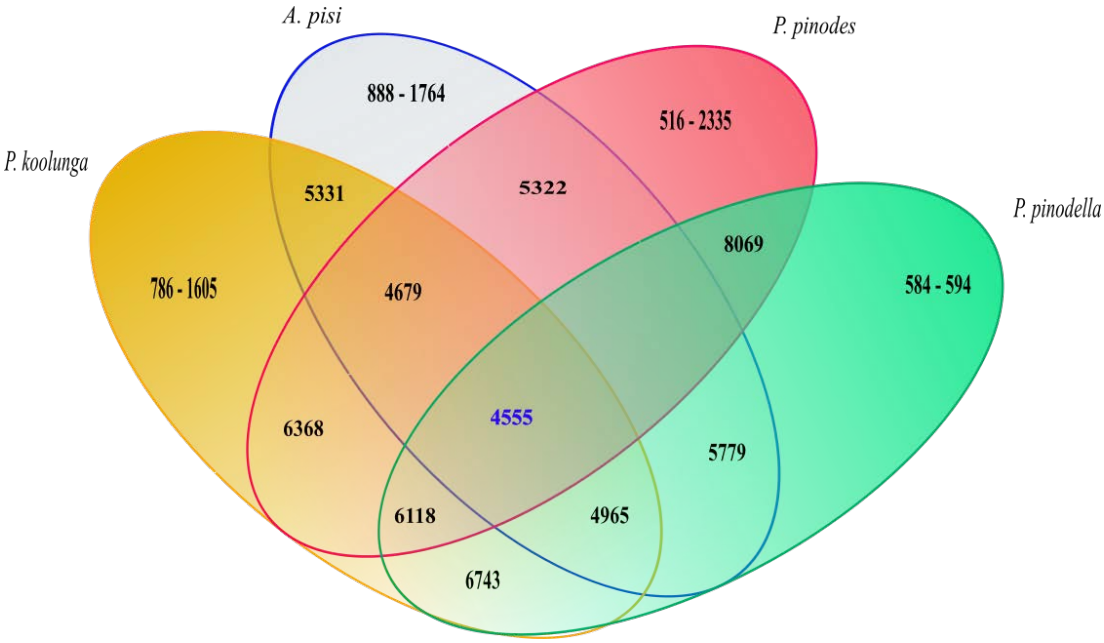


Figure 11 Orthology relationships among proteomes of ABPFPs.

Only proteins with ≥ 50 amino acid sequence length were included in the analysis. The defined core proteins (protein sequences common to *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi*) are indicated in blue font. A total of 2, 2, 2 and 3 isolates of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi*, respectively, included in the analysis.

3.3.5.4 Functional annotation abundance in *P. pinodes*

Conserved protein domain analysis of Pfam annotations indicated that over 8,000 of predicted protein sequences in each species matched conserved protein domains. In addition, over 6,000 of the predicted protein sequences were also assigned GO annotations (Table 4). According to Fisher's exact test, a number of Pfam domain-containing proteins were over-represented in *P. pinodes* compared to the overall average across *P. pinodes, P. pinodella*, *Ph. koolunga* and *A. pisi* (Table 7). These were mostly heterokaryon incompatibility protein (PF06985), cytochrome P450 (PF00067), helix-turn-helix (PF03221 and PF05225), fungal specific transcription factor (PF04082), NACHT domain (PF05729) and alpha/beta hydrolase (PF12695), which appeared to be expanded in gene number in *P. pinodes* relative to other ABPFP (Table 7).

Heterokaryon incompatibility proteins control the viability of heterokaryon after fusion of cells between and within individuals and strictly prevent heterothallism [275]. The enrichment of *P. pinodes* with heterokaryon incompatibility protein may be related to the homothallic nature of this pathogen compared to heterothallic species such as *P. pinodella* and *A. pisi*. Although there were no evidences on mating type studies in *Ph. koolunga,* low heterokaryon incompatibility protein content similar to *A. pisi* and *P. pinodella* may suggest likely existence of heterothallic mating type in *Ph. koolunga.*

Membrane transport proteins (PF07690), alpha/beta hydrolase (PF12697) were slightly higher in *P. pinodes* and *P. pinodella* than *Ph. koolunga* and *A. pisi*. Major facilitator family (MFS) proteins are known to regulate transport of solutes and toxic compounds across the cell membrane of an organism. While the MFS domain can relate to a broad range of biological functions, MFS genes *Bcmfs1* in *B. cinerea* [132] and *MgMfs1 in M. graminicola* [264] are involved in shielding the fungus from natural toxins and fungicides. Families of alpha/beta hydrolase proteins have diverse function including hydrolases, lyases, and transferases activities [187].

Table 7 Proteins domains over-represented in *P. pinodes* compared to average protein domains identified in Didymellaceae

PFAM	*P. pinodes*	*P value*	*P. pinodella*	*P value*	*Ph. koolunga*	*P value*	*A. pisi*	*P value*	Average Pfam	Pfam description
PF00067.17	115	0.02	96	0.06	80	0.03	94	0.06	97	Cytochrome P450
PF03184.14	49	0.00	11	0.03	4	0.00	1	0.00	21	DDE superfamily
PF03221.11	34	0.00	8	0.04	6	0.02	6	0.02	16	Helix-turn-helix
PF04082.13	137	0.01	117	0.05	104	0.03	91	0.01	119	Fungal specific TF domain
PF05225.11	17	0.01	1	0.05	1	0.05	2	0.11	6	Helix-turn-helix
PF05729.7	87	0.02	72	0.07	61	0.04	55	0.02	73	NACHT domain
PF06985.6	122	0.00	64	0.04	38	0.00	36	0.00	75	HET protein
PF07690.11	352	0.03	350	0.03	334	0.03	343	0.03	345	Major facilitator family
PF08659.5	156	0.04	143	0.05	130	0.04	145	0.05	143	KR domain
PF12417.3	6	0.01	6	0.21	0	0.02	2	0.23	4	Protein of unknown function
PF12695.2	146	0.04	137	0.05	126	0.04	137	0.05	136	Abhydrolase_5
PF12697.2	134	0.04	130	0.05	114	0.04	127	0.05	126	Alpha/beta hydrolase 6
PF13191.1	171	0.02	152	0.05	140	0.03	141	0.03	154	ATPases (AAA_16)
PF13358.1	16	0.02	4	0.12	3	0.08	2	0.04	8	DDE superfamily
PF13401.1	175	0.03	164	0.04	144	0.03	144	0.03	161	ATPases (AAA_22)
PF13460.1	114	0.04	104	0.06	90	0.04	105	0.05	102	NAD(P) H-binding
PF13498.1	5	0.01	1	0.38	0	0.09	0	0.09	2	DUF
PF13721.1	7	0.00	0	0.09	0	0.09	1	0.38	2	SecD export protein N terminal

**P* values are *computed P values* according to Fisher's exact test. HET: Heterokaryon incompatibility, TF Transcription Factor, DUF Domain of unknown function

The high structural DNA binding motif observed in *P. pinodes* relative to other pathogens of field pea could correlate to enrichment of DNA transposons encoding DDE superfamily endonuclease (PF13358.1 and PF03184.14) observed in the Pfam domain analysis as well as repeats identified in the genome sequences of *P. pinodes* (Table 3, Table 5).

3.3.5.5 Secondary metabolite gene clusters in ascochyta blight pathogens of field pea

The genome sequences revealed the potential of ABPFP to produce various secondary metabolites using genes encoded by multimodular type I polyketide synthase (PKS-I). There were about 4 to 9 PKS-I gene clusters identified from the draft genome sequences of ABPFP (Supplementary Table 6). Sequence analysis at NCBI indicated that one of the PKS-I found in all isolates of ABPFP has 49 - 59% identity to *L. maculans* PKS2 and 78 - 80% identity to *A. alternata* (PksA). The organization of the former cluster of genes in *P. pinodes* and *P. pinodella* were indicated in Figure 12. *Peyronellaea pinodes, P. pinodella* and *A. pisi* genome also carry a PKS-I homologous to *P. nodorum PKS* (52% identity at protein level) responsible for the production of mellein. The genomes of both *Ph. koolunga* isolates consisted of two PKS-I with high identity to *P. zeae-maydis* (70%) and *C. heterostrophus* PKS2 (66%) identity at amino acid level (Table 8).

In *L. maculans*, PKS6 is involved in phomonic acid production [266]. While not required for virulence of the pathogen to infect canola, it has been suggested that PKS6 plays a role in prevention from other invaders [266]. Similarly, recent studies showed that *P. nodorum* contain PKS associated to mellein synthesis has no impact on fungal virulence on wheat [40, 330]. On the other hands, the PKS2 identified from *A. alternata* is involved in the biosynthesis of melanin. Melanin is reported to be involved in pathogenicity and adaptation to host niche in both plant and animal pathogenic fungi [110, 162, 176, 306]. One PKS involved in melanin synthesis is reported to be conserved across Dothideomycetes [229]. Homologous PKS-I identified in *P. zeae-*

maydis and *C. heterostrophus* are known to produce the pathogenicity determinant PM-toxin [345] and T-toxin (PKS2) [15], respectively. The presence of homologs gene clusters in closely related ABPFP may shed light on their capabilities to produce similar product which are likely to be involved in pathogenesis or assist in their adaptation or provide competitive advantage over other microbes.

Each ABPFP genome also possessed protein domain with distinctive characteristics of NRPSs as observed in other fungal plant pathogens. These fungal pathogens consisted of at least 2 (*P. pinodes* and *P. pinodella*) to 3 NRPS (*Ph. koolunga*) NRPS gene clusters in their genome (Supplementary Table 7). Sequence analysis of some putative NRPS encoding genes showed high identity to *A. alternata* NPS6 in *A. pisi* (68%), *Ph. koolunga* (70%) and *Peyronellaea* species (76%) at amino acid level (Table 9). These homologous genes contain a single adenylation domain (A), three thiolation domains (ACP) and two condensation domains (C) required for NRPS production. Similar NRPS encoding genes identified in *C. heterostrophus* (NPS2, NPS6), *A. brassicicola* (NPS6) and *F. graminaerum* (NPS6, SID1) are known to be involved in siderophore biosynthesis [36, 117, 230]. Siderophore is required for virulence and oxidative stress tolerance via iron uptake mediated siderophore production [145, 186, 230].

The orthologous genes observed in genome of ABPFP likely encode NRPS with similar conserved molecular function. Previous studies showed that introduction of NPS6 homologs from saprophytic *N. crassa* into mutant phenotype of *C. heterostrophus* restored virulence on maize and tolerance to H_2O_2, indicating preserved activities across fungi [230]. Previous studies suggested that accumulation of reactive oxygen species play role in resistance to *P. pinodes* in resistant pea line compared to susceptible one [86]. The ability of the ABPFP to perpetuate under oxidative stress is likely related to their siderophore production. Most known fungal pathogens that produce siderophore as part of their virulence determinants belong to necrotrophic life style [230].

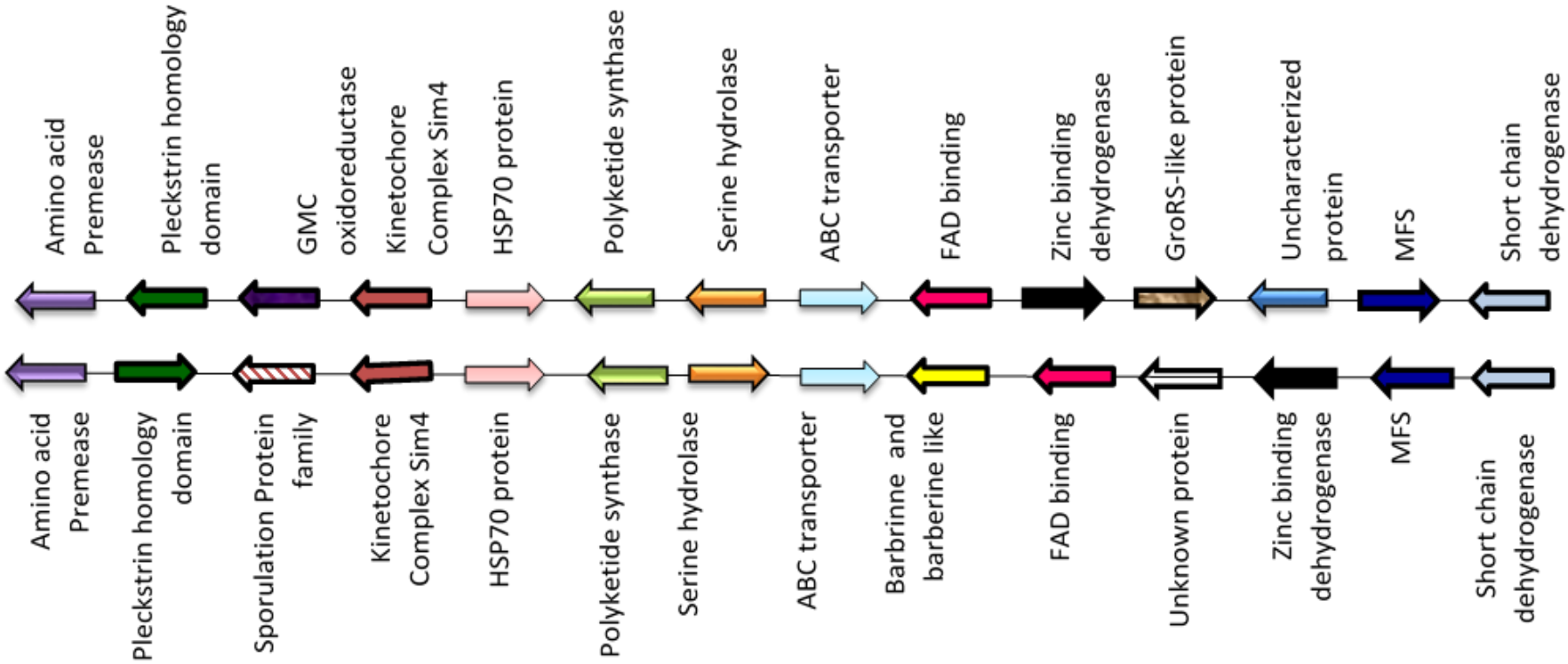


Figure 12 Co-linear organization of polyketide synthase gene cluster in *P. pinodes* and *P. pinodella*.

Upper and lower plots indicate gene order and orientation of *P. pinodes* (M074 Scaffold_30) and *P. pinodella* (410/95), respectively. Orthologous genes are indicated by arrow of same colour. Arrow head indicate direction of the transcripts.

Table 8 Some putative type I PKS encoding genes identified in ascochyta blight pathogens of field pea

Organism	Loci ID	Modular in Query	Gene Bank ID	Identity (%)	Known gene/ predicted product	*Organism*	Modular in Subject
P. pinodes	Ppdes_30.68	KS AT DH ER KR ACP	AGC79958.1	59	Phomonic acid	*Leptosphaeria maculans*	KS-AT-DH-ER-KR-ACP
(M074)	Ppdes_57.33	KS AT DH ER KR ACP	AIA58899.1	38	Brefeldin A	*Penicillium brefeldianum*	KS AT DH ER KR ACP
	Ppdes_57.36	KS AT DH ER KR ACP	XP_007792497.1	50*	unknown	*Eutypa lata UCREL1*	KS AT DH ER ACP
	Ppdes_14.166	KS AT ACP TE	AFN68292.1	80	Melanin (PksA)	*Alternaria alternata*	KS AT ACP TE
	Ppdes_16.63	KS AT DH KR ACP	AIW00670.1	52	mellein	*Parastagonospora nodorum*	KS AT DH KR ACP
	Ppdes_23.93	KS AT ACP TE	XP_001939987.1	46*	unknown***	*Pyrenophora tritici-repentis*	KS AT DH TE
	Ppdes_30.82	KS AT DH KR ACP	AAV66110.2	45	fusaridione A	Fusarium heterosporum	KS AT DH KR ACP ACP ACP
P.pinodella	Pdlla_2691	KS AT DH KR ACP	AIW00670.1	52	mellein	*Parastagonospora nodorum*	KS AT DH KR ACP
(410/95)	Pdlla_448	KS AT ACP ACP TE	AFN68292.1	80	Melanin (PksA)	*Alternaria alternata*	KS AT ACP ACP TE
	Pdlla_2197	KS AT ACP TE	XP_001939987.1	46	unknown***	*Pyrenophora tritici-repentis*	KS AT DH ACP TE
	Pdlla_6178	KS AT DH ER KR	AGC79958.1	59	Phomenoic acid	*Leptosphaeria maculans*	KS-AT-DH-ER-KR-ACP
	Pdlla_9495	KS AT DH KR	AGO86662.1	42	equisetin	*Fusarium heterosporum*	KS AT DH KR ACP ACP
	Pdlla_9801	KS AT DH ER KR	XP_007792497.1	50	unknown	*Eutypa lata UCREL1*	KS AT DH ER ACP

Table 8 continued…

Organism	Loci ID	Modular in Querry	Gene Bank ID	Identity (%)	Known gene/ predicted product	*Organism*	Modular in Subject
Ph. koolunga	Phkol_3871	KS AT DH ER KR ACP	AGC79958.1	49	Phomenoic acid	*Leptosphaeria maculans*	KS AT DH ER KR ACP
(FT04040)	Phkol_6434	KS AT DH ER KR ACP	ABB76806.1	66	T-toxin (PKS2)	*Cochliobolus heterostrophus*	KS AT DH ER KR ACP
	Phkol_3862	KS AT ACP ACP TE	AFN68292.1	79	Melanin (PksA)	*Alternaria alternate(PksA)*	KS AT ACP ACP TE
	Phkol_6027	KS AT DH ER KR	ACM42412.1	46	Radicicol	*Chaetomium chiversii*	KS AT DH ER KR
	Phkol_2655	KS AT ACP	AAR90258.1	73*	polyketide synthase	*Bipolaris maydis*	KS AT DH ER KR ACP
	Phkol_6429	KS AT DH	AAR85531.1	70	PM-toxin	*Peyronellaea zeae-maydis*	KS AT DH ER KR ACP
A.pisi	Apisi_2241	AT DH ER KR ACP	AGC79958.1	58	Phomenoic acid	*Leptosphaeria maculans*	KS-AT-DH-ER-KR-ACP
(Georgia 7)	Apisi_3671	KS AT DH KR ACP	AIW00670.1	52	mellein	*Parastagonospora nodorum*	KS AT DH KR ACP
	Apisi_53	KS AT ACP ACP TE	AFN68292.1	78	Melanin (PksA)	*Alternaria alternata*	KS AT ACP ACP TE
	Apisi_2617	KS AT DH ACP	KFA70666.1	44	Atr6	*Stachybotrys chartarum*	KS AT DH ER KR
	Apisi_7511**	AT ER KR ACP	AFN68292.1	37	Melanin (PksA)	*Alternaria alternata*	KS AT ACP ACP TE

*Most similar hit was considered when there is no characterized product found in the top 100 blast hit at NCBI. BLAST hit E value 0.0. ** KS domain truncated. *** Conidial yellow pigment biosynthesis polyketide synthase. KS= β-ketoacyl synthase; AT=acyltransferase; DH=dehydratase; ER=enoyl reductase; KR = ketreductase; ACP=acyl carrier protein

Table 9 Some putative NRPS encoding genes identified in ascochyta blight pathogens of field pea

Organism/isolate	Loci ID	Gene Bank ID	Identity (%)	Known product/ predicted product	Organism
P. pinodes (M074)	*Ppdes _6.236*	AFN69082.1	76	Siderophore	*Alternaria alternata*
*P. pinodes (*MP1*)*	*Ppdes _6179*	AAX09984.1	56	Siderophore	*Bipolaris maydis*
P. pinodella (410/95)	*Pdlla_6336*	AFN69082.1	68	Siderophore	*Alternaria alternata*
P. pinodella (410/95)	*Pdlla_7784*	AAX09984.1	54	siderophore	*Bipolaris maydis*
P. pinodella (AWPP4B1I0)	*Pdlla_4823*	AFN69082.1	76	NPS6/Siderophore	*Alternaria alternata*
P. pinodella (AWPP4B1I0*)*	*Pdlla_1211*	AAX09984.1	54	Siderophore	*Bipolaris maydis*
Ph. koolunga (FT04040)	*Phkol_2906*	FN69082.1	70	Siderophore	*Alternaria alternata*
Ph. koolunga (FT04040)	*Phkol_6050*	AAX09984.1	55	Siderophore	*Bipolaris maydis*
Ph. koolunga (FT04040*)*	*Phkol_7592*	KGQ06423.1	47	Bacitracin synthase 3	*Beauveria bassiana*
*A. pis*i (Georgia 7)	*Apisi_391*	AFN69082.1	68	Siderophore	*Alternaria alternata*
A. pisi (Georgia 12)	*Apisi_4327*	ABU42595.1	46	NRPS 2	*Alternaria brassicicola*

3.3.5.6 Genomes of Didymellaceae, pathogens of field pea are enriched with carbohydrate active enzymes

The genomes of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* pathogens encode similar number of genes involved in carbohydrate metabolism with slight differences according to their clade. The *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* contained 562 – 581, 592 – 600, 511– 520 and 503 – 509 CAZymes, respectively in order of decreasing total number of CAZymes (Table 10, Supplementary Table 8). *Ph. koolunga* and *A. pisi* consisted of lower number of GT families than *P. pinodes* and *P. pinodella.* The genes encoding putative CAZymes in the genomes of ABPFP were compared to other pathogens from the Dothideomycetes and Sordariomycetes. The number of putative CAZymes encoding genes in *Peyronellaea* spp. are higher than pathogens of broad host range (e.g. *B. cinerea* and *S. sclerotiorum*) and less than the hemibiotrophic pathogen *F. verticillioides* (651) and *N. haematococca* (777), but equivalent to cereal pathogens such as *Colletotrichum graminicola* (578), *F. graminearum* (561). Interestingly, the genomes of ABPFP contain higher number of pectin degrading enzymes than all the pathogens compared except *N. haematococca* (36); a necrotrophic pathogen that infect field pea (Table 11). The expansion of PL gene families acting on pectin in genome sequences of vascular wilt pathogen in *Verticillium* species has been reported [168]. The pectin degrading enzymes observed in the genomes of ABPFP may partially substantiate previous findings. Plant pathogens infecting dicot reported to contain high number of pectin degrading enzymes than pathogens of monocot plants [229]. Dicot plants contain 20 - 35% pectin as compared to 5% in monocots [325]. This host differentiation alone could not explain the high CAZymes content because even pathogens with broad host range contain lower number of CAZymes than ABPFP.

Table 10 Comparison of CAZymes contents across fungal species with diverse life style

Species	Mode of infection	Host preference	GH	PL	CE	CBM	GT	Total
Botrytis cinerea	Necrotrophic	Dicot	243	11	118	15	103	490
Sclerotinia sclerotiorum	Necrotrophic	Dicot	223	5	99	22	91	440
Alternaria brassicicola	Necrotrophic	Dicot	226	23	117	24	89	479
Leptosphaeria maculans	Hemibiotrophic	Dicot	214	18	109	16	99	456
Nectria haematococca	Necrotrophic	Dicot	368	34	194	37	118	751
Cochliobolus heterostrophus	Necrotrophic	Monocot	256	15	132	27	96	526
Magnaporthe oryzae	Hemibiotrophic	Monocot	265	5	136	38	100	544
Colletotrichum graminicola	Hemibiotrophic	Monocot	293	15	137	31	102	578
Fusarium graminearum	*Hemibiotrophic*	Monocot	261	21	140	33	106	561
Fusarium verticillioides	Hemibiotrophic	Monocot	309	25	175	36	106	651
Mycosphaerella fijiensis	Hemibiotrophic	Monocot	244	6	121	15	106	492
Pyrenophora tritici-repentis	Necrotrophic	Monocot	239	10	123	16	104	492
Parastagonospora nodorum	Necrotrophic	Monocot	257	10	140	24	95	526
Neurospora crassa	Saprophytic	Saprophytic	196	4	75	15	95	385
Peyronellaea pinodes	Necrotrophic	Dicot	287	32	139	22	101	581
Peyronellaea pinodes	Necrotrophic	Dicot	282	31	126	27	96	562
Peyronellaea pinodella	Necrotrophic	Dicot	298	31	141	23	107	600
Peyronellaea pinodella	Necrotrophic	Dicot	292	31	142	20	107	592
Phoma koolunga	Necrotrophic	Dicot	260	31	113	18	89	511
Phoma koolunga	Necrotrophic	Dicot	258	31	123	22	86	520
Ascochyta pisi	Necrotrophic	Dicot	256	33	109	21	90	509
Ascochyta pisi	Necrotrophic	Dicot	268	33	101	15	86	503
Ascochyta rabiei	Necrotrophic	Dicot	262	31	126	19	95	533
Ascochyta lentis	Necrotrophic	Dicot	281	34	142	24	102	583
Phoma medicaginis	Necrotrophic	Dicot	264	31	138	19	97	549

GH Glycoside hydrolase; PL Polysaccharide lyase; CE Carbohydrate esterase; CBM Carbohydrate – binding modules; GT Glycosyle transferase

3.3.5.7 Pectin degradation is a common host adaptation mechanism in ascochyta blight pathogens of field pea

In silico analysis revealed the genome of *P. pinodes* encode at least two PL3 highly similar to the PL3 characterized in pea pathogen *N. haematococca* and *Colletotrichum gloesporioides*. In comparative analysis, the PL encoding genes, particularly, the PL1 and PL3 are expanded in Didymellaceae. Similarly, the PL11 gene families are observed only in ABPFP pathogens and in *N. haematococca*. This family is also reported to be present in *V. dahliae*. On the other hand, the broad host range *B. cinerea* is deficient in PL3 families. Conversely, ABPFP contain lower number of (<=12) GH family 28 specifically acting on pectin. This indicate one of the different mechanisms fungal pathogens may deploy to breakdown the same substrate of their host plants.

Pathogen Host Interaction database (PHI-base) search was conducted to identify potential homologous genes known to be involved in fungal pathogenesis. Analysis of the reference isolate M074 using reciprocal best BLAST hit (RBBH) returned a well characterized gene encoding PL family from *N. haematococca* and *C. gloesporioides* (causal agent of anthracnose on Avocado). The phylogenetic analysis of the protein sequences predicted to encode pectin lyases 3 (PL3) from ABPFP were clustered with that of pectate lyases (A, B, C, D) from *N. haematococca* and *C. gloesporioides* (Figure 13). Importantly, the genes from each ABPFP that formed a unique clade together with genes from *N. haematococca* are very similar (paralogues) within a species. The *N. haematococca* encodes at least four PLs two of which are very similar. Previous studies of double mutants of *pelA* and *PelD in N. haematococca* indicated dramatic reduction of virulence, which could be recovered through supplementation suggesting its role as virulence factor on pea [263]. In *C. gloesporioides* pectin lyase (PLB) is identified as a virulence factor on avocado fruit [341]. Similarly, the *C. lindemuthianum Clpnl2* gene is related to pectin lyase and involved in fungal virulence [179]. Recent finding also indicated that *Alternaria alternata Pl1332*, which encode pectate lyase is required for virulence of its host [39]. As shown in Figure14, the PL genes closely associated to *N. haematococca* PLs (PLC, PLA or PLB), existed in

duplicate and such redundancy is relevant for virulence to pea as single gene disruption has no impact on virulence [263].

To complement the relevance of PLs during infection and subsequent disease development, RNA sequence analysis generated from *P. pinodes* time point infection assay used in candidate effector gene idenytification was evaluated (see chapter 6 for further information).The result indicated that 28% of the predicted genes are expressed at early stage while 62% of the genes are expressed at medium or late stage of disease development. Most pair of orthologues proteins encoding PLs were also co-expressed in at least two time points in *P. pinodes* and *P. pinodella* during *in planta* infection studies. Similar pattern of up-regulation of CAZymes encoding genes observed in *B. cinerea* and *S. sclerotiorum* [7].

The early expression of PL genes, its high similarity to the known *N. haematococca* PLs and its extracellular localization suggest ABPFP may deploy pectinolytic enzymes to overcome host defence reaction. Some of the PLs in *P. pinodes* (M074) are constitutively expressed while others are induced at latter infection stages and disease development. Absence of expression at early infection stage may not rule out the importance PLs in pathogenesis. Carrying genes with redundant function has a cost to the pathogen for energy expenditures but could be necessary for re-activation of the genes when required. The existence of inducible genes with redundant function has been demonstrated in various pathogens that use CWD arsenals during pathogenesis [263]. In ABPFP, only few PL encoding genes were expressed at early stage while most of the genes activated at mid and late stages; suggesting that most of the PL of ABPFP may be inducible in response to host as observed in *N. haematococca* - pea interaction. The typical symptom appearance on succulent shoot tips of pea before 24 HPI, may also suggest the fungus could use CWDE such as PLs to overcome host response.

Table 11 Comparison of polysaccharide lyase encoding genes from ascochyta blight pathogens of field pea and other pathogenic fungal species

Organism	PL1	PL11	PL14	PL15	PL20	PL22	PL3	PL4	PL7	PL9	GH28	GH43	GH78	GH88	GH95	GH105	GH115
P. pinodes M074	11	1	1	0	0	0	12	5	0	2	11	21	9	2	2	6	2
P. pinodes MP1	11	1	1	0	0	0	11	5	0	2	11	21	9	2	2	6	3
P. pinodella 410/95	11	1	0	0	0	0	11	5	1	2	12	21	10	2	2	6	2
P.pinodella AWPP4B110	11	1	0	0	0	0	11	5	1	2	12	21	9	2	2	6	2
Ph.koolunga FT04040	13	0	0	0	0	0	11	4	1	2	12	18	4	2	2	6	1
Ph.koolunga FT0713	13	0	0	0	0	0	11	4	1	2	12	18	5	2	1	6	1
A. pisi Georgia -7	14	0	0	0	0	1	10	6	0	2	8	20	8	1	1	5	1
A. pisi Georgia 12	14	0	0	0	0	1	10	6	0	2	9	21	7	2	1	5	1
A. rabiei	12	0	0	1	0	0	10	6	0	2	11	18	5	1	2	6	1
A. lentis	14	0	0	0	0	1	10	6	1	2	10	21	9	2	2	6	1
B. cinerea	7	0	0	0	0	1	2	0	1	0	19	6	7	1	3	2	1
C. sativus	6	0	0	0	0	0	5	4	0	0	4	15	3	1	2	2	2
L. maculans	8	0	0	0	0	0	6	4	0	0	7	10	1	1	1	3	1
M. oryzae	2	0	0	0	1	0	1	1	0	0	4	18	4	1	1	3	3
P. tritici-repentis	3	0	0	0	0	0	3	4	0	0	6	15	4	1	2	3	1
S. sclerotiorum	4	0	0	0	0	0	0	0	1	0	16	5	4	0	2	1	1
A. brassicola	8	0	0	0	0	0	11	3	0	1	7	17	1	1	1	4	2
C. graminicola	7	0	0	0	0	0	4	3	0	1	7	16	4	1	1	5	2
F. graminearum	9	0	0	0	1	0	7	3	0	1	6	18	7	1	2	3	2
F. verticillioides	11	0	0	0	0	1	7	3	0	3	8	21	8	2	1	3	2
M. fijiensis	2	0	0	0	0	1	1	2	0	0	5	15	8	0	1	4	3
P. nodorum	4	0	0	0	0	0	2	4	0	0	4	12	4	1	2	2	1
N. haematococca	14	1	0	0	1	1	11	6	1	1	11	32	11	5	3	4	2
C. fulvum	3	0	0	0	0	1	3	2	1	0	14	20	5	2	0	6	1

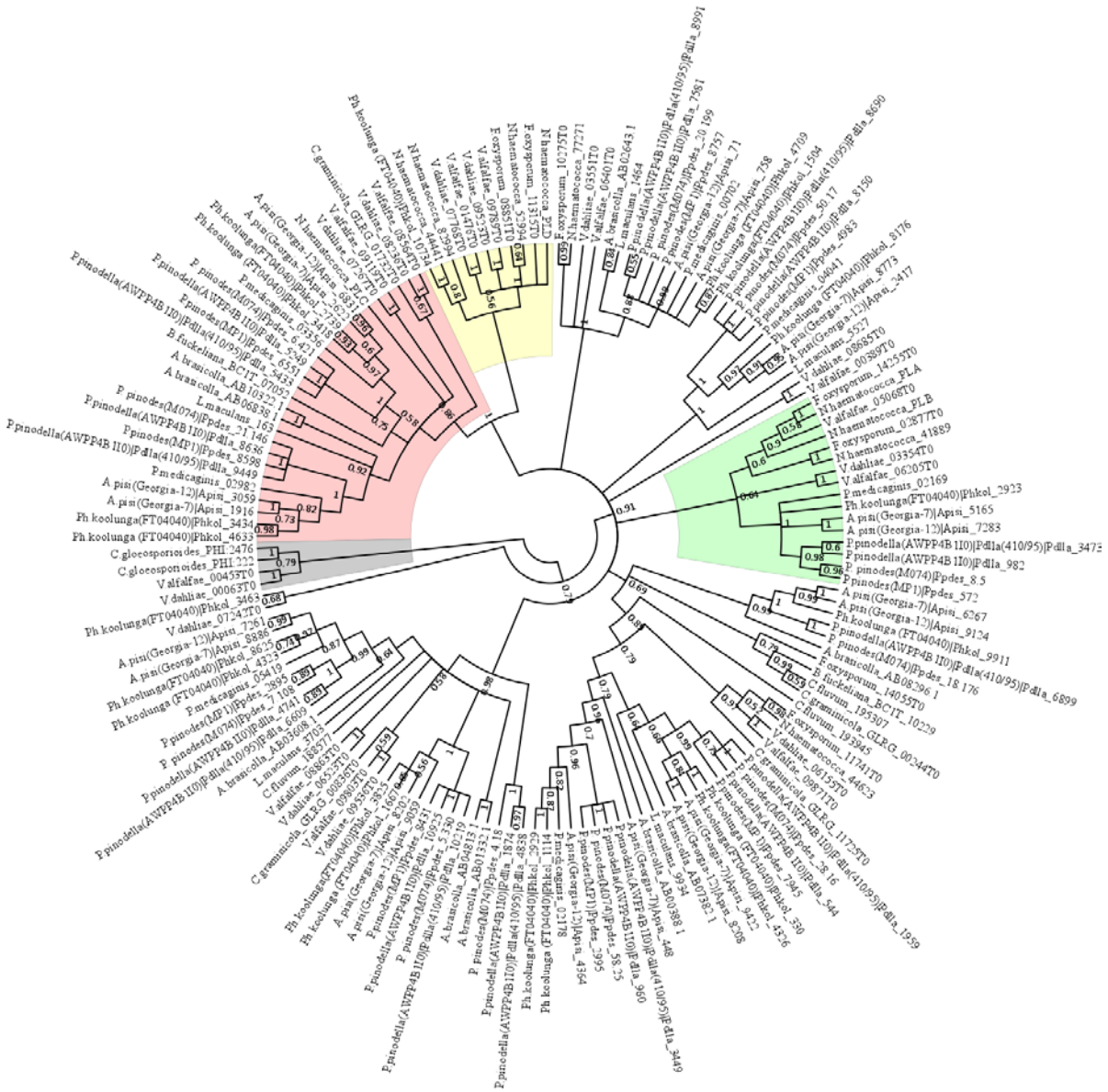


Figure 13 Phylogenetic analysis of pectin lyase 3 in some fungal species.

Phylogenetic tree was constructed from 153 amino acid sequences using the Neighbor-Joining method following Poisson correction after elimination of positions with less than 95% coverage. The numbers indicate the bootstrap support from 1000 pseudo replicates and branched with less than 0.50 bootstrap support were collapsed. Coloured branches indicate the position of functionally characterized PLs from *N. haematococca* ((PLA and PLB (light green), PLC (brown) and PLD (yellow)) and *C. gloesporioides* (grey).

3.4 Conclusion

The genome sequences of pathogens that cause ascochyta blight in field pea showed similar assembly sizes equivalent to those predicted using pulse field gel electrophoresis [2]. Protein family analysis indicated significant amount of genes shared among ABPFP which may imply their common strategies to adapt to same host species. The identification of necrosis inducing proteins, genes involved in phytoalexin detoxification (pisatin demethylase) and pectin lyases shed light onto the mechanism of pathogenicity and disease development in ABPFP – pea interaction. Moreover, the potential of secondary metabolite production coupled with high CAZymes content may help in understanding of survival strategies of ABPFP. Further studies will uncover the competitive or antagonistic effects of secondary metabolite production of pathogens adapted to the same ecological niche. The antagonistic effect of co-occurrence of *P. pinodes* and *P. pinodella* on inoculated pea was previously reported [184]. Similarly *Ph. koolunga* has an antagonistic effect on both *P. pinodes* and *P. pinodella* as determined by DNA quantification using real-time quantitative PCR (qPCR) [58].

In comparative analysis, over 100 potential gene families were found specific to ABPFP which may determine their host specificity. Significant differences were also noticed within ABPFP using Pfam domain enrichment analysis. Over-represented domains in *P. pinodes* were associated to osmotic adjustment, sexual reproductive cycle, and proteins with alpha/beta hydrolase that have diverse catalytic activities. These findings could explain adaptive advantage of *P. pinodes* to the hostile environment during infection that result in rapid disease development and host colonization compared to other ABPFP. The sequence data is a substantial contribution to the scarce molecular information currently available for pathogens in Didymellaceae. The data should serve in discovering and characterisation of key pathogenicity related genes in ABPFP. It will also help to identify host specificity factors among these complex pathogen–pea interaction to develop breeding tools that will enhance crop improvement.

Chapter Four Attribution Statement

Title: Evaluation of Nep1-like Genes in Didymellaceae Provide Insight to Subtype Predominant in Fungal Pathogens of Dicots

Chala Jefuka Turo, Bernadette Henares, Johannes Debler and Judith Lichtenzveig

This chapter includes a collaborative research that include bioinformatics analysis as well as functional assay. As a result, not all research activities were performed by the PhD candidate. The main contributors to this chapters are:

Chala Jefuka Turo (PhD candidate)

- Conceived the project
- Performed comparative genome analysis
- Performed phylogenetic tree analysis
- Performed protein evolutionary analysis
- Performed *in planta* transcriptome sequencing and expression analysis in connection to Chapter six.
- Involved in writing non-functional assay part of the chapter

Bernadette Henares

- Involved in cloning, protein expression and purification
- Involved in protein infiltration assay
- Involved in writing functional assay part

Johannes Debler

- Involved in cloning, protein expression and purification

Judith Lichtenzveig

- Involved in designing the project
- Lead the project

I, Chala Jefuka Turo, certify that this attribution statement is accurate record of my contribution to this chapter.

_____ _____

Chala Jefuka Turo (PhD candidate) Date

I, James Hane, certify that this attribution statement is accurate record of Chala Jefuka Turo's contribution to results presented in this chapter.

_____ _____

James Hane (Principal Supervisor) Date

4. Evaluation of Nep1-like Genes in Didymellaceae Provide Insight to Subtype Predominant in Fungal Pathogens of Dicots

4.1 Introduction

It is generally accepted that genes diverged after speciation event undertakes equivalent function [95], a concept that has been employed in comparative genomics to associate unknown gene-to-function relationship [95, 170]. Compelling evidences suggest partial functional replaceability of orthologues genes in distantly related organisms [149, 225]. The divergence of orthologous genes are often driven by positive selection [113, 302]. A large number of genes also acquire new function through duplication [133]. In pathogenic fungi, an evolutionary arms race of pathogen with host enhances the evolution of pathogenicity related genes via change in selection pressure which provides chances for rapid adaptive evolution.

The necrosis and ethylene-inducing peptide (Nep1-like protein, NLP) is an extracellular protein initially identified from *Fusarium oxysporum* culture filtrate [13], [14]. This protein is present in various microorganisms including fungi, bacteria and oomycetes [12, 106]. Also, *NLP* genes are not exclusive to pathogens that infect dicotyledons but are also present in the genomes of pathogens that infect monocotyledons. One of the hallmarks of an NLP is a conserved heptapeptide motif 'GHRHDWE' located in the central region of the protein is the hallmark of the NLP [234]. Furthermore, over, 90% of the identified NLPs contain an N-terminal signal sequence, which suggest that the protein is destined for secretion [233]. The NLPs can be classified based on the number of disulphide bonds where type I, II and III form one, two and three disulphide bridges, respectively [106]. Fungi all the three NLP types, bacteria contain type I and II while oomycetes have an expanded type I NLP.

Several studies have shown that NLPs causes necrosis only in dicot plants while monocots are unaffected [13, 241, 253, 276, 293]. They are involved in the virulence of *Colletotrichum coccodes*, *Abutilon theophrasti*, and *Rhynchosporium commune* [9, 83, 165]. Although the exact mechanism is yet to be determined, [234] showed that NLP contribute to disease development through the destruction and cytolysis of plasma membranes. The heptapeptide motif is understood to be essential for the phytotoxic activity [234, 276]. In addition, amino acid substitution studies in *Verticillium dahliae* revealed that several conserved amino acid residues are important for necrotic activities [354]. Whole genome analysis showed that most NLPs exist as multiple copies in different microbes with only a few exceptions such as *Mycosphaerella graminicola*, which contain only one NLP [222]. However, only two of the nine NLPs in *V. dahliae* cause necrosis [354]. Likewise, one out of five NLPs in *Moniliophthora perniciosa* [347] and eight out of nineteen representative NLPs in *Phytophthora sojae* [74] induce necrosis. This suggests that the existence of multiple copies of NLPs in a given species do not necessarily imply a greater degree of their phytotoxic activity. The proteins have either functionally diverged or inactivated to minimize functional redundancy as observed in many pseudogenized NLPs in *P. sojae* [346].

In addition to necrotic activities, NLP stimulates host immune responses in dicot plant (*Arabidopsis thaliana)* [12, 24, 82, 234, 253]. These include pathogen-associated molecular patterns, high expression of pathogenesis-related genes, induction of ethylene and salicylic acid biosynthesis, production of biosynthetic enzymes and reactive oxygen species. Intact protein sequence are not required for such responses. Experimental evidences in *Arabidopsis* indicated that 20 - 24 amino acids in the central region of NLP are sufficient to trigger plant immune responses [24, 233]. Both phytotoxic and non - phytotoxic NLP induce immune response in *Arabidopsis* [24].

Most fungal pathogens belonging to the family Didymellaceae in the class Dothideomycetes attack cool season pulse crops (Fabaceae). *Peyronellaea pinodes*, *Peyronellaea pinodella*, *Phoma koolunga* and *Ascochyta pisi* are the key pathogens that infect field pea while *Ascochyta rabiei*, *A. fabae* and *A. lentis* are principal

pathogens of chickpea, faba bean and lentil, respectively. These pathogens induce a typical necrotic symptom on their respective host collectively called ascochyta blight (AB). Regardless of the economic importance of pathogens belonging to Didymellaceae, limited molecular information is available in the public databases. This study aims to identify, compare and functionally test the necrosis inducing activity of the putative NLP(s) among key fungal pathogens in Didymellaceae.

4.2 Materials and Methods

4.2.1 Identification of NLP in Didymellaceae

Conserved functional domains present in the whole protein sequences of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* were identified by searching the Pfam-A database [250] using HMMER3 [77]. The NLP containing domain identified in *P. pinodes* was used to retrieve homologous protein sequences from *A. rabiei, A. lentis, A. fabae* and *A. vicea-vilosae* predicted proteomes using BLASTp [6]. The NLP protein sequences from other fungal pathogens including *Colletotrichum graminicola, Cochliobolus heterostrophus, Parastagonospora nodorum, Pyrenophora tritici-repentis, Fusarium graminearum, Fusarium. oxysporum f.sp lycopersici, F. verticillioides, Botrytis cinerea, Sclerotinia sclerotiorum,* and *M .oryzae* were identified from proteins clustered into families through whole proteome comparative analysis using Proteinortho version Proteinortho (version 5.1); parameters -singles –selfblast [185]. Sixty two NLP-containing sequences from various fungal, bacterial and oomycetes were retrieved from the genbank protein database using *P. pinodes* NLP as query (Supplementary Table 7). Presence of a NLP domain in all protein sequences from various sources of organisms was confirmed by sequence analysis at InterProScan [147].

The signal peptide domain of each proteins was predicted using SignalP 4.1 [243]. Glycosylation sites were predicted using the NetOGlyc 4.0 server hosted

at http://www.cbs.dtu.dk/services/NetOGlyc/. Linear structure of the protein illustration were drawn using DOG1 [257].

Full length protein sequences were aligned using Clastl W [180]. Phylogenetic tree was generated using maximum likelihood method according to Whelan and Goldman [331] model implemented in MEGA6 [307]. Sequence alignment coverage of less than 95% was excluded from phylogenetic analysis. The bootstrap branch support was obtained from 1,000 replications. Tree was edited using Evolver online tool (http://www.evolgenius.info/evolview) and exported in scalable vector graphics (svg) for further processing.

4.2.2 Test for amino acids under positive selection

A total of 20 and 21 protein coding nucleotides were aligned based on amino acid sequences using TranslatorX [1] (http://translatorx.co.uk/). Aligned coding sequences were tested for evidence of positive selection using CODONML as implemented in PAML version 4.7b [342].

4.2.3 Genome region context analysis

To assess the scale of sequence variation in genomic regions comprising NLP in a closely related species, whole genome sequences were aligned using PROmer via MUMer 3.0 [174] with max-match option. Coordinates of matching genomic regions were generated from the delta filtered Promer output using show-coords. Genomic regions (Scaffolds) matching to NLP regions were extracted from the whole genome and re-aligned using Promer with default parameters. Mummerplots were generated in scalable vector graphics (svg) file format for visualisation.

4.2.4 *In planta* expression assay of *P. pinodes*, *P. pinodella* and *Ph. koolunga*

Seeds of pea cultivar 'KASPA' commonly grown in Australia were planted on vermiculite and perlite mixed in 3:1 and grown on porous plastic pot for 27 days. The plants were inoculated with a conidial suspension of $5x10^5$ spore/ml determined by haemocytometer and covered with plastic sleeves to retain humidity over 95%. The plastic sleeves were removed 48 hours post inoculation. Plants were retained in a controlled growth chamber 'Conviron' (Argus Control System Ltd, Canada) adjusted to 20^0 C ± 2 and relative humidity of 60% followed by manual daily water spraying to maintain high humidity. Samples were collected at 0, 3, 6, 9, 21, 33, 48, 72 and 96 hours post inoculation (HPI). Samples were immediately frozen in liquid nitrogen and stored at -80^0C. Total RNA was extracted using Trizole method (Trizole® life technologies, USA), treated with DNase (Ambion® life technologies, USA). The integrity of RNA was analysed on 1% agarose gel electrophoresis using SYBR safe staining (SYBR® life technologies, USA). Based on cytological observation, approximately equal amount of RNA samples were pooled into early (0 – 6HPI), medium (9 – 33HPI) and late (48 – 96 HPI). The RNA was subjected to strand specific sequencing at Australian Genome Research Facilities (AGRF, Melbourne, Australia) following Illumina NGS standards with paired end reads of 100 bp. Low quality phred score (<28) and adaptors were trimmed from the reads and mapped to the whole genome sequences of respective species using TopHat (v2.0.9) [161]. Expression levels of *NLP* genes of interest were determined using Cufflinks version 2.1.1 [319].

4.2.5 Cloning, expression, and purification

4.2.5.1 Heterologous expression of *Ppdes_NLP1* in *Pichia pastoris*

For functional studies, NLPs from *P. pinodes* were used to test for its necrosis inducing activity. The full length cDNA of *Ppdes_nlp1* was amplified using Q5 polymerase (New England Biolabs) with primers that harbor EcoRI and XhoI restriction sites at the 5' and 3' end, respectively. The primers are 5´-cgcgGAATTCATGgTCTTCAATATTCAGTCTC-3´ forward and

5´-taGCGGCCGCtGAAGTAAGCATCCGCAAG-3´ reverse, respectively. The digested PCR product was cloned into appropriately digested pGAPZ (Invitrogen) to produce a C-terminal His6-tag. The resulting vector was transformed into Top10 cells and positive transformants were selected on LB plates containing Zeocin (30 ug/ml). The sequence of the clones was confirmed by Sanger sequencing (Macrogen sequencing facility, Korea). Expression of *Ppdes_NLP1* was carried out in *Pichia pastoris* X33 (Invitrogen) according to the manufacturer's instructions. Briefly, the cells were grown in yeast extract peptone dextrose broth supplemented with Zeocin (100 ug/ml) at 30^0C with shaking at 220 rpm for 48 hours. The culture supernatant was collected by removing the cells by centrifugation, then concentrated using amicon spin filter (MWCO: 3kDA; Millipore). Purification of His-tagged protein secreted into the culture medium was achieved using metal affinity chromatography (Nickel-Nitrilotriacetic acid agarose column, GE life sciences) equilibrated with Buffer A (50 mM Tris-HCl, pH8, 300mM KCl). Bound proteins were eluted with Buffer A containing 250 mM Imidazole. Fractions containing Ppdes_NLP1 were pooled and dialysed against 25mM Tris-HCl, pH 8.0. The protein purity was assessed by SDS-PAGE and protein concentration was determined (Coommassie Protein Assay Reagent kit; Pierce USA) using bovine serum albumin (BSA) as standard. Immunoblotting was performed using Anti-His6 antibody (Biosensis) and the bands were visualized using chemiluminescence (Chemidoc, Biorad). The amino acid sequence was confirmed by LC-MS (Proteomics International, Australia).

4.2.5.2 Heterologous expression of *Ppdes_NLP2* in *E. coli*

The coding region of the mature Ppdes_NLP2 (without the signal sequence) was amplified from *Ppdes_nlp2* cDNA using iProof polymerase with primers containing PciI and XhoI restrictions sites incorporated at the 5´end and 3´end, respectively. The locus specific primers are 5´-CATACATGTCACCTACTCCTGCCACGCT-3´ in forward and 5´-TCTTGAAGCCCTCCACTGGTGTTACCCGCA-3´ in the reverse direction. The amplified PCR product was digested with appropriate restriction enzymes and cloned in pET28a cut with NcoI (New England Biolabs) and XhoI (New England Biolabs). Transformation, selection and sequence confirmation of positive clones were carried as described in section 4.2.5.1. *Escherichia coli* Rosette cells

(Novagen) containing the positive clone was used to express Ppdes_NLP2 His6-tag fusion protein. Cells were grown in Terrific Broth (Yeast extract, glycerol) supplemented with 30 ug/ml kanamycin (Amresco, USA) and 34 ug/ml Chloramphenicol (Amresco, USA) at 37^0C and 220 rpm. Over expression was induced by 400 uM isopropyl-1-thio-d-galactopyranoside (IPTG) at mid-logarithmic stage. The culture was further incubated for 4 hours at 37^0C at 200 rpm after addition of IPTG. Cells were harvested and resuspended in Buffer A and lysed using sonicator (Bandelin, Germany). Soluble fractions were clarified by centrifugation. Purification and characterization of Ppdes_NLP2 were performed as described for Ppdes_NLP2.

4.2.5.3 Infiltration bioassay

Seeds of *Pisum sativum* cultivar 'KASPA', *Triticum aestivum* cultivar 'Mace', *Hordeum vulgare* cultivar 'FLAGSHIP', *Cicer arietinum* cultivar 'MOTI' were planted on vermiculite and perlite mixed in 4:1 and grown on porous plastic pots for 2 - 3 weeks. Protein infiltrations were performed using needleless 1cc syringe and pressure infiltrated on stipules of *P. sativum* (2-3 weeks old), leaves of *T. aestivum* and *B. vulgare* (2-3 weeks old) at 2 µM protein concentration. Necrosis was allowed to develop 2-5 days post infiltration and leaf tissues were excised and scanned at 1200 dpi (Epson XP200). All experimeintes were conducted in three replicaion.

4.3 Results and Discussion

4.3.1 Fungal pathogens in Didymellaceae contain two NLP

Bioinformatics and comparative genome analysis enabled us to identify a pair of NLP encoding protein sequences in eight key fungal pathogens from Didymellaceae, causal agent of ascochyta blight in pulse crops. All closely related fungal pathogen species from Didymellaceae including *P. pinodella, Ph. koolunga, A. pisi, A. lentis, A. fabae, A. rabiei* and *A. vicea-vilosae* contain a pair of NLP collectively named as NLP1 and NLP2. The number of observed NLPs in Didymellaceae are equivalent to those

reported in *B. cinerea* [51, 293], *S. sclerotiorum* [53] but less than that observed in *V. dahliae* [354] and *F. oxysporum* [14].

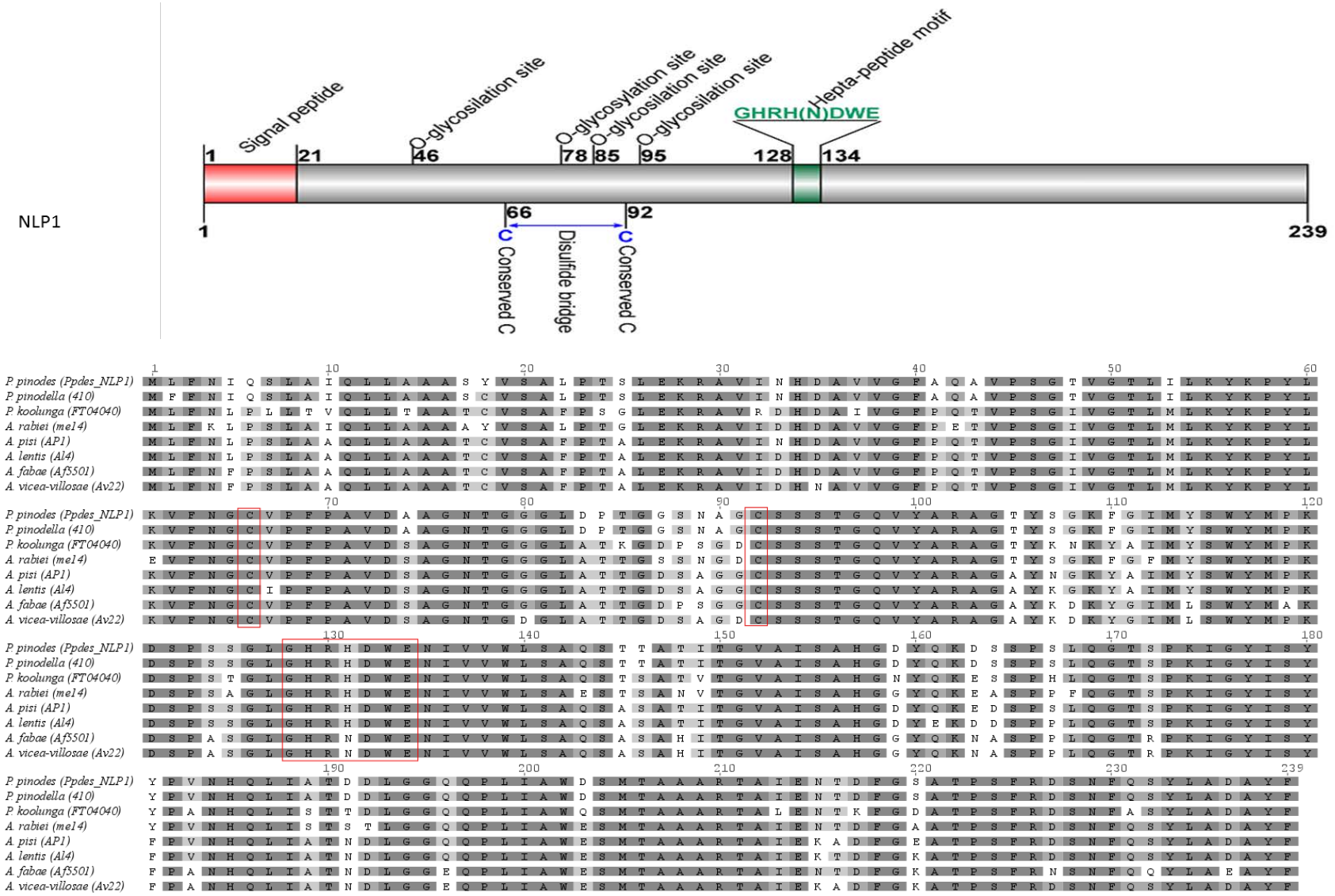
All of the eight pathogens from Didymellaceae contain type I NLP characterized by two conserved cysteine residues predicted to form a disulphide bridge. The NLP1 proteins are composed of 239 amino acid residues while NLP2 has 244 amino acids in *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. rabiei* and 242 amino acids in *A. pisi*, *A. lentis*, *A. fabae* and *A. viciae-villosae.* Both NLPs contain two conserved cysteine residues, thus belonging to type I category. All sequences contain a signal peptide at the N-terminus and a perfect heptapeptide motif at the central region except in *A. fabae* and *A. viciae-villosae* NLP1 as shown in Figure 14. Previous research indicated that the two conserved cysteines are essential for phytotoxic activity of NLPs in various microbes [82, 234, 354]. Predicted glycosylation sites are higher in NLP1 than NLP2 but most of the predicted O-glycosylation sites in NLP1 and NLP2 are similar and conserved across fungal species from the Didymellaceae (Figure 14). In addition, NLP2 predicted to contain two N-glycosylation sites not observed in NLP1. Glycosylation is important for the secreted protein to retain structure and activity as well as to prevention from degradation [265, 323].

The two NLPs have only 36 – 42% identical amino acid sites within each species. There is high sequence variation between the two paralogues (NLP1 and NLP2) in all species. Similar levels of paralog divergence was also reported in *Botrytis* species [293]. Some mutations are also observed in the heptapeptide motif. For instance, histidine is substituted with asparagine in the heptapeptide motif ‘GHRHDWE’ in *A. fabae* and *A. viciae-villosae*. Previous studies have shown the importance of this heptapeptide motif in necrotic activity [234, 354]. The impact of positively charged side chain histidine substitution with that of uncharged polar side chain asparagine in *A. fabae* and *A. viciae-villosae* need further studies to determine if it could abolish necrotic activities in *A. fabae* and *A. viciae-villosae*.

4.3.2 NLP2 is absent in closely related cereal pathogens

The phylogenetic relatedness among 96 NLP sequences from fungi, bacteria and oomycetes were analysed using maximum likelihood methods that enabled us to distinguish subtype 1 NLP (NLP1 and NLP2) that exist in a wide array of dicot fungal pathogens. According to the phylogenetic analysis, the fungal pathogens from the Didymellaceae can be placed into two distinct groups. Group A contain NLP1 which are closely associated to pathogens of monocots including C. *heterostrophus, P. tritici-repentis* and *P. nodorum* (Figure 15) The branching pattern of group A along the phylogenetic tree is congruent to species phylogeny. On the other hand, group B consists NLP2 which are closely related to most pathogens infecting dicotyledonous plants. The oomycetes (group C) are closely related to basidiomycete (*M. perniciosa).* Finally, group D consists of type II NLP from bacteria and fungi. Interestingly, three consecutive residues ('NVA' and their variants) were absent from NLP1 which could be used as a signature to distinguish the two NLPs (Supplementary Figure 1). No non-phytotoxic type 1 NLPs from oomycetes contain these three residues in their protein sequences.

Comparative genome analysis of *P. pinodes* against genomes of cereal pathogens (*P. tritici-repentis, P. nodorum, M. graminicola,* and *C. heterostrophus)* was performed to substantiate the phylogenetic distribution of NLP1 (closely related to monocot pathogens) and NLP2 (more associated to pathogens of dicots) based on hosts. The result reveals lack of sequence matches surrounding *P. pinodes* NLP2 (Ppdes_NLP2) loci in cereal pathogens. Further TBLASTn search of the aforementioned genomes using *P. pinodes* NLP2 (Ppdes_NLP2) as a query returned no significant hit. Protein clustering using Proteinortho showed absence of NLP2 orthologues in *P. tritici-repentis, P. nodorum, M. graminicola,* and *C. heterostrophus.* In addition, there is no type I NLP2 paralogues was reported from fungal pathogens of cereals tested. Both phylogenetic and sequence analysis suggests NLP2 orthologous are absent from the tested fungal pathogens of cereals suggesting NLP2 confers an adaptive advantage to pathogens of dicot plants.


NLP1
Signal peptide
O-glycosilation site
O-glycosylation site
O-glycosilation site
O-glycosilation site
GHRH(N)DWE
Hepta-peptide motif
1
21
46
78
85
95
128
134
239
66
92
Conserved C
Disulfide bridge
Conserved C
P. pinodes (Ppdes_NLP1)
P. pinodella (410)
P. koolunga (FT04040)
A. rabiei (me14)
A. pisi (AP1)
A. lentis (Al4)
A. fabae (Af5501)
A. vicea-villosae (Av22)

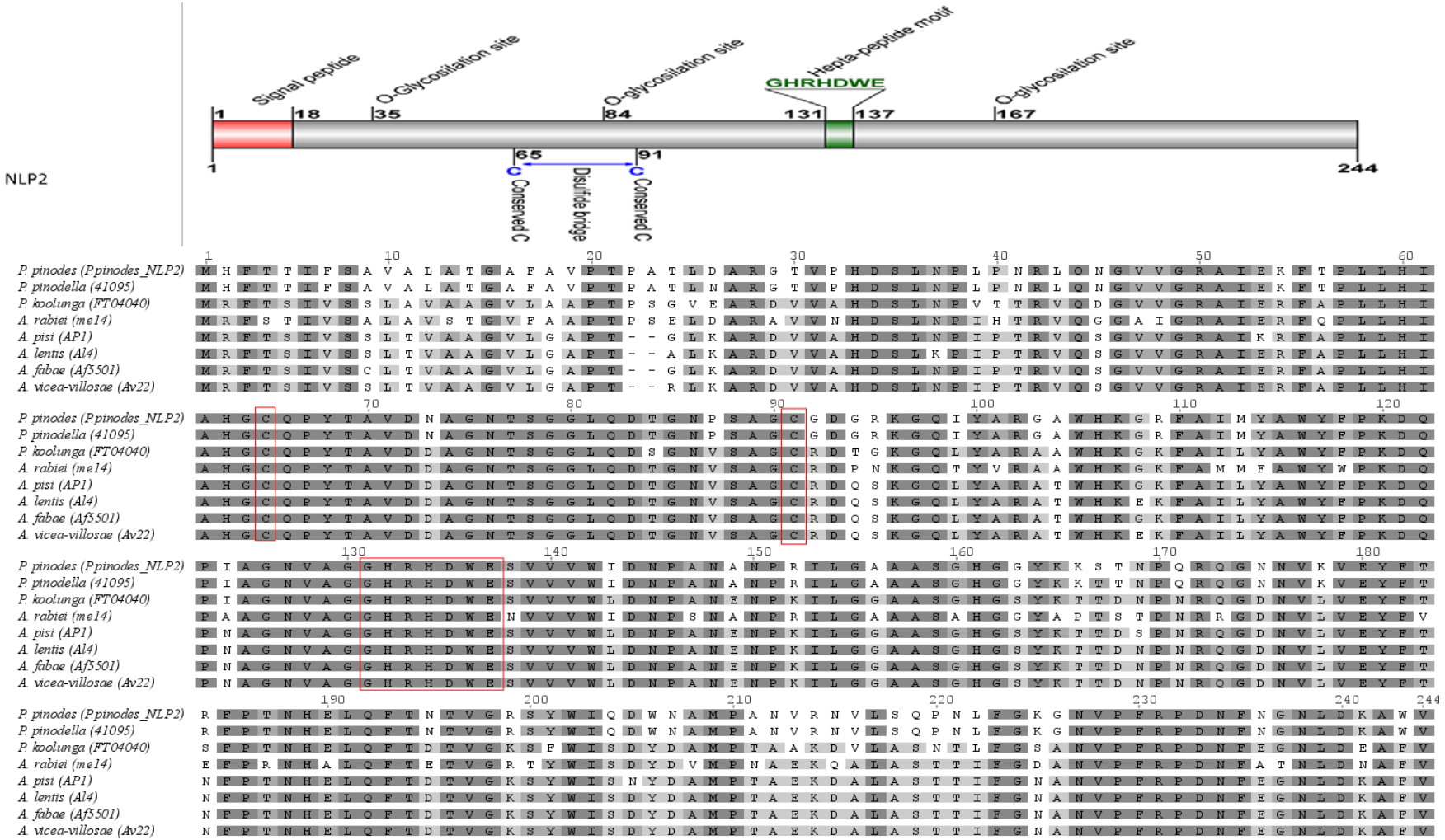


Figure 14 Amino acid sequence alignment of the necrosis- and ethylene-inducing peptide (Nep1-like protein, NLP) from Didymellaceae.

Sequences are highlighted according to amino acid percent similarity. Conserved cysteine and GHRHDWE motif is in black box.

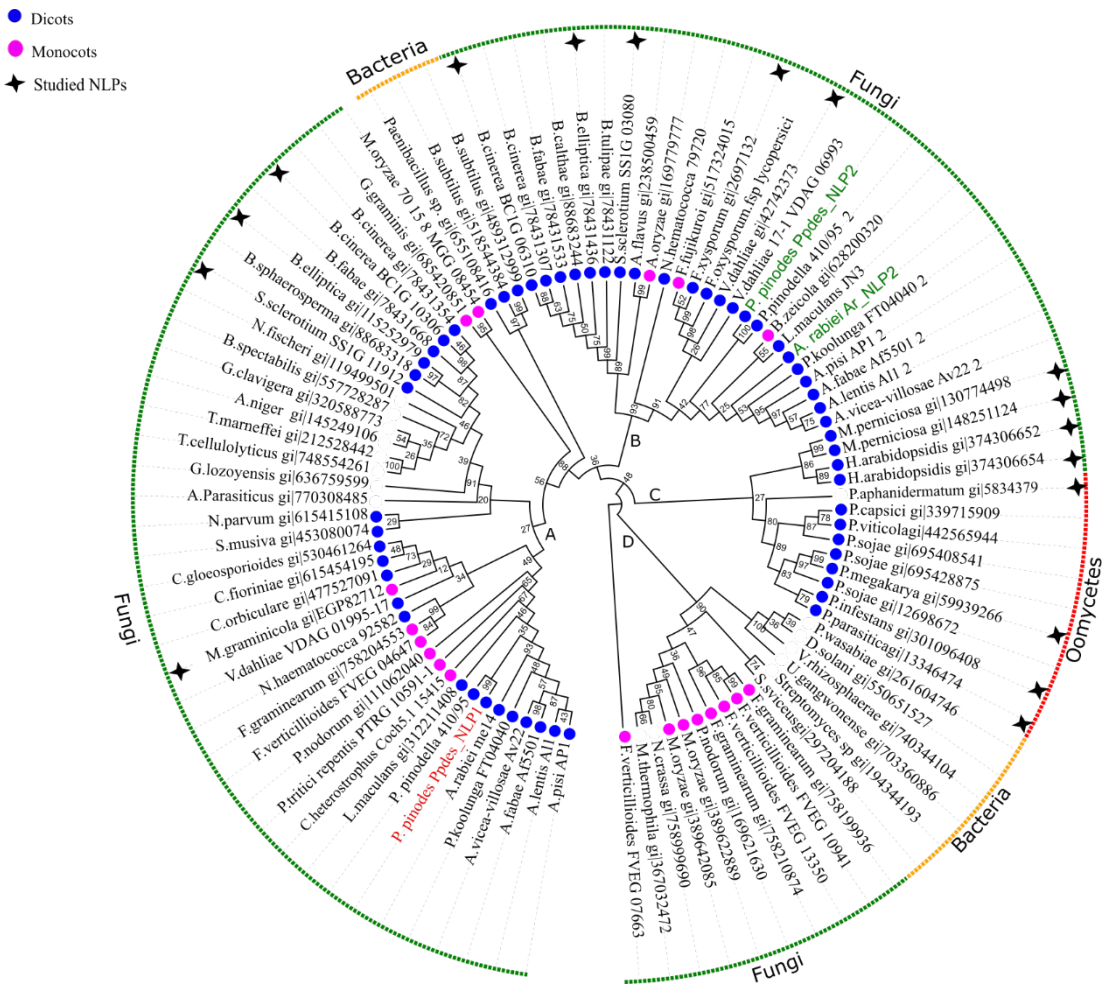


Figure 15 Phylogenetic distribution of Nep1-like proteins (NLPs).

Node numbers indicate bootstrap values from 1,000 replication. The analysis involved 96 amino acid sequences from oomycetes, bacteria and fungi. All positions with less than 95% site coverage were eliminated. Functionally characterized proteins were indicated in green (active) and red (inactive) font colours.

4.3.3 NLP resides in non-conserved loci of Didymellaceae genome

The genomic region containing NLPs in pathogens from the Didymellaceae were compared through six frame translation using MUMmer (PROmer). The loci surrounding NLP1 and NLP2 surrounding loci in *P. pinodes* show differences in sequence conservation compared to other pathogens in Didymellaceae. However, macrosynteny is conserved in scaffolds that carry NLPs between closely related species such as *P. pinodes* and *P. pinodella* as shown by high sequence similarity matches (red >90%) drawn across diagonal line in the dot-plot (Figure 16). When *P. pinodes* sequences were compared to *A. rabiei, A. lentis* and *A. fabae*, a break of sequence co-linearity is observed at NLP loci. Absence of matching sequences are reflected in dot plots (particularly between 10 and 15 kb) comparison between *P. pinodes* Scaffold 73 against *A. rabiei, A. lentis* and *A. fabae*(Figure 16). Similar result is observed when genomic loci (Scaffold) containing NLP2 in *A. rabiei* (Scaffold_85) is compared against *P. pinodes* and *P. pinodella*.

Given the limitations imposed by the fragmented nature of current genome assemblies, NLPs in Didymellaceae seems to reside in clade specific genomic regions within Didymellaceae. Existence of reminiscent retrotransposon (Gypsy 47 bp) and DNA transposon (Molley 57 bp) observed near NLP2 in *P. pinodes* might have contributed to sequence variation through RIP leakage. In *F. oxysporum* three of the NLP loci reside in lineage specific genomic regions rich in transposable elements [201]. Similarly, one NLP is also located on lineage specific chromosome in *F. verticillioides*.

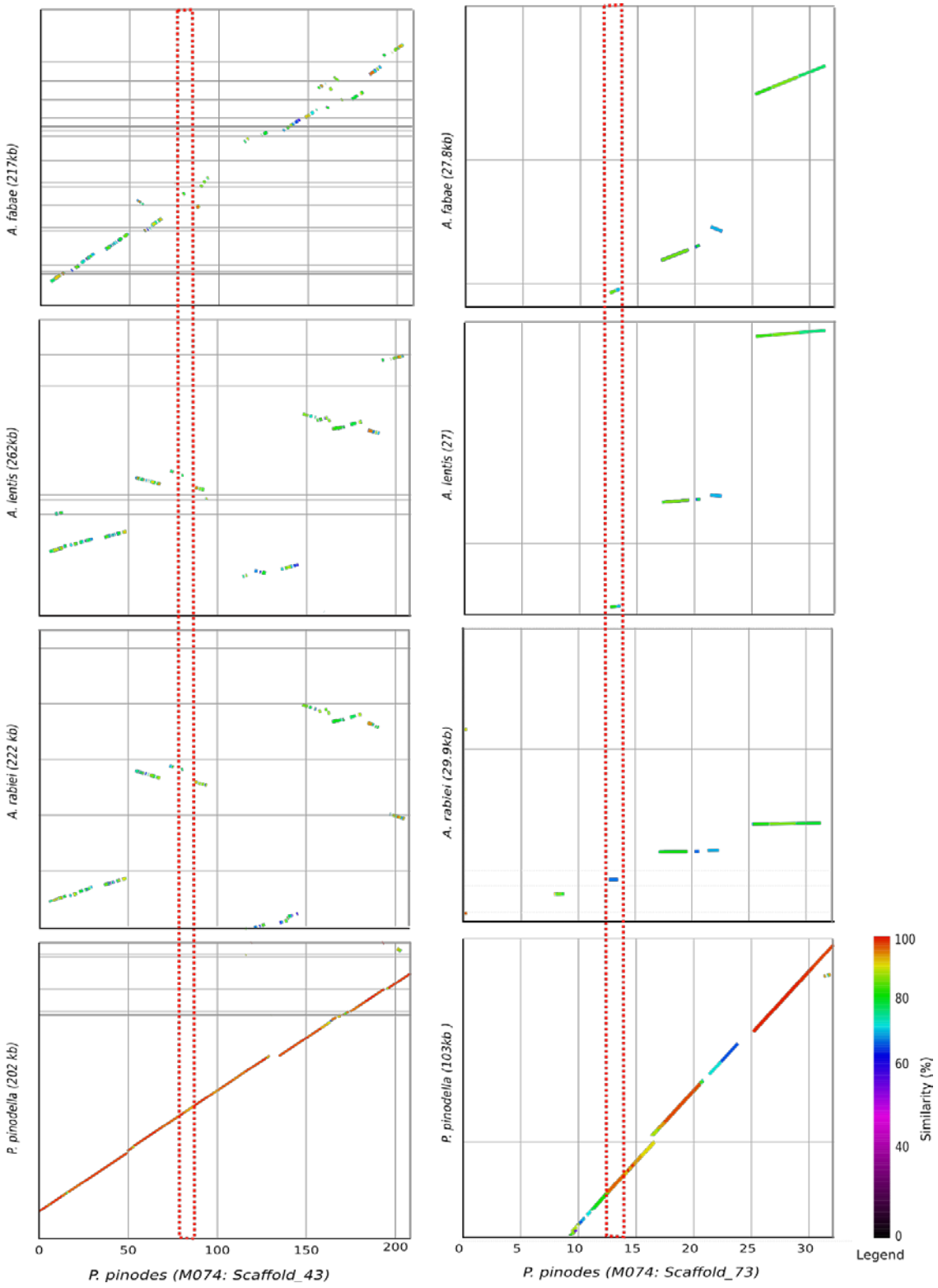


Figure 16 Dot-plot comparison of genome loci containing Nep1-like proteins (NLPs) in *P. pinodes* and closely related species in Didymellaceae.

Matching sequences were colour coded according to their similarity. Red broken box indicate position of NLPs on *P. pinodes* scaffold.

4.3.4 NLP is under purifying selection

Tests for positive selection under random-site model was performed using CODEML for NLP1 and NLP2. A total of 21 and 20 full length amino acid coding sequences orthologous to NLP1 and NLP2, respectively, were used in the analysis. A ω value of 0.183 in NLP1 and 0.107 in NLP2 is observed under M8 indicating that these genes are under purifying selection. The M2a model suggests 0.36% of the sites in NLP1 are under positive selection at ω =7.54 (Table 12). This corresponds to position 226A (based on alignment) at ≥95% posterior probability following Bayes Empirical Bayes (BEB) probabilities. This position is not polymorphic between the two NLPs within a species. There is no positively selected sites detected in NLP2 in this analysis indicating that it is under tight regulation. Previous studies indicated that NEP1 and NEP2 in *Botrytis* species showed a general trend of purifying selection with the exception of few amino acid residues under positive selection [293]. Amino acids under positive selection were also observed in *NIP1* genes of *R. secalis* [277]. The results imply that host response that might have been developed in response to NLP in the Fabaceae do not interfere with NLP phytotoxicity.

4.3.5 NLP is expressed during infection and disease development

The expression pattern of both NLPs was determined for *P. pinodes, P. pinodella* and *Ph. koolunga* using *in planta* RNA sequencing. In *P. pinodes,* NLP2 is constitutively expressed during early to late stage of infection while NLP1 is expressed from mid to late stage of infection (Figure 18) indicating differential activation of these genes in response to host. In *P. pinodella,* NLP1 shows consistent expression from early to late stage of infection while NLP2 is expressed from mid to late stage of infection. Likewise, NLP2 showed high level of expression from early to late stage of infection in *Ph. Koolunga* (Figure 18). No expression of NLP1 is detected at early stage of infection in both *P. pinodes* and *Ph. koolunga.*

The *in planta* expression of NLPs in *P. pinodes*, *P. pinodella* and *Ph. koolunga* suggests that these genes may contribute to disease development by through eliciting host cell death in a manner typical of a necrotrophic fungal lifestyle. In *P. pinodes* and *Ph. koolunga* the high expression at early to mid-stage of infection is related to early onset and expansion of necrosis. Previous work has shown high expression of NLP during transition from biotrophy to necrotrophy in *Phytophthora sojae, P. infestans* and *Colletotrichum higginsianum* [155, 167, 252]. Similarly, *B. cinerea Bcnep1* and *Bcnep2* expression were detected at early hours after treatment and during onset of necrotic lesion/biomass build-up, respect [51, 53]. As opposed to this, *HaNLP* was expressed during early stage of infection in *Hyaloperonospora arabidopsidis* [34] and *M. graminicola* [222].

4.3.6 Functional characterization of NLP

Several studies have shown that the NLP gene family members from different microorganisms differ in their ability to induce necrosis in dicotyledonous plants [13, 241, 253, 276, 293]. In order to test whether Didymellaceae NLPs have cell-death activity, the proteins were heterologously expressed in *P. pastoris* for Pdes_NLP1 and in *E.coli* for Ppdes_NLP2. Necrotic lesions on pea cultivar ‘Kaspa’ infiltrated with purified Pdes_NLP2 started to appear as early as 2 days post infiltration (dpi) and clear necrosis was observed 3 dpi (Figure 17). Infiltration of pea stipules with controls purified from cells containing the corresponding empty vector did not show any signs of necrotic lesions. On the other hand, infiltration of pea cultivar ‘Kaspa’ with the same concentration of purified Ppdes_NLP1 failed to induce necrosis. Ppdes_NLP2 did not elicit any response from barley cultivar ‘Flagship’ even after 10 dpi (Figure 17). Western blot analysis confirmed the presence of the proteins that were infiltrated into the stipules of pea ‘Kaspa’. Taken together, these data strongly suggest that Ppdes_NLP2 has a cell-death inducing activity only on dicots while Ppdes_NLP1 lacks the ability to induce cell death under current experimental conditions. This finding is consistent with cell-death inducing activities of several NLPs in dicotyledonous plants but not against monocotyledonous plants [13, 241, 253, 276, 293].

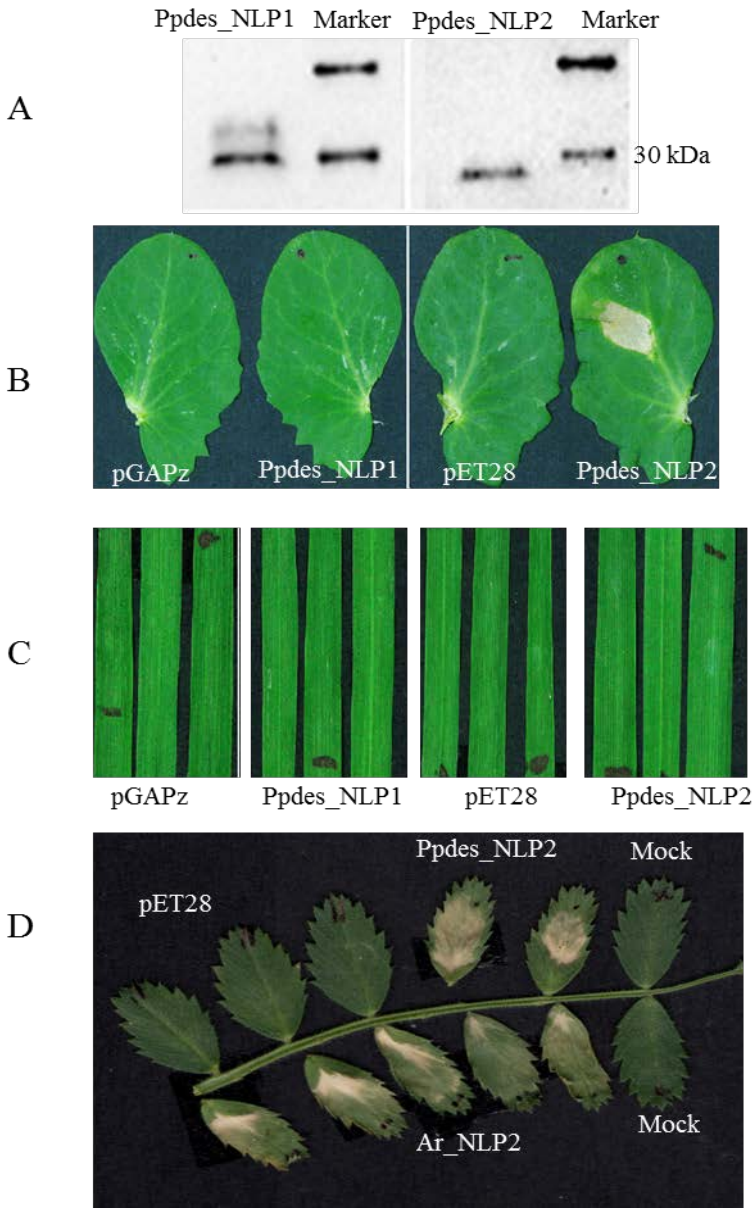


Figure 17 Ppdes_NLP2 and Ar_NLP2 induces cell death in pea and chickpea plants but not in barley.

A) Western blot analysis of purified Ppdes_NLP1 and Ppdes_NLP2 after staining
B) NLP from *P. pinodes* infiltrated on pea cultivar 'Kaspa'
C) NLP1 and NLP2 from *P. pinodes* infiltrated on barley cultivar 'Flagship' NLP2 from *P. pinodes* and *A. rabiei* infiltrated on pea cultivar chickpea cultivar 'Moti'. Photographs were taken 3 days post infiltration (dpi) for pea and 7 days (dpi) for barley. Only Ppdes_NLP2 triggers necrosis in pea, while both Ppdes_NLP2 and Ppdes_NLP1 do not elicit response from barley under the current conditions.

To determine whether orthologues of Ppdes_NLP2 from *A. rabiei*, a member of Didymellaceae pathogen of chickpea, has the same cell-death –inducing activity, Ar_NLP2 was amplified using 5´-CATACATGTCTCCCACTGGGCTCGAGAA-3´ forward and 5´-CATGCGGCCGCGAAGTAAGCATCAGCAAGGT-3´ reverse sequences, expressed in *E. coli*. Infiltration of purified Ar_NLP2 into the leaflets of chickpea cultivar 'Moti' showed necrosis 3 dpi, while the control vector failed to induce necrosis (Figure 17). In addition, infiltration of Ar_NLP2 showed the same necrosis-inducing activity on pea cultivar 'Kaspa'.

Protein assay using SDS-PAGE revealed presence of exta band from Ppdes_NLP1 expressed in *P. pastori* (Figure 17), which implied post translational modification of the expressed protein. The purified protein infiltrated on dicot plants did not induced necrosis. Speculatively, post-translational modification might have inactivated the necrotic activity of Ppdes_NLP1. This contradict with previous finding where post-translational modification (glycosylation) is important for Slp1 effector protein stability and function. Similarly, glycosylation appear to enhance infection and subsequent disease development in *C. albicans* ([327]).

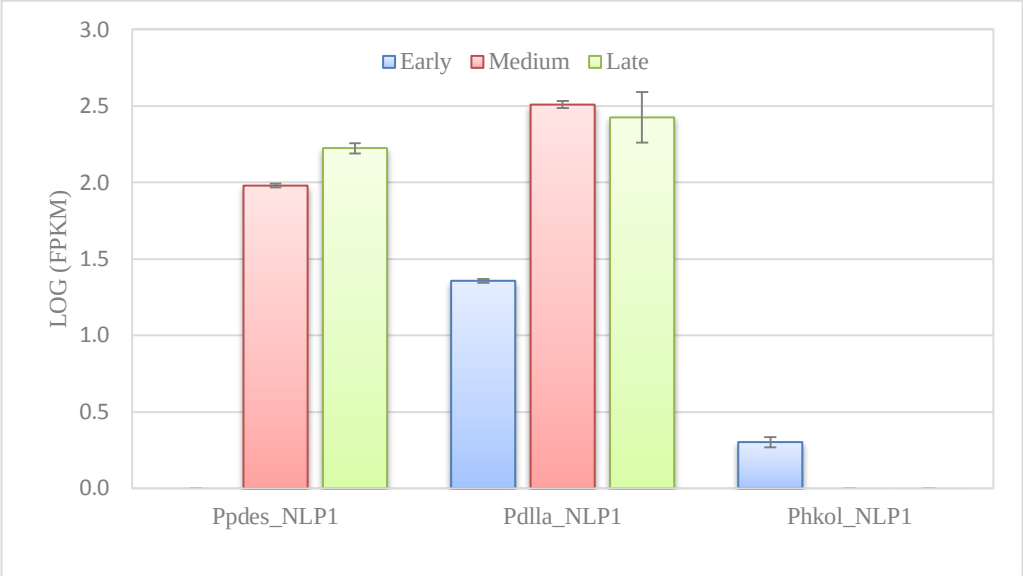
We were unable to express Ppdes_NLP2 in *P. pastori* to observe similar result. However, Ppdes_NLP2 expressed in *E. coli* showed strong necrotic activity. The *E. coli* expression environment may minimize glycosylation of Ppdes_NLP2 and improved its subsequent cytotoxic activity against dicot plants since post translational modifications such as glycosylation is less common in bacteria [70].

The difference in necrotic activities between NLP1 and NLP2 in Didymellaceae also suggests functional diversification between the two paralogues. Duplicate genes acquire new function via neofunctionalisation and sub-functionalisation as observed in yeast [33], *M. grisea* cutinase gene family [288], plant UbiA gene family [329]. In *V. dahlia*, VdNLP is involved in asexual reproduction, vegetative growth, as well as pathogenesis [272, 354]. It was also suggested that non-phytotoxic NLP may

contribute to disease development without necrotic effect [272]. The fact that both NLP expressed *in vitro* and *in planta* plausibly support the idea of additional role of NLP1 and NLP2 under various circumstances yet to be verified.

4.4 Conclusion

The eight fungal pathogens from Didymellaceae family contained a pair of similar NLP proteins sequences. The genomic loci of these two genes lack synteny conservation except in sister species. Both NLPs are actively expressed during pea infection by *P. pinodes*, *P. pinodella* and *P. koolunga*. However, only NLP2 possessed phytotoxicity against dicot plants. The potent necrosis inducing protein, NLP2, wasn't detected at least in four pathogens infecting cereals suggesting niche adaptive evolution of Didymellaceae pathogens to dicotyledonous hosts. The evolution of NLP under purifying selection indicates lack of deleterious host defence response established to halt phytotoxic activities of NLP in Fabaceae. Identifying and studying the NLP interacting molecule in the host may help to develop novel pre-breeding techniques. Determining the contribution of NLP2 in disease development in ABPFP warrants further knockout/down studies. It is also important to clarify the impact of post-translational modification of Ppdes_NLP1 and Ppdes_NLP2 will influence disease development during host and pathogen interaction.



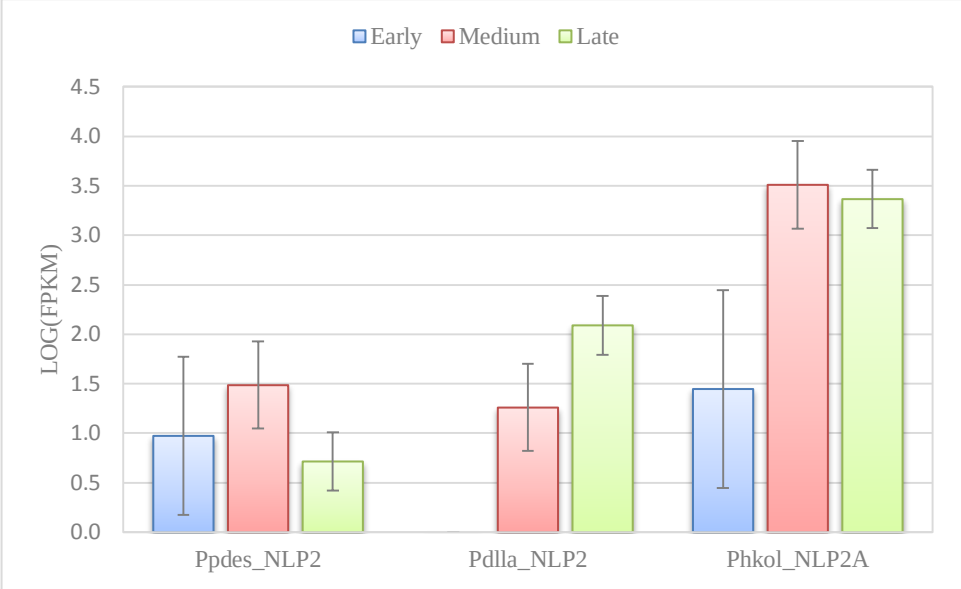


Figure 18 *In planta* expression pattern of necrosis - and ethylene inducing - peptide like proteins (NLPs) from *P. pinodes, P. pinodella* and *Ph. koolunga*.
P. pinodes NLP = Ppdes_NLP1 & Ppds_NLP2; *P. pinodella* NLP = Pdlla_NLP1 & NLP2; *Ph. koolunga* NLP = Phkol_NLP1 & Phkol_NLP2A. Error bars indicate standard error of mean.

Table 12 Log likelihood values and parameter estimates under variable ω ratios for NLP1 and NLP2

	Model code	lnL	Parameter estimates	Positive selected site
NLP1	M0	-8364.14	0.13487	Not allowed
	M1a	-8236.35	ω_0=0.11019 (ω_1= 1), p_0 =0.76087(p_1=0.23913)	Not allowed
	M2a	-8235.52	ω_0=0.11064(ω_1=1), ω =7.54665, p_0=0.7592, p_1=0.2372, p_2=0.00356	**226 A**
	M7 (b)	-8123.28	p = 0.60750 q = 2.96372	Not allowed
	M8 (beta & ω)	-8120.46	p_0 = 0.99561(p_1 = 0.00439) p = 0.62470, q = 3.17086, ω = 5.02143	**226 A**
NLP2	M0	-7078.79	0.08495	Not allowed
	M1a	-6948.92	ω =0.07358 (ω_0=1) p_0= 0.82129 (p_1= 0.17871)	Not allowed
	M2a	-6948.92	ω_0=0.07358 (ω_2=1), ω_2=32.80982, p_0=0.82129, p_1=0.17871, p_2=0.00	None
	M7 (b)	-6857.35	p = 0.57373 q = 4.62350	Not allowed
	M8 (beta & ω)	-6854.5	p_0 = 0.97035(p_1 = 0.02965), p = 0.64319, q = 6.35815, ω =1.00	37A, 166T,198N

Sign for positive selection was analysed using CODEML implemented in PAML. Model M2a is compared against model M21a and Model M8 is compared against M7 for detecting positive selection. Twenty one and twenty amino acid sequences were included in the analysis. Positively selected sites were identified at posterior probability cut-off value of $p \geq 95\%$ is shown in bold.

Note: NLP1: M2a suggest 0.00356*100 is under positive selection at ω =7.54.

Note: NLP2: M2a suggest 0.00*100 sites are under positive selection at ω =32.80982. No site under positive selection

5. Known and Conserved Effector Candidate Genes Associated with Ascochyta Blight Pathogens of Field Pea

5.1 Introduction

Plant pathogenic fungi can be perceived by the host plant immune system and induce host defence responses, which are triggered by molecular epitopes referred to as pathogen-associated molecular patterns (PAMPs) [197]. Chitin oligosaccharides (N-acetyl chito-oligosaccharides) are an example of common PAMPs that induce defence responses in various plant species [25, 283]. Recognition of chitin through membrane-localized pattern recognition receptors leads to the induction of defence responses, called PAMP-triggered immunity (PTI) [146].

Fungal pathogens have developed various mechanisms to avoid, overcome, suppress or detoxify pre-existing or induced host defence responses using a wide variety of "effector" molecules and other pathogenicity-related proteins. Many effectors remain unknown, however, several have been identified and characterised to varying degrees. LysM effectors are widely conserved effector molecules that have been characterized in several economically important fungal pathogens. For instance, *Ecp6* of *Cladosporium fulvum*, *Slp1* of *Magnaporthe oryzae*, and *Mg1LysM*, *Mg3LysM* of *Mycosphaerella graminicola* are all LysM effectors. LysM effectors bind to extracellular fungal chitin and shield hyphae from plant derived hydrolytic enzymes [63, 210, 216] which bind to chitin and are thus able to suppress chitin-triggered PTI, which leads to a compatible interaction in their respective host [63, 210, 216]. Similar preventive mechanisms are employed by *Mycosphaerella fijiensis* (*MfAvr4*) and *C. fulvum* (*Avr4*), via Avr proteins that are able to subvert host defence mechanisms [296].

Many fungal pathogens are also able to detoxify induced antimicrobial compounds into nontoxic compounds. For example, fungal pathogens of crucifers such as *Alternaria brassicola* and *Leptosphaeria maculans* are able to detoxify the

antimicrobial compound camalexin into its non-toxic form [236, 237]. The detoxification of brassinin by *A. brassicola* is also able to suppress host defence responses [239].

Available literature has indicated that *Peyronellaea pinodes* produces small glycoprotein molecules (suppressors) that condition pea to become susceptible to fungal infection [284, 286]. The three ascochyta blight pathogens of field pea viz *P. pinodes, Peyronellaea* and *Ascochyta pisi* are able to detoxify pisatin, a phytoalexin produced by pea through a demethylation process similar to that of the wilt pathogen *Nectria haematococca* [71, 105]. Research finding show that pisatin is demethylated directly by a cytochrome P450 monooxygenase [105, 213].

The corresponding genes governing demethylation have not yet been identified in Ascochtya Blight Pathogens of Field Pea (ABPFP), or even predicted prior to this analysis. The purpose of investigations in this chapter is to identify and compare putative known and well characterized effectors and other pathogenicity genes in the new genome resources of *P. pinodes*, *P. pinodella*, *A. pisi* and *Phoma koolunga*, such as LysM effectors and pisatin demethylase genes, for which prediction is likely to be relatively reliable.

5.2 Materials and methods

5.2.1 Protein sequence acquisition, annotation and phylogenetic analysis

Whole proteome sequences of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* were obtained as described in Chapter 2. Conserved functional domains present in the whole protein sequences of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* were identified through scanning versus the Pfam-A database [250] using HMMER3 (3.0, default parameters) [77]. Sequences with a significant hit ($P < 0.01$) to the Pfam record for the LysM conserved domain (PF01476) were retrieved from each genome annotation.

Likewise, protein sequences that were functionally assigned the annotation for cytochrome P450 pisatin demethylase (PDA, GO: 0005506) via BLAST2GO (go_201303-assocdb-data) [50] were retrieved from the whole genome sequences of *P. pinodes, P. pinodella, A. pisi* and *Ph. koolunga.* Protein orthology analysis using Proteinortho (version Proteinortho 5.1; parameters -singles -selfblast) [185] was used to find orthologous protein sequences from various fungal species including *P. nodorum* [128], *C. heterostrophus* [229], *L. maculans* [266], *Botrytis cinerea* and *Sclerotinia sclerotiorum* [7], *N. crassa* [96], *Pyrenophora tritici-repentis* [204], *N. haematococca* [46], *C. graminicola* [227], *Rhizoctonia sola* [332], *Fusarium graminearum* [52], *Verticillium dahliae* [168], *Magnaporthe grisea* [67] and *Fusarium verticillioides* [200]. The sources of proteome data used in this study are indicated in the (Table 13).

In addition, functionally characterized LysM - and PDA-encoding protein sequences were retrieved from NCBI database (Table 13). The secretion domain of each protein was predicted using SignalP 4.1b (-f short, -k, -t euk) [243]. Linear structure of the protein illustration were generated using DOG 1.0 [257]. BLASTp (blastall version 2.2.26, -p blastp, -m 8, -e 10^{-3}) [6] analysis of *P. pinodes* and *P. pinodella* putative PDA against functionally characterized *PDA* genes of *N. haematococca* and *F. oxysporum fsp pisi* was also performed to determine the level of sequence identity.

Full length protein sequences were aligned using CLUSTAL W with BLOSUM protein weight matrix (gap open penalty 10, gap extension penalty 0.2 in multiple alignment [180]. Although orthologous protein sequences were detected through orthology analysis, only functionally characterized LysM and PDA containing protein sequences from non-Didymellaceae and putative LysM and PDA from ABPFP were included in the phylogenetic analysis. A phylogenetic tree was generated using the neighbour joining method [269, 331] implemented in MEGA6 [307]. Bootstrap branch support was obtained from 1,000 replications. Branches with less than 50% bootstrap support were collapsed and the consensus trees were edited using the online tool EvolView [348] (http://www.evolgenius.info/evolview) .

Table 13 Sources of characterized genes used in this study

Organism	Domain	Gene name	NCBI ID	Reference
Aspergillus niger	LysM	-	EHA27166	Mentlak *et al.*, 2012
Aspergillus A niger	LysM	-	XP_001402350	Mentlak *et al.*, 2012
Cladosporium fulvum	LysM	*Ecp6*	4B9H_A	de Jonge *et al.*, 2010
Cladosporium fulvum	CBM50	*Avr4*	CAA69643	Joosten *et al.*, 1997
Colletotrichum lindemuthianum	LysM	*CIH1*	CAA04765	Perfect *et al.*, 1998
Magnaporthe oryzae 70-15	LysM	*Slp2*	XP_003716457	Mentlak *et al.*, 2012; Gohari *et al.*, 2015
Mycosphaerella fijiensis	LysM	-	XP_007929246	Ohm *et al.*, 2012
Mycosphaerella graminicola IPO323	LysM	*Mg2LysM*	XP_003850050	Marshall *et al.*, 2011; Gohari *et al.*, 2015
Mycosphaerella graminicola IPO323	LysM	*Mg3LysM*	XP_003848663	Marshall *et al.*, 2011
Magnaporthe oryzae 70-15	LysM	*Slp1*	XP_003717420	Mentlak *et al.*, 2012
Sclerotinia sclerotiorum	LysM	-	XP_001595446	Mentlak *et al.*, 2012*
Fusarium oxysporum f.sp. pisi HDV247	PDA	*FoPDA1*	AAR32716	Coleman *et al.*, 2011
Nectria haematococca mpVI	PDA	*PDA6-1*	P38364	Reimman and VanEtten, 1994
Nectria haematococca mpVI	PDA	*PDA1*	XP_003044225	Coleman *et al.*, 2009
Nectria haematococca mpVI	PDA	*PEP1*	AAK11166.1|AF294788_1	Han *et al.*, 2001
Nectria haematococca mpVI	PDA	*PEP2*	AAK11167/AF294788_2	Han *et al.*, 2001
Nectria haematococca mpVI	PDA	*PEP5*	AAK16922.1|AF315315	Han *et al.*, 2001

LysM and PDA indicate lysine motif and pisatin demethylase, respectively.

5.3 Results and discussion

5.3.1 Conserved LysM might be involved in infection and disease development

The availability of genome sequences for *P. pinodes, P. pinodella, A. pisi* and *Ph. koolunga* allowed the survey of their whole proteomes to identify conserved LysM domain containing protein sequences. As typically observed across a number of Ascomycetes species, all the four fungal pathogens contained LysM domain-containing protein sequences. Forty-five LysM domain containing protein sequences were identified from the genome sequences of *P. pinodes*, *P. pinodella*, *A. pisi* and *Ph. koolunga* (Table 14). Over 88% of these predicted LysM proteins had molecular weights of 34 kDa or less. Of these, 60% were cysteine-rich (≥ 5 cysteines). Moreover, 51% of these LysM domain-containing sequences from ABPFP were predicted for extracellular secretion (Table 14). These cysteine-rich secretory LysM containing proteins were predicted as effector candidates identified in ABPFP.

Phylogenetic analysis grouped the LysM-containing protein sequences into four main groups, as indicated by coloured branches in Figure 19. These groups indicated with green and blue branches contained 1- 4 LysM domain with a signal sequence at the N-terminus. Proteins with such domain configuration has been previously described as "LysM effectors" [62]), which were secreted and possess multiple LysM motifs without known catalytic domain [3]. Groups indicated with red branches were predicted to contain single or multiple LysM domain but lack a predicted signal peptide. According to the phylogenetic analysis, all functionally characterized LysM effectors such as *Ecp6* from *C. fulvum*, *Mg2LysM* and *Mg3LysM* from *M. graminicola,* and *Slp1* and *Slp2* from *M. oryzae* were grouped together. The predicted LysM effectors from *Pk. koolunga* (*Phkol_FT04040_9842, Phkol_FT04040_8631, Phkol_FT0713_10520,* and *Phkol_FT0713_2974*) and *A. pisi* (*Apisi_11254, Apisi_AP1_11254, Apisi_AP1_9426, Apisi_Georgia-7_4297,* and *Apisi_Georgia-*

12_5867) also grouped with the known LysM effectors with 98% bootstrap support (Figure 19).

Various researches indicated that LysM-containing proteins function as virulence factor in plant pathogenic fungi. Over-expression of extracellularly secreted *C. fulvum Ecp6* increased the virulence of *Fusarium oxysporum* on tomato (Bolton et al., 2008). *Ecp6* facilitates virulence through prevention of chitin-triggered host immunity by sequestering fungal chitin fragments [63]. Similarly, *Mg1LysM* and *Mg3LysM* from *M. graminicola* are involved in protection of fungal hyphae against plant derived hydrolytic enzymes [210]. The genes are specifically expressed during symptomless infection stage. The *Slp1* from *M. oryzae,* is also required for full virulence on rice and functionally similar to *C. fulvum Ecp6*. The close association of LysM domain containing proteins from *Ph. koolunga* and *A. pisi* are grouped together with known LysM effector, with similar domain architecture and presence of signal sequences at N terminus indicating their potential role during infection and subsequent disease development. A similar group of predicted LysM effectors from ABPFP (indicated in green branches in Figure 19 may also play similar role.

Similar to LysM, cyanovirin-N homology domain (CVNH) is an 11-kDa sugar binding protein domain conserved in eukaryotes [242]. All the four ABPFPs contain a CVNH protein that includes a LysM domain [Pfam: PF01476] sandwiched between CVNH [Pfam: PF01476] domains as indicated in Figure 19. The N terminus of these protein sequences contain Rick-17kDa domain but lack a signal peptide, consistent with previous finding. Percudani *et al.* [242] indicated that CVNH is found as a non-secretory domain in filamentous ascomycetes. It is CVNH is widely observed in bacteria, fungi and plants [242]. However, BLASTp analysis versus NCBI NR Proteins using *P. pinodes* CVNH as a query returned no significant hits in the Taphrinomycotina (Ascomycotina taxid:451866), Basidiomycota (Agaricomycotina, taxid:5032; Pucciniomycotina, taxid 29000; Ustilaginomycotina, taxid 452284 or Oomycetes (taxid 4762). Further sequence search was conducted in the reference proteome database using HMMER (http://www.ebi.ac.uk/Tools/hmmer/) at p-value of

0.001. This result corroborated the previous result from BLAST analysis –an absence of significant hits in fungi belong to Saccharomycotina, Taphrinomycotina and Oomycota. Furthermore, very few sequences with significant hits are observed across the Basidiomycota (*Pleurotus ostreatus, Galerina marginata, Piriformospora indica,* and *Botryobasidium botryosum*). Based on combined analysis of BLASTp and Phmmer, the data suggested that CVNH is discontinuously distributed across fungi. Percudani *et al.* [242] suggested that CVNH might have been more recently acquired by some organisms after the separation of the main evolutionary lineages. Further evolutionary studies may reveal the evolutionary history of CVNH across a wider taxonomic range.

Only a little information is available with regards to contribution of the CVNH domain to fungal virulence. In *M. oryzae,* the CVNH domain protein MGG_03307 is localized to the appresorium and suggested to play crucial role in early stage of infection [169]. It would be interesting to functionally characterize CVNH homologs in ABPFP and related necrotrophs to determine its role in inducing rapid cell death in host plants.

Table 14 Amino acid properties of proteins with LysM motif

Organism	Protein Id/ known effector	Isolate	Length	Cysteine count (%)	Molecular weight	Secreted
P. pinodes	Ppdes_24.2	M074	356	15(4.2)	29.32	Yes
	Ppdes_26.69	M074	342	3(0.9)	27.70	No
	Ppdes_5.67	M074	371	16(4.3)	32.61	Yes
	Ppdes_6.2	M074	275	14(5.1)	22.70	No
	Ppdes_10730	MP1	371	16(4.3)	32.61	Yes
	Ppdes_11617	MP1	275	14(5.1)	22.70	No
	Ppdes_11951	MP1	133	5(3.8)	10.30	No
	Ppdes_9992	MP1	301	13(4.3)	25.34	Yes
P. pinodella	Pdlla_3628	410/95	498	24(4.8)	43.53	Yes
	Pdlla_4301	410/95	877	27(3.1)	76.51	No
	Pdlla_8854	410/95	342	3(0.9)	27.67	No
	Pdlla_11090	AWPP4BI10	401	16(4)	33.65	Yes
	Pdlla_11222	AWPP4BI10	158	6(3.8)	12.71	No
	Pdlla_1717	AWPP4BI10	645	24(3.7)	56.64	No
	Pdlla_3029	AWPP4BI10	498	24(4.8)	43.53	Yes
	Pdlla_8336	AWPP4BI10	342	3(0.9)	27.67	No
A. pisi	A pisi_10786	AP1	391	19(4.9)	33.58	Yes
	Apisi_10788	AP1	257	11(4.3)	21.17	No
	Apisi_11254	AP1	217	8(3.7)	17.86	Yes
	Apisi_11282	AP1	139	6(4.3)	11.82	Yes
	Apisi_7926	AP1	364	3(0.8)	29.57	No
	Apisi_9426	AP1	120	2(1.7)	9.51	Yes
	Apisi_9612	AP1	340	13(3.8)	30.17	No
	Apisi_2488	Georgia-12	340	13(3.8)	30.13	Yes
	Apisi_ 4812	Georgia-12	347	2(0.6)	28.22	No
	Apisi_ 5867	Georgia-12	120	2(1.7)	9.51	Yes
	Apisi_ 9529	Georgia-12	238	5(2.1)	20.49	Yes

Table 13 Continued…

Organism	Protein Id/ known effector	Isolate	Length	Cysteine* count (%)	Molecular weight	Secreted
	Apisi_ 4297	Georgia-7	120	2(1.7)	9.51	Yes
	Apisi_ 6618	Georgia-7	364	3(0.8)	29.60	No
	Apisi_ 8916	Georgia-7	340	13(3.8)	30.17	Yes
Ph. koolunga	Phkol_ 4171	FT04040	343	3(0.9)	27.85	No
	Phkol_8631	FT04040	120	2(1.7)	9.78	Yes
	Phkol_9728	FT04040	365	16(4.4)	32.02	No
	Phkol_9842	FT04040	199	7(3.5)	16.11	Yes
	Phkol_10520	FT0713	216	9(4.2)	17.53	Yes
	Phkol_11622	FT0713	185	3(1.6)	15.47	No
	Phkol_2974	FT0713	120	2 (1.7)	9.78	Yes
	Phkol_3309	FT0713	343	3 (0.9)	27.85	No
	Phkol_3738	FT0713	371	16 (4.3)	32.81	Yes
	Phkol_714	FT0713	264	10 (3.8)	22.17	No
	Phkol_8714	FT0713	321	5 (1.6)	28.45	No
	Phkol_9452	FT0713	365	16 (4.4)	32.02	No
	Phkol_9543	FT0713	580	29 (5)	49.17	No
	Phkol_9893	FT0713	319	15 (4.7)	27.45	Yes
C. fulvum	Ecp6	-	228	9 (3.9)	19.64	Yes
	avr4	-	135	8 (5.9)	12.13	No
C. lindemuthianum	CIH1	-	230	6 (2.6)	20.39	Yes
M. graminicola	Mg2LysM	-	97	4 (4.1)	8.04	Yes
	Mg3LysM	-	232	9 (3.9)	19.74	Yes
M. fijiensis	gi|145259376	-	413	8 (1.9)	32.23	Yes
M. oryzae	Slp1	-	162	6 (3.7)	13.84	Yes
	Slp2	-	285	6 (2.1)	22.12	Yes
S. sclerotiorum	SS1G_03535	-	228	8 (3.5)	19.56	Yes
A. niger	gi|145259376	-	224	8 (3.6)	19.22	Yes
	gi|350638810	-	190	6 (3.2)	16.58	No

*values in the bracket indicate percent cysteine content.

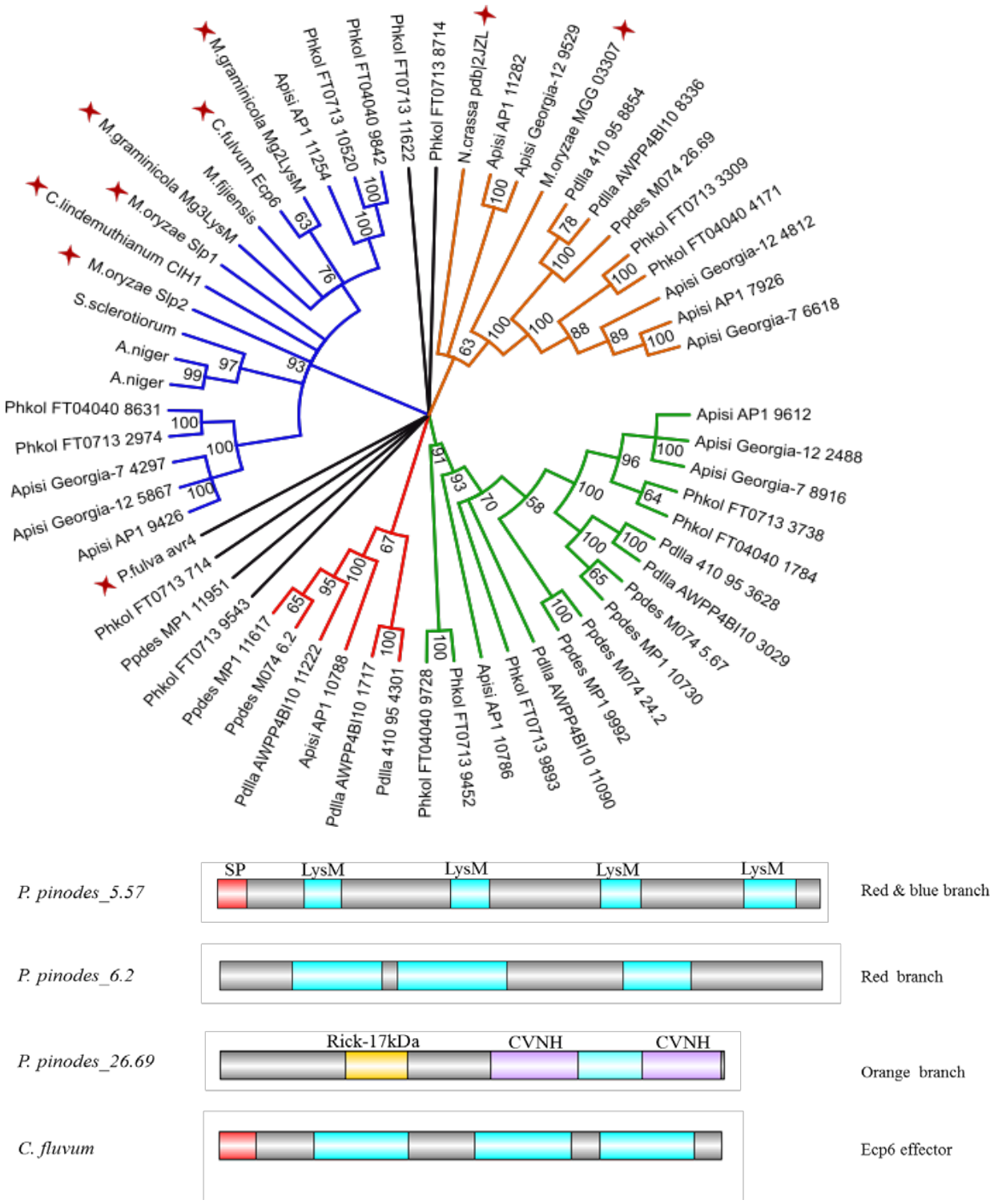


Figure 19 Phylogenetic relationship of LysM effectors from *P. pinodes*, *P. pinodella* (A) and *Ph. koolunga* and their domain organization (B)

Colored branches indicate relatedness of protein domains as depicted in figure 19B. SP: Signal peptide, LysM: Lysin motif, CVNH cyanovirin-N homology. Functionally characterized LysM are indicated by ✦

5.3.2 Pathogenicity related LysM genes exist in cluster in ascochyta blight pathogens of field pea

LysM-domain proteins of ascochyta blight pathogens of field (ABPFP) also frequently contain other domains including F-box protein, GH18 and lipase is *P. pinodes* and *P. pinodella*. Lipase was substituted with MFS and Py-redox in *A. pisi* and *Ph. koolunga,* respectively (Figure 19). The organization of these domains might have a role in disease development. Previous research in a number of plant pathogenic fungi including *F. oxysporum* [76]; *Fusarium oxysporum* [217], *Fusarium graminearum (Fbp1*) (Jonkers et al., 2011), *M. oryzae* (*Pth1*) [303] indicated that the F-box domain, while involved in a broad range of biological functions, can in some cases be important for fungal virulence on their respective host.

The outermost plant surface is the first line of defence against fungal invaders. The epicuticular layer of pea leaves contain 74 - 83% total wax on leaf surfaces [108]. This may indicate that pathogens which infect pea are likely able to degrade these lipids on the leaf surface. In fungal pathogen *B. cinerea* and *A. brassicola*, lipase helps the fungus to penetrate the host or adhesion to host surface, respectively [20, 256]. Similarly, lipase act as a virulence factor and contribute to infection and disease development in *F. graminearum* [326].

Many fungi encode chitinase proteins of the GH18 CAZyme family [278]. Fungal chitinases are involved in decomposition of exogenous chitin as well as fungal cell wall degradation and remodelling [130]. Since perception of fungal chitin by plant leads to activation of host defence signalling pathway [152, 328], sequestering of fungal fragments is likely to be important to reduce host recognition of PAMPs. In ABPFP, since the LysM domain containing protein sequence lack catalytic domain, co-regulation of LysM and chitinase (GH18) are essential to subvert host recognition through chitin binding and breakdown, respectively.

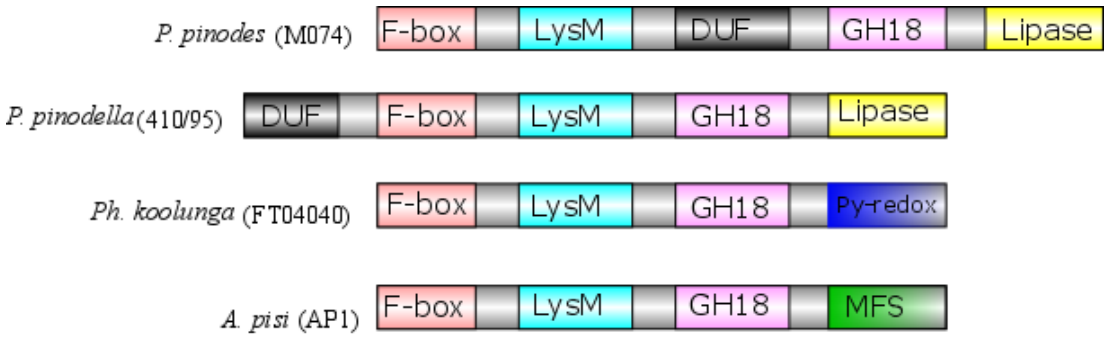


Figure 20 Organization of cluster of genes surrounding LysM in Ascochyta blight Pathogens of field pea.

DUF=domain of unknown function. Boxes with same colour indicate putative homologous genes in each species.

5.3.3 Cytochrome P450 pisatin demethylase may contribute to ascochyta blight disease development

Pisatin is an important disease resistance mechanism in pea [337]. Proteomes predicted from 25 fungal genomes were analysed and compared using Proteinortho and proteins that were orthologous to a putative pisatin demethylase P450 from *P. pinodes* M074 (*Ppdes_58.12*) were extracted. *Peyronellaea pinodes, P. pinodella* and *N. haematococca* had high numbers of P450 domain-containing proteins (10 – 12), compared to *A. pisi* (3) and *Ph. koolunga* (3). No orthologs were observed in the wheat-infecting and distantly related species *F. graminearum* and *R. solani* (Figure 21). This may imply an expansion of the pisatin-specific P450 protein in *P. pinodes, P. pinodella* and *N. haematococca*, due to a requirement to overcome pea phytoalexin (pisatin) *via* demethylation.

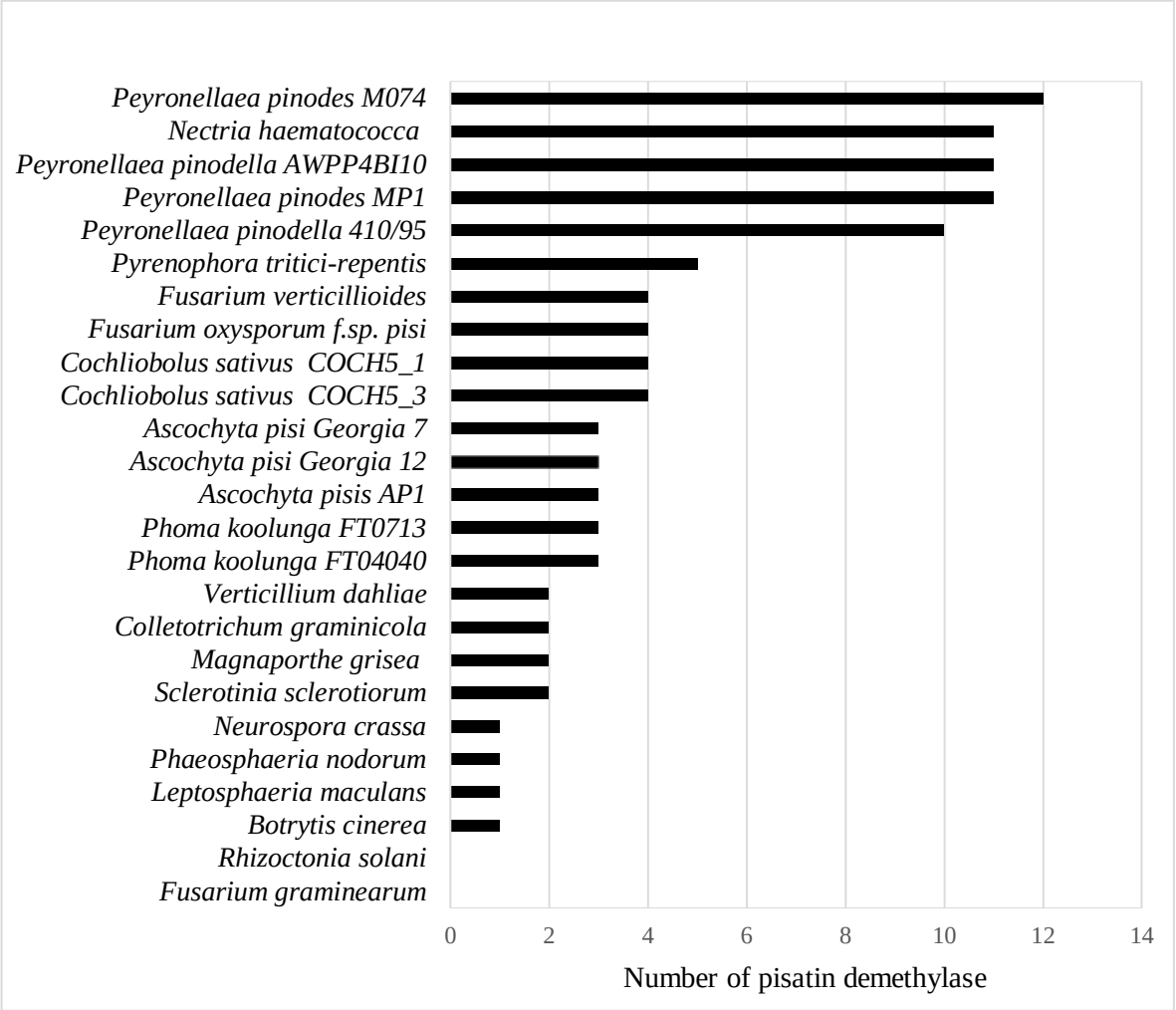


Figure 21 Comparison of P450 pisatin demethylase orthologues in various fungal pathogens

Orthologous proteins with the predicted pisatin-specific P450 of *P. pinodes* were identified *via* Proteinortho and gene counts were compared across species.

The genes responsible for detoxification of pisatin were determined as pisatin demethylase (*PDA*) [104, 105]. One of the gene named *PDA1* reside on a 1.6 Mbp accessory chromosome [124] and clustered with other pea pathogenicity genes (*PEP1, PEP2, PEP5*) where disruption of either of these genes resulted in decreased virulence [124]. Interestingly, these chromosomes on which the *N. haematococca* PDA cluster resides can be transferred into *F. oxysporum*, conferring a virulent pisatin demethylation phenotype [201, 218]. In contrast to *N. haematococca*, none of the putative *PDA* genes in ABPFP were observed to be clustered in the draft genome sequences, although it could be possible that some of the genes are located on short

scaffolds (e.g. *Ppdes_99.2* on Scaffold_99 which are less than 10kb). Further improvement of the genome assemblies will confirm the relative locations of *PDA* genes in the ABPFP genomes.

Putative PDA sequences from *P. pinodes* and *P. pinodella* were analysed against functionally characterized genes from *N. haematococca* (*PDA1, PDA6-1*) and *F. oxysporum f.sp pisi* (*FoPDA1*). Result indicated that putative PDA from *P. pinodes* and *P. pinodella* are 41 – 44% identity with over 94% sequence coverage (Table 15). Multiple sequence alignment comparison also indicated that a number of residues were conserved (Supplementary Figure 2). In addition to the P450 signature residue FGAGSRSCIG, there is a highly conserved motif (AGSDTTA) observed upstream of P450 signature, which may specify P450 pisatin (Supplementary Figure 2). This information could be utilized in comparative analysis to fish out more specific PDA related genes across genomes of various fungal pathogens deemed to undertake pisatin detoxification. Amino acid sequence composition of the two characterized PDA (*PDA1* and *FoPDA1*) indicated that the protein is rich in leucine as shown in Supplementary Figure 3).

Phylogenetic analysis was conducted to better understand the evolutionary relationship among the PDA genes. The analysis indicated that one protein sequence from *P. pinodes* (*Ppdes_58.12*), and two sequences from *P. pinodella* (*Pdlla_410/95_10685* and *Pdlla_AWPP4BI10_9093*) were closely related to the characterized PDA from *N. haematoccoca* (*PDA1*) and *FoPDA1* (Figure 22). The three sequences shared 35 – 45% and 45 – 50 % identical sites with the *PDA1* and *FoPDA1*, respectively.

Previous studies indicated that pisatin is induced by *PDA* in *P. pinodes, P. pinodella* and *A. pisi* [71, 105]. However, because of lack of DNA hybridization between the PDA from *N. haematococca* and *P. pinodes, P. pinodella* and *A. pisi*, the gene was not identified [71]. The authors showed the existence of PDA in *F. oxysporum f.sp pisi* using the same techniques. In current studies, at least two PDA genes from *N. haematococca* and *F. oxysporum f.sp pisi* were highly similar to predicted PDA genes

from *P. pinodes* (*Ppdes_58.12*) and *P. pinodella* (*Pdlla_410/95_10685* and *Pdlla_WPP4BI10_9093*) at the amino acid level. Given the high specificity for pisatin induction of *PDA1* and *FoPDA1* [47, 105] and their close relationship with predicted genes from *P. pinodes and P. pinodella,* it is likely that they contribute to virulence of ABPFP. On the other hand, the three pea pathogenicity genes (*PEP1, PEP2, and PEP5)* from *N. haematococca* formed monophyletic groups.

Table 15 BLASTp analysis of *P. pinodes and P. pinodella* putative pisatin demethylase with functionally characterized genes from *N. haematococca* and *F. oxysporum f.sp pisi*

	Gene name*	Identical site (%)	Query coverage (%)
Ppdes_M074_58.12	*PDA6-1*	44.1	94.5
	PDA1	43.4	97.3
	FoPDA1	44.4	94.9
Pdlla_41095_10685	*PDA6-1*	41.3	94.9
	PDA1	40.9	96.9
	FoPDA1	41.5	95.8
Pdlla_AWPP4BI10-9093	*PDA6-1*	44.1	94.5
	PDA1	42.8	97.3
	FoPDA1	43.9	94.9

**N. haematococc*a PDA (PDA1, PDA6-1); *F. oxysporum f.sp pisi* PDA (*FoPDA1).*

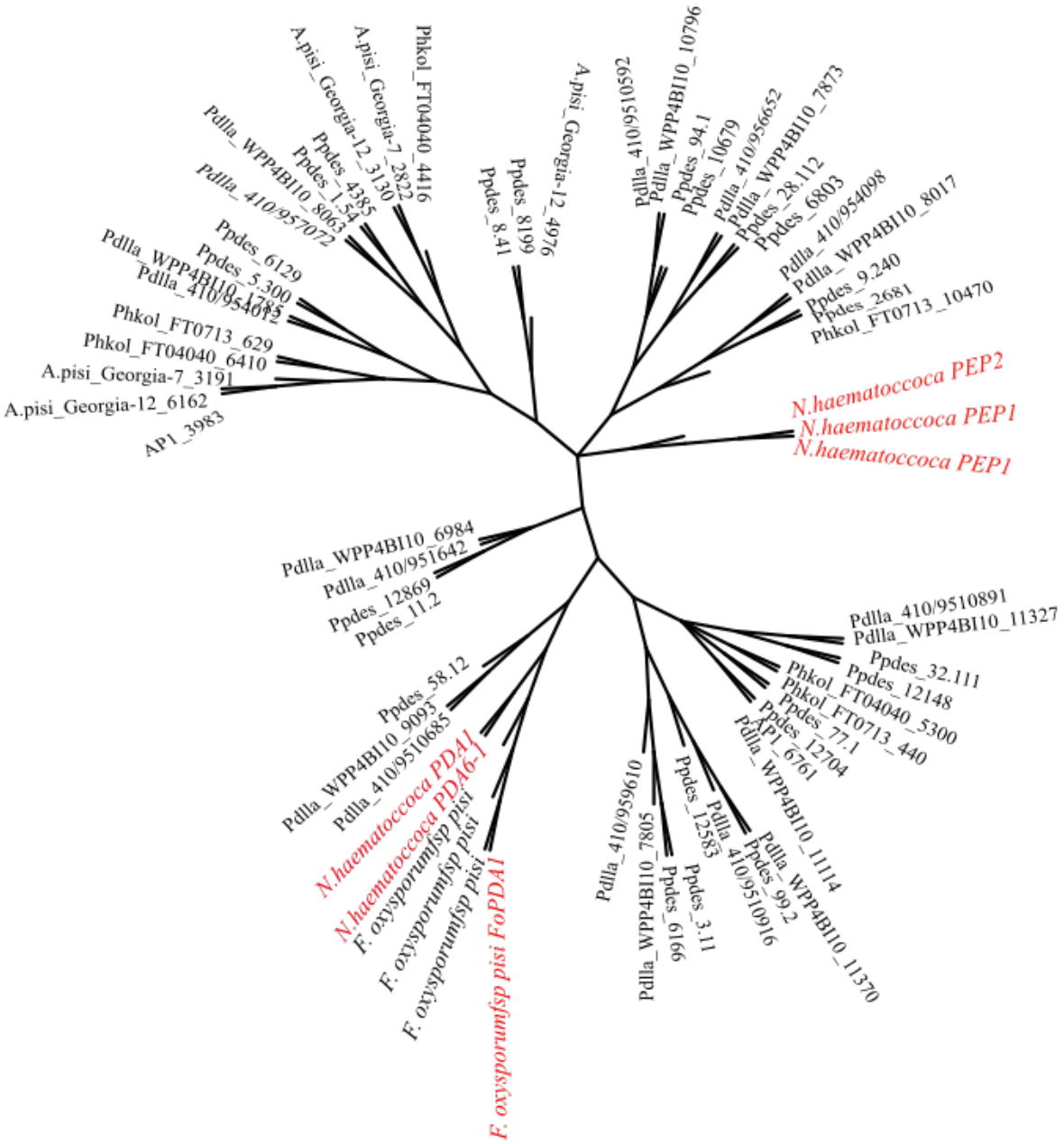


Figure 22 Phylogenetic relationships of putative pisatin demethylase inferred from Neighbor-joining consensus tree constructed from 1000 replicates using Poisson correction model

Characterized protein sequences from *N. haematococca* and *F. oxysporum f.sp pisi* were indicated in red font. Branches with less than 50% bootstrap support were collapsed. The phylogenetic was constructed from 56 amino acid sequences following the Poisson correction model implemented in MEGA6 after removal of all ambiguous positions from each sequence pair.

5.4 Conclusion

Great effort across the discipline of plant pathology has been made to identify and elucidate the possible roles of conserved protein domains related to fungal virulence. The conserved LysM domain is widespread across the Ascomycetes [3, 32, 62]. Some LysM domain proteins were demonstrated to act as fungal effectors and contribute to virulence via different mechanisms [26, 63, 144, 216, 219]. The occurrence of 4 conserved LysM effector-like protein sequences and a single CVNH domain sandwiched between the LysM domains were identified in the genomes of the ascochyta pathogens of field pea (ABPFP). Comparative sequence and phylogenetic analyses presented here will contribute to the potentially increased importance of LysM effectors in ABPFP disease development in pea.

Genomic regions containing LysM revealed the presence of other clustered genes nearby. The putative functional domains of these genes included F-box, GH18 and major facilitator superfamily and/or domain of unknown function. It is likely that at least the LysM motif, F-box, GH18 and lipase could potentially contribute to ABPFP virulence, and their collinearity may suggest transcriptional co-regulation with the LysM-containing proteins. All of these genes are biologically relevant to fungal pathogens at early to late stages of infection. It would be interesting to validate the contribution of these individual genes during infection and disease development.

Plant phytoalexin detoxification is considered to play an important role in plant pathogen infection and subsequent disease development in several phytopathogens [47, 105, 236, 237]. Through comparative and phylogenetic analysis, I identified and narrowed down the gene likely responsible for pisatin detoxification in ascochyta blight pathogens of field pea. The analysis may also imply that pathogens infecting field pea (*N. haematococca, P. pinodes, P. pinodella* and *A. pisi*) have adapted to their niche through evolution of phytoalexin detoxification. Further functional studies are required to confirm the role of potential PDA identified in this study to develop pre-breeding tools in the complex pea improvement program.

6. *In planta* Transcriptome and Secretome Analysis of Ascochyta Blight Pathogens of Field Pea

6.1 Introduction

The ease of sequencing a wide range of phytopathogens at relatively low cost and continuous improvements in bioinformatics have accelerated pathogenomic studies investigating the molecular basis of interaction between the pathogen and their host. *Peyronellaea pinodes*, *Peyronellaea pinodella* and *Phoma koolunga* are destructive necrotrophic fungal pathogens of field pea. Necrotrophic fungi such as these secrete effectors to breach host-defence responses ahead of colonization [41, 93, 308]. Effectors are proteins or small metabolites that promote disease through altering host cells structure or and function [137, 154]. This chapter focusses on proteinaceous effectors, as these are the most amenable to bioinformatic prediction.

According to Lo Presti *et al.* [197], effectors can either promote the virulence of fungi or allow symbiotic-association with the plant. In *Pyrenophora tritici-repentis*, *Parastagonospora nodorum, Cochliobolus victoriae,* host-selective toxins are often essential to infect their respective hosts [43, 91, 320]. On the other hand, some necrotrophic pathogens with broader host-ranges like *Botrytis cinerea* and *Sclerotinia sclerotiorum*, *Colletotrichum coccodes*, *Abutilon theophrasti*, and *Rynchosporium commune* and *Alternaria alternata*, [9, 83, 165, 322], kill their hosts by secreting toxic metabolites and proteins with activity across multiple plant hosts.

Plant pathogenic fungal effectors are often under selection pressure to diversify their sequence or structure due to avoid potential recognition by host defences, yet despite their diversity share some common characteristics. The current repertoire of known fungal effectors has been previously reviewed [61, 65, 197, 267, 290, 295, 324]. Some fungal effectors may undergo enhanced rates of mutation [231], exhibit dynamic structural evolution [197, 229, 266, 305], or exhibit mutations that indicate under

diversifying selection [292]. They may also be horizontally-transferred between pathogens [259, 271, 301]. Regardless, most of the known proteinaceous effectors share common characteristics of being secreted, cysteine-rich and low molecular weight and often lacking sequence similarity to any other known proteins [61, 65, 258, 295].

The bioinformatics prediction of effectors varies slightly between studies, but often shares common parameters. Generally, it has been common practice to filter the complete set of proteins for those that are cysteine-rich (e.g. >2 cysteines), predicted to be secreted, and of less than 300 amino acids in length should be considered strong effector candidates [119, 274, 290, 305]. Optionally this set of proteins may be filtered for those without detectable orthologous beyond the genus under investigation or often lacking conserved domains (i.e. no Pfam matches). However a recent review [197] proposed that some effectors may also be larger than 300 amino acids and could also contain functionally conserved protein domains; thus they considered all secreted proteins as potential candidates, or used multiple and flexible criteria that allowed some conditions not to be met [126]. Sperschneider *et al.* [290] also indicated that many cysteine rich, small, secreted proteins are not effectors, suggesting discrimination through further integration of several lines of evidence including *in planta* expression, diversifying selection and deeper analysis of the taxonomic range of homologs.

Regardless of the various challenges and unknowns in effector prediction, the prediction and characterisation of effector candidate genes plays pivotal role in molecular plant breeding. Unbiased prediction and characterisation of effector genes can lead to rapid and effective genotype screening and disease monitoring using either detection of effector gene sequences or assay of responses to effector protein infiltration as molecular markers [103, 324], For instance, semi purified ToxA protein from *Pyrenophora tritici-repentis* is being delivered to breeders as a functional assay to screen commercial cultivars in Australian breeding programs [324]. Knowledge of the sequences of *Avr* effectors in *Leptosphaeria maculans* is also being employed for pathogen monitoring and cultivar selection in Australia [324]. There are also prospects

to apply similar methods to for crops where resistance is only partial, as in the field pea-ABPFP interaction. Reliable identification of effectors and their cognate host-interactome will help breeders to specifically exploit partial resistance.

As a first step in the exploration of the effector content of *P. pinodes*, Kessie [158] predicted 16 potential small-secreted proteins identified from culture filtrate through a proteomic approaches. Two of the proteins (DPIT_05229 and DPIT_10135) identified from culture filtrate that induced necrosis on pea were reported promising effector candidates. Kessie [158] also predicted a number of additional small-secreted protein sequences through *in silico* analysis of a preliminary draft genome sequence of *P. pinodes* draft genome sequence. However, the expression of these predicted putative effectors during infection and disease development is not known. Since then, the genome assembly of *P. pinodes* has been significantly improved, through the incorporation of new data from multiple fragment-length libraries (chapter 3). In this study, we have also employed *in planta* RNA sequencing approach to identify an updated catalogue of effector candidate genes from an improved version of *P. pinodes* genome assembly, as well as new catalogues of effector candidates for the new and draft genome assemblies of *P. pinodella* and *Ph. koolunga*.

6.2 Materials and Methods

6.2.1 Infection assay and sample collection

The seeds of pea cultivar KASPA, which is commonly grown in Australia and susceptible to *P. pinodes*, were used in disease infection assays. The seeds were surface sterilised prior to infection in 3% bleach for three minutes, rinsed three times in sterile water (autoclaved) to remove the bleach. Seeds were transferred to damp petri dishes lined with tissue paper and incubated at room temperature for two days. Uniformly germinated seeds were planted in sterile vermiculite and perlite mix (3:1) (autoclaved) filled in a porous plastic boxes. The mixes were moistened with water and autoclaved

before use. The plants were regularly watered to keep the growth mixtures moist. Grow A and B fertilizers were supplemented to the irrigation water at a rate of 20 ml of each fertilizer mixed in 5 litre of water twice a week throughout the experimental period.

A spore suspension was prepared from 10 days old culture grown on 50% potato dextrose agar (18.5 gram of potato dextrose, 15 gram of agar mixed in 1 litre of molecular water) by flooding with 3-5 ml sterile water followed by gentle scraping using a glass rod to detach the pycnidiospores. The suspension was filtered by passing through a sterile cotton plug fitted into a 50 ml syringe to remove any mycelial fragments. The spore concentration was determined using haemocytometer and adjusted to a concentration of $5x10^5$ spore/ml. Tween 20 was added to the spore suspension at a rate of approximately 2% prior to inoculating the plants.

Four weeks old plants were inoculated by spraying the spore suspension until "run off", sprayed manually by hand. The inoculated plants were immediately covered with black plastic sleeves, misted with water and transferred to a controlled growth chamber 'Conviron' (Argus Control System Ltd, Canada) adjusted to 20^0 C ± 2 and relative humidity of 60%. The temperature and RH within the plastic sleeves were monitored using a hygrometer sensor. The inoculated plants were uncovered 24 hours post inoculation and retained in the 'Conviron' with but sprayed with sterile water to increase leaf surface wetness. An initial experiment consisted of *P. pinodes* infection assay with three replication and nine time points. The second and third experiment consisted of *P. pinodes*, *P. pinodella* and *Ph. koolunga* with three replications and nine time points. Mock-inoculated plants were included in all the experiments as negative control in all experiments. The experimental setup is shown in Table 16.

Samples were collected immediately after inoculum spraying (0 HPI), as well as 3, 6, 9, 21, 33, 48, 72 and 96 HPI post inoculation and again at 4 days post infection for both microscopic examination and RNA extraction. For RNA extraction, the first two nodes of each plant with stipules were cut with sterile scissors, wrapped in aluminium foil, immediately dried in liquid nitrogen and stored at -80^0C until RNA extraction.

6.2.2 RNA extraction, processing and sequencing

Total RNA was extracted as per the Trizol extraction protocol [260]. Collected plant tissues were ground using an oven-baked (150^0C, 2 hours) RNase-free mortar and pestle under liquid nitrogen. About 100 – 150 mg of each ground tissue was transferred to 1.5 ml pre-cooled Eppendorf tubes and 1 ml of Trizol was immediately added, vortexed 3 times and incubated at room temperature (RT) for 15 minutes. The samples were then centrifuged for 10 minutes at 12000 revolutions per minute (rpm). After transferring the upper liquid phase to new 1.5 ml tubes, 200 µl cold chloroform was added and gently mixed for 15 seconds followed by incubation for 5 minutes at RT and then centrifuged for 15 minutes at 12000 rpm. The upper phase was transferred to a new 1.5 ml tube and RNA was precipitated by adding 300 µl of 1.2M NaCl and 300 µl isopropanol incubated at RT for 10 minutes and subsequently centrifuged for 10 minutes at 12000 rpm. Finally, the upper liquid phase was decanted and the RNA pellet washed twice in 75% ethanol. After the ethanol wash, excess ethanol was allowed to evaporate from the pellet allowed for 5 minutes at RT, resuspended in 50 µl diethylpyrocarbonate (DEPC) treated water and incubated at 60 ^{0}C for 10 minutes to dissolve the RNA. The integrity and quality of the RNA was analysed on 1% agarose gel electrophoresis after CYBR Safe staining. Each of the samples were treated with DNase to remove plant DNA contaminants from extracted RNA samples and checked by running a polymerase chain reaction (PCR) using primer that amplified elongation factor 1-α (*EF1-α*) of *Pisum sativum* with 5'-CGACTCTGGAAAATCAACAACAACTGG-3' and 5'-GTCCATCCTTAGAGATACCAGCTTCAA-3' forward and reverse primers, respectively[318].

Based on biological observation of fungal development during infection and disease establishment (Supplementary Figure 4), cleaned RNA samples from 0, 3, and 6 HPI (including three replications each) were pooled to represent "early stage infection". Similarly, RNA samples from 9, 21, 33 HPI and 48HPI were pooled to represent "medium stage infection", and 48, 72 and 96 HPI infection were pooled to represent "late infection" (Table 17). The RNA pooling was repeated for the three experiment

and the three replications. Final samples consisted of nine libraries for *P. pinodes* six libraries for each *P. pinodella* and *Pho ma koolunga* and (Table 17, Figure 23).

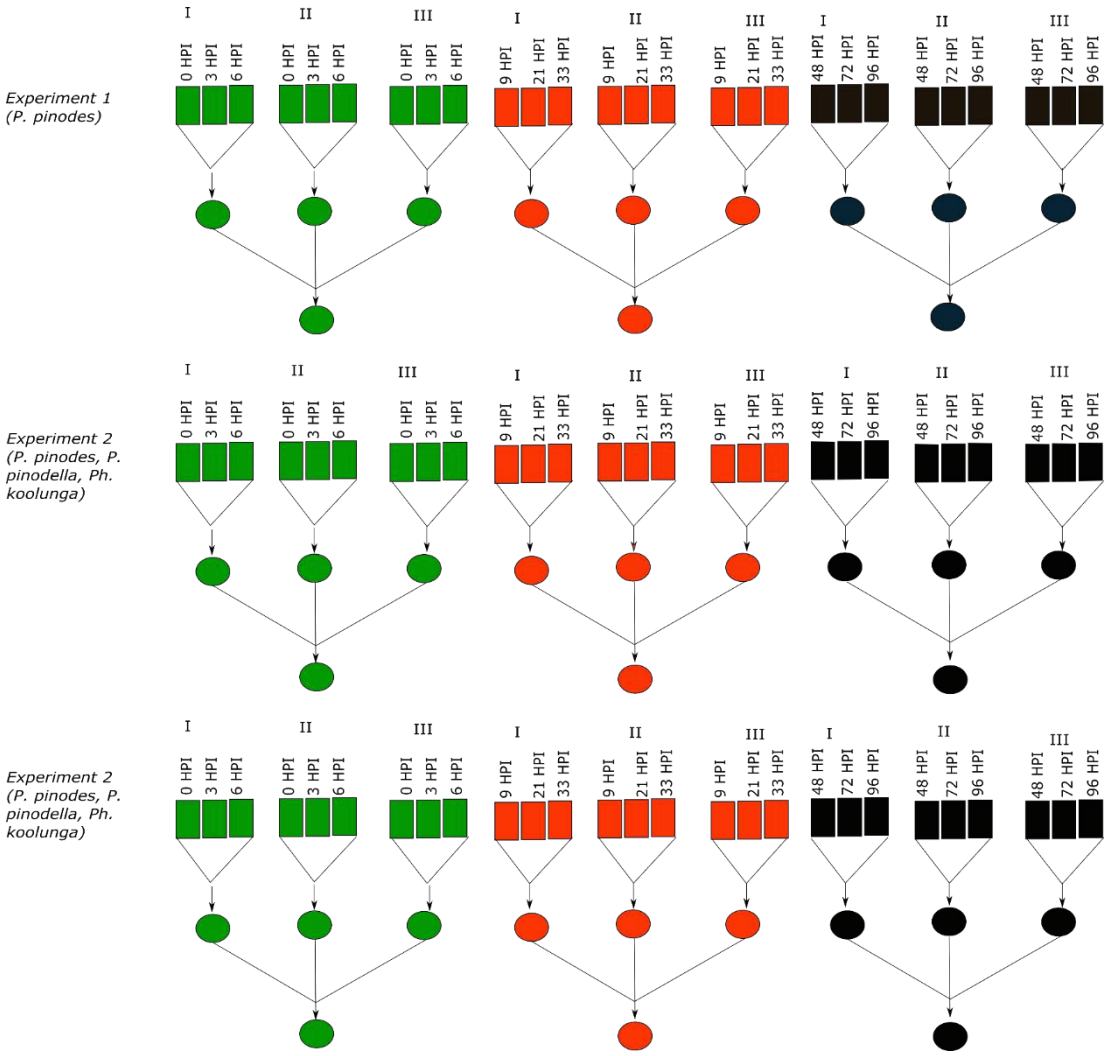


Figure 23 RNA sample preparation procedures employed for in planta transcriptome sequencing.

Green, red and black colours indicate early, medium and late time points, respectively.

Table 16 Experimental setup of in planta infection assay of *P. pinodes*, *P. pinodella* and *Ph. koolunga*

	*P.pinodes**	*P. pinodella*	*Ph. koolunga*	Mock*
Number of experiments	3	2	2	3
Number of replication per experiment	3	3	3	3
Spore concentration	$5x10^5$	$5x10^5$	$5x10^5$	-
Number of initial time points for infected leaf collection	9	9	9	9
Total samples collected	81	54	54	81
Total samples after pooling for RNA-seq	9	6	6	9

*The first experiment included only *P. pinodes* and mock

RNA of *in planta* libraries sequencing was performed at the Australian Genome and Research Facility (AGRF) in Melbourne, Australia using an Illumina HiSeq 2500 with 100 bp read length. Illumina TruSeq Stranded mRNA sample preparation was performed at AGRF as part of their sequencing services. *In vitro* RNA library preparation was undertaken following Illumina TruSeq unstranded RNA sequencing protocol. Transcriptome sequencing conditions were summarized in Table 16.

The RNA sequence quality assessment was carried out as described in Chapter 2 section 2.2.1. Reads were mapped to the reference genome and splice sites were predicted using TopHat2 v2.0.9 [161] (parameters: minimum intron length 20, maximum intron length 5000, no coverage search, splice mismatches 2, very sensitive, read mismatches 2, segment mismatches 2, maximum insertion length 3, maximum deletion length 3, report secondary alignments, library-type first-strand). Differential expression analysis was performed using Cuffdiff (Cufflinks version 2.1.1) [319] with minimum alignment count of 5 reads and multi read correction settings. Comparisons

between *in planta* gene expression libraries only were time dependent and considered as time series in Cuffdiff analysis. The data from each pooled time points (early, medium and late) was also re-analysed *via* Cuffdiff as an independent experimental conditions for effector candidate gene prediction, in which *in vitro* libraries were also included.

Table 17 Transcriptome sequencing conditions and library type used in this study

In vitro Library	Condition	Library type*	Organism
F2-3D	Fries 2 media 2 days old	U, paired, 100 bp	*P. pinodes*
F3-3D	Fries 3 media 3 days old	U, paired, 100 bp	*P. pinodes*
F2-W4	Fries 2 media 4 weeks old	U, paired, 100 bp	*P. pinodes*
F3-W4	Fries 3 media 4 weeks old	U, paired, 100 bp	*P. pinodes*
In planta conditions			
Early	0, 3 and 6 (HPI)	TruSeq, S, paired, 100 bp	*P. pinodes, P. pinodella, Ph. koolunga*
Medium	9, 21, 33 HPI	TruSeq, S, paired, 100 bp	*P. pinodes, P. pinodella, Ph. koolunga*
Late	48, 72, 96 HPI	TruSeq, S, paired, 100 bp	*P. pinodes, P. pinodella, Ph. koolunga*

*S= stranded, U= Unstranded. HPI= Hour Post Inoculation

6.2.3 Effector candidate gene prediction

Genes were filtered and ranked based on their likelihood of encoding a proteinaceous effector, which was predicted based on multiple factors. In the effector candidate gene prediction, the translated protein sequence was assumed to be secreted if it was predicted as such by either, SignalP4.1 [243], Phobius 1.01 [153] or WOLFPsort 0.2 [139] and also requiring at most one transmembrane domain predicted by TMHMM 2.0 [172]. Predicted secreted proteins were then ranked according to various predicted properties, including: molecular mass of 50 kDa or less, lack of known conserved domains at Pfam database [250] analysed through HMMER3 [77], either diversifying ($dN/dS \geq 1.2$) or purifying ($dN/dS \leq 0.8$) selection, species and isolate specificity,

when comparing whole-genome alignments of multiple isolates of the same species (Chapter 2). All of the above properties were assigned a score of 2. Similarly, proteins with 2 - 4 cysteines, or greater than 4 cysteines were scored 1 and 2, respectively. Significantly upregulated genes at early or medium time points were assigned a score of 3, and expressed genes with no significant *in planta* expression were assigned a score of one. In the case of *P. pinodes*, where there were also 4 *in vitro* and three *in planta* conditions, additional scores relating to genes that were upregulated *in planta* relative to each of the *in vitro* conditions were also scored 3, or were scored 1 if expressed *in planta* but not significantly up-regulated relative to the *in vitro* libraries.

6.3 Results and Discussion

6.3.1 The genome ABPFPs contain similar level of secretome

For predicting a set of candidate effector genes from the predicted proteome in *P. pinodes*, *P. pinodella* and *Ph. koolunga* and stringent parameters were applied. Secreted protein prediction (see methods) indicated that a total of 1681, 1451 and 1518 protein sequences were secreted in *P. pinodes*, *P. pinodella* and *Ph. koolunga,* respectively. Further ranking based on multiple predictors, including: *in planta* expression, presence/absence of known domains, species/isolate specificity, protein size, cysteine content and purifying/diversifying selection pressure were applied in the effector selection, with each property assigned a score and each protein ranked on the sum of these scores. A plot of total scores indicated that a high-priority candidate effector-like gene dataset of reasonable size (feasible for future experimental validation) could be obtained by applying scoring cut-off values of 8 in *P. pinodes*, 12 in *P. pinodella* and 11 in *Phoma koolunga* (Figure 24). These thresholds resulted in the short-listing of high priority 223 effector like candidates in *P. pinodes*, 226 in *P. pinodella* and 140 proteins in *Ph. koolunga* (Table 18).

Protein domain analysis using the Pfam database [250] through HMMER3 [77] showed that approximately 60%, 34% and 80% of the effector like sequences in the

respective species had either no known Pfam domains, matched uncharacterized proteins or corresponded to conserved protein domains of unknown function (Table 18 , Supplementary Table 8). Furthermore, 8.9 % (21) of shortlisted candidates in *P. pinodes*, 3.7% (8) in *P. pinodella* and 47% (67) in *Ph. koolunga* were species-specific effector like proteins, as previously identified by cross-species orthology analysis (Chapter 2).

Table 18 Secreted effector like candidate proteins of *P. pinodes, P. pinodella* and *Ph. koolunga*

	P. pinodes	*P. pinodella*	*Ph. koolunga*
Total genes/protein annotations	11,352	11,058	10,084
Predicted secreted proteins	1,681	1,451	1,518
Percent of proteome predicted as secreted (%)	14.8	13.1	15.1
Predicted candidate effector like proteins (CELPs)	223	212	140
Species-specific CELPs	21	8	67
CELPs with unknown function/domain	136	78	113
Percent of CELPs with unknown function/domain (%)	60	36.7	80

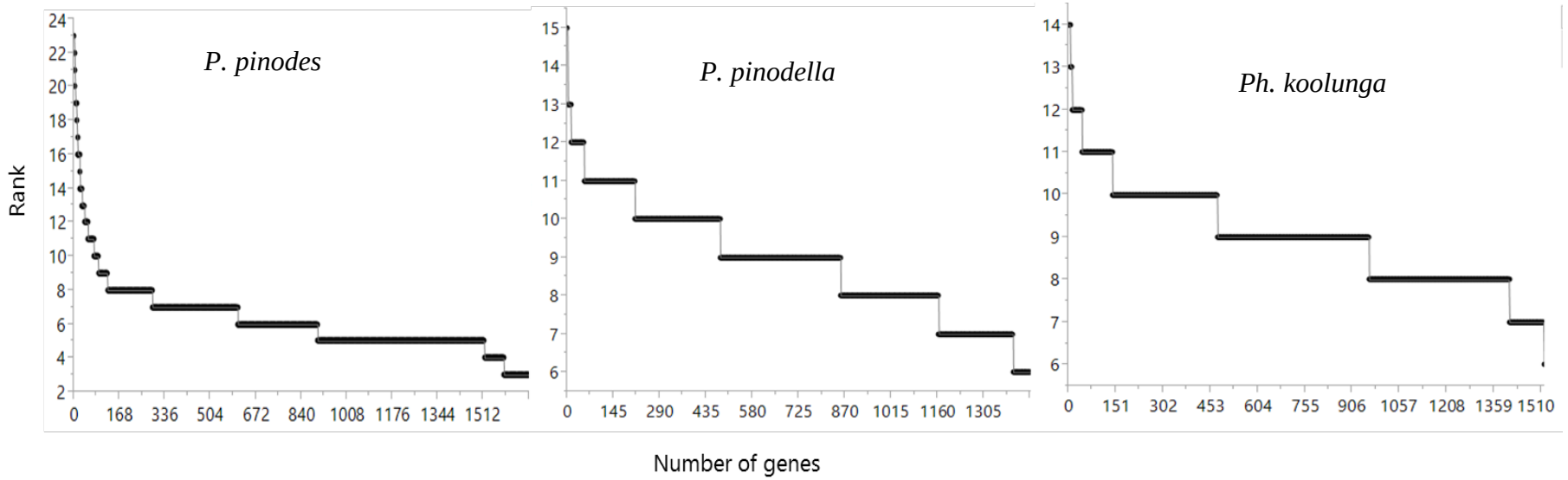


Figure 24 Graph showing the distribution of effector like candidate genes in *P. pinodes*, *P. pinodella* and *Ph. koolunga.*
The ranked score date was plotted as geometric distribution of occurrence of rare events. Vertical axis is scoring scale, horizontal axis is number of genes.

Effector-like gene prediction using two methods: 1) evidence-based ranking of cumulative scores derived from the presence or absence of a combination of previously known effector characteristics and 2) using a recently developed probabilistic effector prediction tool, EffectorP [291]. The analysis showed a low level of agreement between the two techniques (Figure 25). However, in *P. pinodes*, a comparison between cumulative ranking and EffectorP probability showed that the higher the cumulative ranking score, the more likely proteins would also be predicted as an effector by EffectorP (Figure 26). This trend is similar but relatively weaker in *P. pinodella* and *Ph. koolunga*, which lack *in vitro* RNA-seq data and notably consisted of low coverage *in planta* RNA-seq data. This may indicate the importance of RNA-seq data generated under different conditions to improve the accuracy of initial gene predictions in in *P. pinodella* and *Ph. koolunga*, which in turn would improve effector prediction accuracy. EffectorP could be useful in predicting effector-like candidate genes spanning average length of about 400 amino acids as employed in training the program [291]. Notably, EffectorP was effective in distinguishing the necrosis-inducing proteins from non-cytotoxic paralogues genes in *P. pinodes* and *P. pinodella*. However, because late-expressed genes were given a lower priority score relative to other time points in the evidence-based ranking method, known domains like NPP1 that are associated with necrosis-inducing proteins were not picked up.

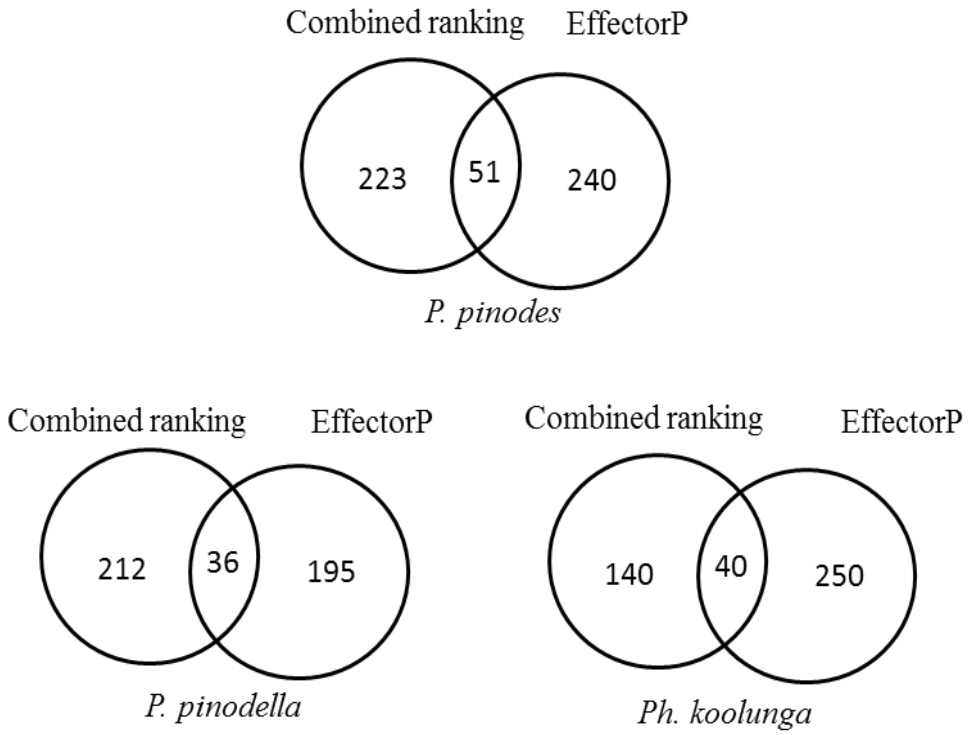


Figure 25 Venn diagram representing effectors predicted using two methods 1) evidence-based rankings and 2) EffectorP.

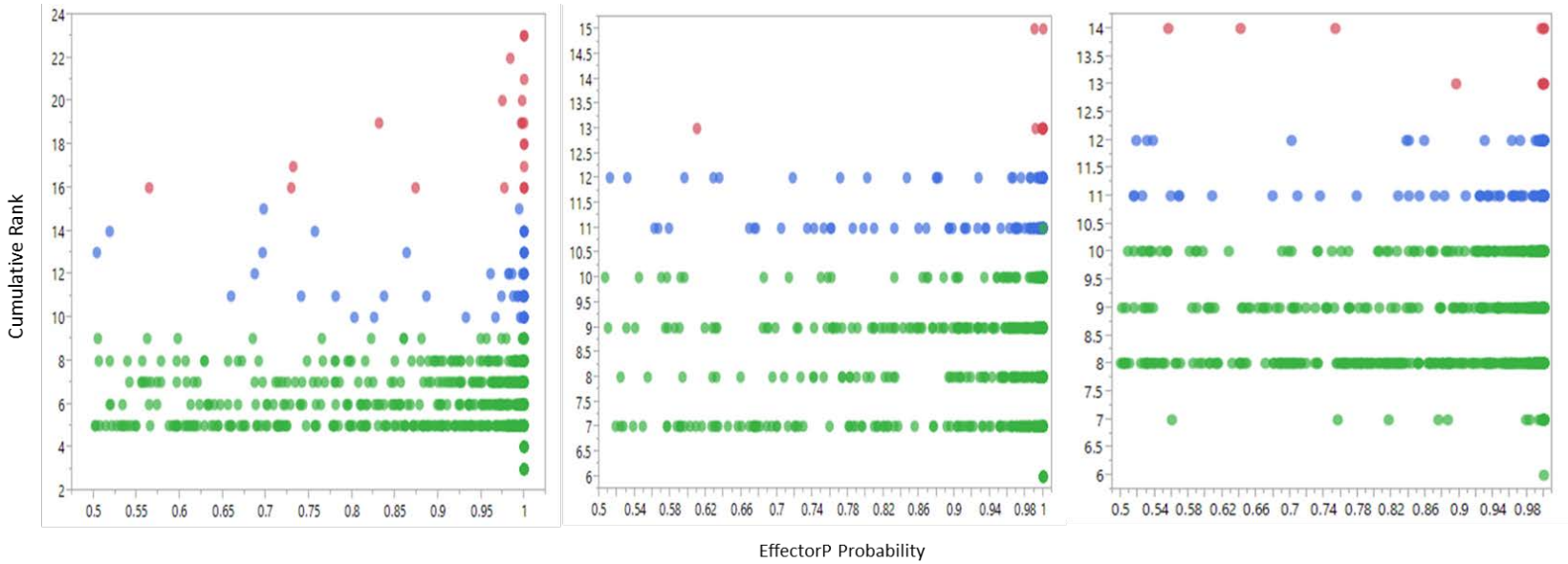


Figure 26 Scatterplot showing the level of agreement between effectors predicted using evidence-based cumulative ranking and EffectorP.

From left to right *P. pinodes*, *P. pinodella* and *Ph. koolunga*, respectively.

6.3.2 *In planta* expression pattern of effector like candidate genes in *P. pinodes*

Among the 223 candidate effector-like proteins (CELPs) in *P. pinodes*, the corresponding gene expression of 44 (20%) were significantly up-regulated at early time points *in planta* relative to at least one *in vitro* conditions. The diverse range of expression patterns under different conditions for the 100 top ranked candidates predicted by the scoring method are a colour-coded heat map generated from the normalised fragments per kilo base of exon per million fragments mapped (FPKM) values (Figure 27). A group of genes including *Ppdes_16.9, Ppdes_32.5* and *Ppdes_56.67* showed a consistently high level of expression from early to late infection. Likewise, a cluster of genes consisted of *Ppdes_5.48, Ppdes_1.96, Ppdes_1.220, Ppdes_5.53, Ppdes_38.47* and *Ppdes_2.193* showed highest expression at early infection stages (Figure 27, Table 19, and Supplementary Table 8). One gene (*Ppdes_1.220*) was significantly upregulated ($P > 0.05$) at early time points relative to all four *in vitro* conditions. Protein domain analysis indicated that *Ppdes_1.220* encodes a FAD binding domain [Pfam: PF00890], which is related to succinate dehydrogenase. Succinate dehydrogenase plays central role in mitochondrial metabolism and targeted for crop fungal disease management through fungicide development [191]. Fungicides targeting succinate dehydrogenase were directed to block tricarboxylic acid cycle (TCA) cycle at early stage of succinate oxidation to fumarate, which subsequently halt respiration. On the other hand, it could be possible that succinate dehydrogenase activity may assist ABPFP to survive under oxidative stress. This scenario has been proposed for the human pathogenic prokaryote *Mycobacterium tuberculosis* [131].

Similarly, *Ppdes_2.193* showed significant up-regulation under all conditions except in F3-W4. In addition, 9 genes showed significant up-regulation at early time points relative to at least two *in vitro* conditions while the remaining 33 genes showed up-regulation in only one of the *in vitro* conditions. Most of those *in planta* upregulated genes generally belong to the families of CAZymes pectate lyase [Pfam: PF00544,

PF03211], glycoside hydrolase (GH) 61 [Pfam: PF03443], GH28 [Pfam: PF00295]), GH5 (cellulase) [Pfam: PF00150] and GH43 [Pfam: PF04616]. Esterase, tannase and feruloyl esterase [Pfam: PF07519], carboxylesterase [Pfam: PF00135], protease (subtilase [Pfam: PF00082], peptidase [Pfam: PF01828], eukaryotic aspartyl protease [Pfam: PF00026), ribosomal biogenesis (ADP-ribosylation factor family [Pfam: PF00025], ribosomal L29 protein [Pfam: PF00831]), FAD binding [Pfam: PF01494], cytochrome P450 (pisatin demethylase, cytochrome alkane) [Pfam : PF00067], fungal specific extracellular membrane proteins (CFEM domain, Pfam: PF05730) and protein domains involved in protein–protein interaction (PAN domain, Pfam: PF14295) were also included in putative effectors.

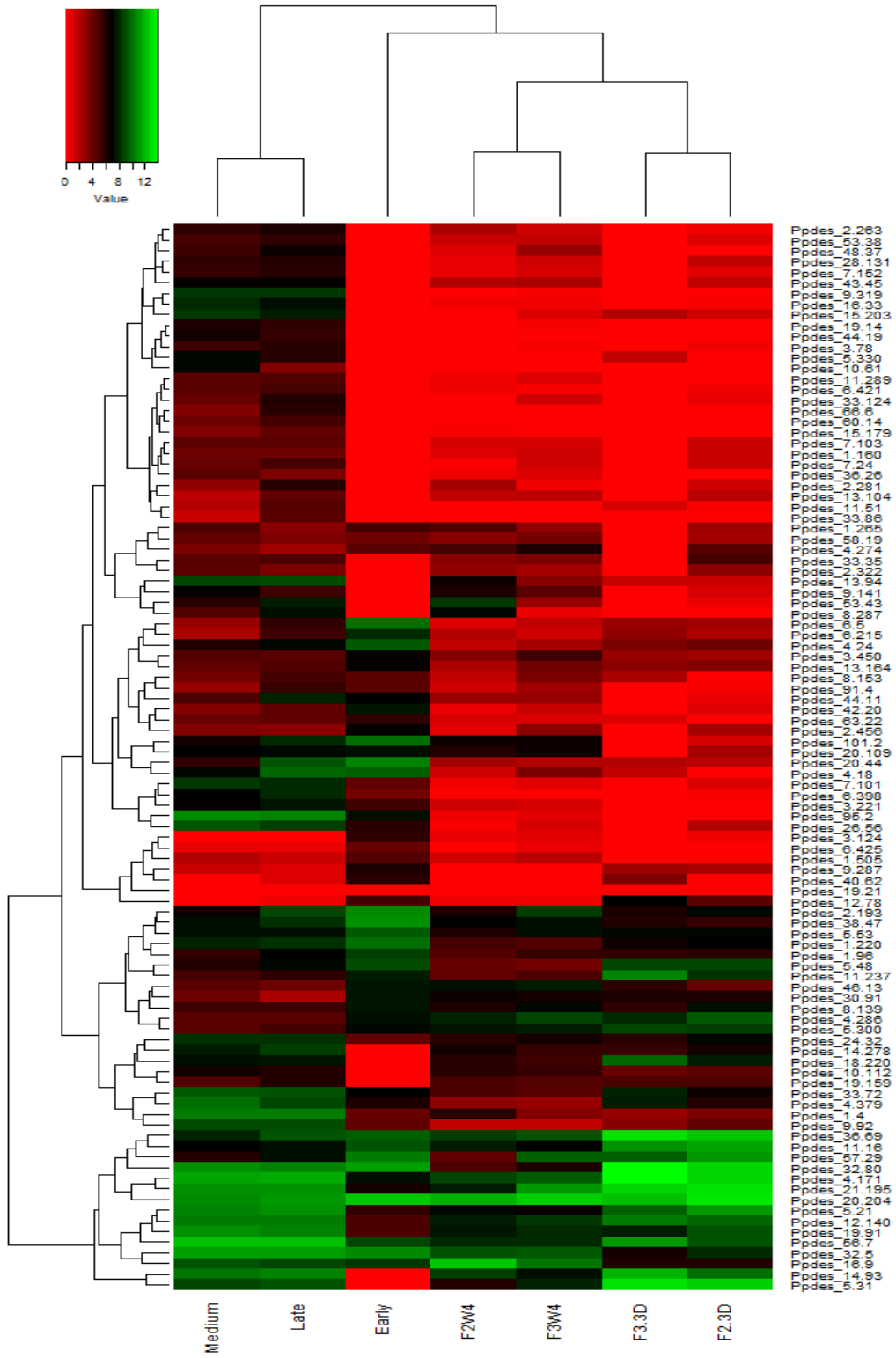


Figure 27 Heat map of relative expression levels of candidate effector-like genes in *P. pinodes* isolate M074.

Early = 0 – 9 hour post infection (HPI); Medium = 9 – 33 HPI; HPI; 48 - 96 HPI. F2-3D, F3-3D, F2W4 and F3W4 indicate Fries 2 media 2 days, Fries 3 media 2 days, Fries 2 media 3 weeks and Fries 3 media 3 weeks old, respective

Table 19 Top 60 effector-like candidate genes in *P. pinodes* isolate M074 as predicted by the evidence-based ranked score method and their associated supporting data.

Gene ID	*In vitro* FPKM			*In planta* FPKM				Amino acid length	Cysteine count	Molecular weight (kDa)	Cumulative Rank	EffectorP[a]	Pfam ID	Pfam description	Top BLAST hit description[b]
	F2W4	F3W4	F3-3D	F2-3D	Early	Medium	Late								
Ppdes_2.193	88	427	70	142	1614	114	487	457	17	48	23	N	-		cnd3 protein
Ppdes_60.14	0	0	0	0	0	15	36	634	10	69	23	N	-		saponin hydrolase precursor
Ppdes_95.2	0	1	0	0	172	1702	1401	130	6	14	22	E	PF08881	CVNH domain	hypothetical protein H072_1743
Ppdes_1.220	36	26	99	119	1061	245	302	638	4	69	21	N	PF00890	FAD binding domain	fumarate reductase
Ppdes_4.18	2	11	2	0	739	137	937	264	12	27	20	E	PF03211	Pectate lyase	pectate lyase
Ppdes_5.330	0	0	2	0	0	136	55	240	10	25	20	E	PF03211	Pectate lyase	pectate lyase
Ppdes_10.61	0	0	0	0	0	138	10	295	10	32	19	N	PF00067	Cytochrome P450	cytochrome p450 monooxygenase
Ppdes_101.2	93	96	0	1	972	82	267	115	6	12	19	E	-	-	hypothetical protein
Ppdes_32.5	617	645	84	266	1689	2421	2381	178	0	20	19	N	PF00831	Ribosomal L29 protein	60s ribosomal protein l35
Ppdes_56.7	282	285	2160	608	460	4692	5040	158	0	16	19	N	-		signal peptide-containing protein
Ppdes_44.19	0	0	0	0	0	89	42	487	7	55	18	N	PF00067	Cytochrome P450	cytochrome p450 family protein
Ppdes_6.398	0	0	0	0	13	129	280	432	4	49	18	N	PF00150	Cellulase	glycoside hydrolase family 5
Ppdes_19.91	176	266	212	624	28	1989	1524	346	2	35	17	N	PF03443	Glycosyl hydrolase family 61	glycosyl hydrolase family 61
Ppdes_40.62	0	0	12	0	60	0	1	148	0	16	17	E	PF06864	Pilin accessory protein (PilO)	Nohit
Ppdes_11.289	0	1	0	0	0	24	28	283	5	30	16	N	PF00657	GDSL-like Lipase/Acylhydrolase	rhamnogalacturonan acetylesterase
Ppdes_38.47	133	171	65	47	1854	162	296	186	1	21	16	E	PF00025	ADP-ribosylation factor family	adp-ribosylation factor 6
Ppdes_44.11	6	7	0	1	120	31	251	394	9	40	16	N	PF00295	Glycosyl hydrolases family 28	glycoside hydrolase family 28
Ppdes_48.37	1	6	0	0	0	39	101	307	6	33	16	N	PF10503	Esterase PHB depolymerase	esterase lipase
Ppdes_6.215	4	2	8	4	260	4	39	1592	35	169	16	N	PF14295	PAN domain	apple protein
Ppdes_8.287	142	1	0	0	0	27	159	230	6	24	16	N	PF03443	Glycosyl hydrolase family 61	glycoside hydrolase family 61
Ppdes_9.319	0	0	0	0	0	368	342	373	6	38	16	N	PF00089	Trypsin	serine protease
Ppdes_26.56	0	1	0	4	57	618	420	309	5	34	15	N	PF13583	Metallo-peptidase family M12B	cellulose-binding family ii
Ppdes_42.20	1	2	0	1	194	10	19	228	17	24	15	N	PF09792	Ubiquitin 3 binding protein But2	gpi anchored cell wall protein
Ppdes_19.14	0	0	0	0	0	72	50	554	7	58	14	N	PF02065	Melibiase	carbohydrate-binding module 35

Table 19 continued

Gene ID	*In vitro* FPKM				*In planta* FPKM			Amino acid length	Cystiene count	Molecular weight	Cumulative Rank	EffectorP[a]	Pfam ID	Pfam description	Top BLAST hit description[b]
	F2W4	F3W4	F3-3D	F2-3D	Early	Medium	Late								
Ppdes_24.32	61	83	58	136	19	317	315	874	8	95	14	N	PF00933	Glycosyl hydrolase family 3	beta-glucosidase
Ppdes_3.124	1	1	0	0	52	0	0	240	4	26	14	N	PF08657	DASH complex subunit Spc34	hypothetical protein SNOG_10148
Ppdes_3.78	0	0	0	0	0	39	58	287	4	29	14	N	PF00080	Copper/zinc superoxide dismutase	cytosolic cu zn superoxide dismutase
Ppdes_33.124	0	2	0	0	0	18	62	227	4	24	14	N	PF03443	Glycosyl hydrolase family 61	glycoside hydrolase family 61
Ppdes_33.72	31	23	236	96	123	665	568	454	4	49	14	N	PF03198	Glucanosyltransferase	glycoside hydrolase family 72 protein
Ppdes_5.21	144	117	799	2266	52	1512	2019	392	0	39	14	N	PF02469	Fasciclin domain	beta-ig-h3 fasciclin
Ppdes_14.93	400	164	3646	1012	0	1112	1489	197	5	21	13	E	-	-	cell wall protein
Ppdes_16.9	5343	1082	67	62	369	572	477	373	4	39	13	N	PF01849	NAC domain	Nascent polypeptide-associated complex subunit beta
Ppdes_20.109	64	100	0	5	168	120	144	369	0	38	13	E	-	-	-
Ppdes_3.221	2	1	0	0	40	109	192	142	4	15	13	N	PF00545	ribonuclease	guanyl-specific ribonuclease f1
Ppdes_32.80	31	77	16193	7266	2635	1902	1424	209	8	19	13	N	PF05730	CFEM domain	proline-rich antigen
Ppdes_36.69	398	575	8744	5554	750	239	622	405	8	45	13	N	PF00264	central domain of tyrosinase	tyrosinase
Ppdes_4.171	454	681	14820	6784	166	3005	3427	161	6	17	13	E	-	-	major allergen alt
Ppdes_5.53	76	155	137	137	648	165	163	188	8	17	13	N	PF05730	CFEM domain	mediator of rna polymerase ii transcription subunit
Ppdes_58.19	9	12	0	5	16	20	11	439	6	49	13	N	PF07519	Tannase and feruloyl esterase	feruloyl esterase b
Ppdes_1.160	1	2	0	2	0	15	15	393	8	40	12	N	PF00544	Pectate lyase	pectin lyase
Ppdes_1.265	27	7	0	6	35	29	8	268	3	29	12	N	PF01828	Peptidase A4 family	peptidase a4 family protein
Ppdes_1.505	1	3	0	0	26	4	2	254	2	28	12	N	-	-	predicted protein
Ppdes_11.16	214	122	1920	2652	614	134	168	260	1	29	12	N	-	-	hypothetical protein
Ppdes_12.78	0	0	120	24	33	0	0	333	28	35	12	E	-	-	hypothetical protein W97_01096
Ppdes_2.456	1	9	0	5	115	10	8	135	2	12	12	N	PF07174	Fibronectin-attachment protein	hypothetical protein SNOG_09735

Table 19 continued…

Gene ID	*In vitro* FPKM				*In planta* FPKM			Amino acid length	Cystiene count	Molecular Weight	Cumulative Rank	EffectorP[a] prediction	fectorPa p value	Pfam ID	Pfam description	Top BLAST hit description b
	F2W4	F3W4	F3-3D	F2-3D	Early	Medium	Late									
Ppdes_4.286	258	465	275	703	160	23	22	1058	10	110	12	N	1.0	-	-	glycoside hydrolase family 18
Ppdes_4.379	8	7	204	69	75	1069	492	305	4	32	12	N	1.0	PF00194	Eukaryotic-type carbonic anhydrase	carbonic anhydrase
Ppdes_43.45	4	4	0	3	0	113	109	447	10	48	12	N	1.0	PF00295	Glycosyl hydrolases family 28	glycoside hydrolase family 28
Ppdes_6.421	0	0	0	0	0	22	33	238	16	25	12	E	1.0	PF03211	Pectate lyase	pectate lyase
Ppdes_6.425	0	1	0	0	18	1	1	398	3	40	12	N	1.0	PF00082	Subtilase family	subtilisin-like protease pr1a
Ppdes_63.22	1	1	1	0	53	17	19	570	6	60	12	N	1.0	PF00135	Carboxylesterase family	para-nitrobenzyl esterase
Ppdes_7.101	0	1	0	1	19	360	289	316	4	33	12	N	1.0	PF00544	Pectate lyase	pectate lyase b
Ppdes_1.4	58	10	6	12	18	1382	1238	505	6	56	11	N	1.0	-		hypothetical protein SNOG_04543
Ppdes_1.96	26	46	50	59	424	52	120	328	10	37	11	N	1.0	PF00445	Ribonuclease T2 family	ribonuclease t2
Ppdes_2.263	4	2	0	0	0	47	72	262	0	29	11	N	1.0	-	-	carbohydrate-binding module family 1 protein
Ppdes_2.322	7	5	0	9	0	21	11	239	6	26	11	N	0.7	-	-	protein
Ppdes_20.44	5	3	4	3	1517	50	595	207	4	22	11	N	1.0	PF07510	Protein of unknown function (DUF1524)	secreted protein
Ppdes_3.450	10	41	6	5	106	25	24	756	10	83	11	N	1.0	PF00266	Aminotransferase class-V	aminotransferase family protein
Ppdes_30.91	95	80	77	76	181	16	4	274	4	30	11	N	0.7	PF13460	NADH(P)-binding	hypothetical protein W97_08467
Ppdes_33.35	9	11	0	34	0	23	28	288	3	31	11	N	0.8	PF00106	short chain dehydrogenase	polysaccharide deacetylase family

EffectorP is a machine-learning tool developed to predict effectors in fungi. a E and N indicate genes predicted as effector and non-effector by EffectorP, respectively. b indicate top blast result at NCBI (nr). FPKM = Fragments Per Kilobase of transcript per Million mapped reads

6.3.3 *In planta* expression pattern of effector like candidate genes in *P. pinodella*

Twenty-two of the effector like candidate genes showed ≥ 2 fold change when early time is compared to mid time points. Sixteen of the genes expressed only at early time points indicating differential activation of effector like genes during infection and disease development. Some of these genes include *Pdlla_10646* (Alpha/beta hydrolase family [Pfam: PF12697], *Pdlla_10746* (pectinesterase [Pfam: PF01095]), *Pdlla_124* (FAD dependent oxidoreductase [Pfam: PF01266]), *Pdlla_1302* (asparaginase [Pfam: PF00710]), *Pdlla_1430* (arginase [Pfam: PF00491]), *Pdlla_3572* (GlcNAc-PI de-N-acetylase [Pfam: PF02585]), *Pdlla_771* (thioesterase-like superfamily[Pfam: PF13279) and *Pdlla_9106* (short chain dehydrogenase[Pfam: PF00106]). Putative genes encoding asparaginase and arginase showed highest expression among all the candidate genes at earliest time points.

Likewise, 66 candidate genes were induced in response to host at mid-time point post infection. Thirty-two of these genes significantly up-regulated compared to early infection. Most genes induced at mid-stage of infection are associated to CAZymes such as GH43 [Pfam: PF04616], GH10 [Pfam: PF00331], GH61 [Pfam: PF03443], GH88 [Pfam: PF07470], chitin binding domain [Pfam: PF03067 and pectate lyase [Pfam: PF00544]. Cutinase [Pfam: PF01083], FAD binding domain [Pfam: PF01494] and protease (putative peptidase family [Pfam: PF13933], trypsin [Pfam: PF00089] and eukaryotic aspartyl protease [Pfam: PF00026] were also induced at mid-stage infection (Figure 28, Table 20 , and Supplementary Table 10).

About 18 genes were constitutively expressed from early to late stage of infection with fold change varying from -0.8 to 0.9 (excluding 0). Some of the proteins include *Pdlla_1196, Pdlla_1551, Pdlla_2835, Pdlla_3392, Pdlla_340/Pdlla_9039* and *Pdlla*_9293 which belong to FAD binding [Pfam: PF01494], NPP1 [Pfam: PF05630], GH76 [Pfam: PF03663], GH16 [Pfam: PF00722], chitin binding [Pfam: PF03067] and cutinase [Pfam: PF01083] protein families, respectively. Heat map and two way clustering of expression levels of the first top 100 genes are indicated in Figure 28.

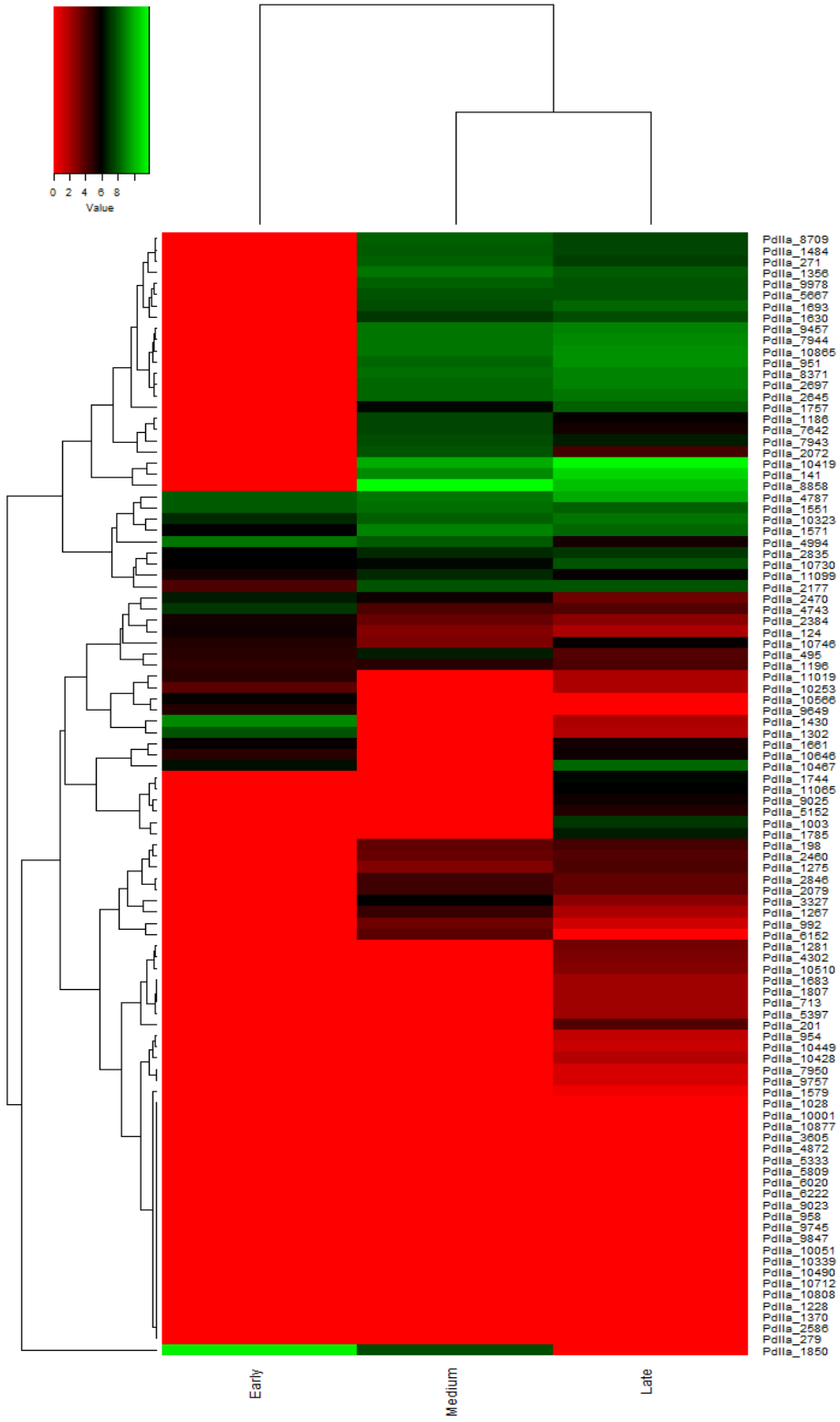


Figure 28 Heat map of relative expression levels of candidate effector-like genes in P. pinodella isolate 410/95.

Early = 0 – 9 hour post infection (HPI); Medium = 9 – 33 HPI; HPI; 48 - 96 HPI.

Table 20 Top 60 effector-like candidate genes in *P. pinodella* isolate 41095 as predicted by the evidence-based ranked score method and their associated supporting data.

Gene ID	Early	Medium	Late	Early vs Medium		Early vs Late		Medium Vs Late		Sequence length	Cysteine count	Molecular weight (kDa)	Rank	EffectorP	PFAM Description	Top BLAST hit description
				Log2(Δ)	Q value	Log2(Δ)	Q value	Log2(Δ)	Q value							
Pdlla_1356	0	402	255	+	0.00	+	0.00	-0.7	0.75	326	9	35.8	15	N	PF04616 Glycosyl hydrolases 43	hypothetical protein
Pdlla_8709	0	277	178	+	0.01	+	0.04	-0.6	0.80	257	5	26.6	15	N	PF00331 Glycosyl hydrolase 10	
Pdlla_1186	0	183	53	+	0.00	+	1.00	-1.8	0.60	450	7	47.5	13	N	PF03663 Glycosyl hydrolase 76	hypothetical protein
Pdlla_2072	0	225	20	+	0.02	+	1.00	-3.5	0.60	279	6	27.4	13	N	PF11327 Protein of unknown function (DUF3129)	hypothetical protein
Pdlla_2645	0	309	386	+	0.00	+	0.00	0.3	0.88	391	6	41.1	13	N	PF00544 Pectate lyase	-
Pdlla_7642	0	176	42	+	0.00	+	1.00	-2.1	0.60	443	3	47.2	13	N	PF01565 FAD binding domain	similar to FAD linked oxidase domain protein
Pdlla_7943	0	204	96	+	0.00	+	1.00	-1.1	0.64	389	7	43.2	13	N	PF01494 FAD binding domain	similar to salicylate hydroxylase
Pdlla_7944	0	382	517	+	0.00	+	0.00	0.4	0.84	226	6	22.2	13	N	PF01083 Cutinase	hypothetical protein
Pdlla_8858	0	3686	1271	+	0.00	+	0.00	-1.5	0.57	106	6	11.1	13	E	PF10083 Uncharacterized protein conserved in bacteria (DUF2321)	hypothetical protein
Pdlla_9457	0	400	511	+	0.00	+	0.00	0.4	0.87	305	6	31.3	13	N	PF03067 Chitin binding domain	hypothetical protein
Pdlla_951	0	308	598	+	0.00	+	0.00	1.0	0.63	305	6	31.2	13	N	PF03443 Glycosyl hydrolase 61	similar to endoglucanase II
Pdlla_9978	0	269	218	+	0.00	+	0.00	-0.3	0.90	327	9	34.2	13	N	PF13933 Putative peptidase family	hypothetical protein
Pdlla_10001	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	254	12	27.1	12	E	-	-
Pdlla_1028	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	197	14	21.6	12	E	-	-
Pdlla_10323	111	283	372	1.4	0.63	1.7	0.60	0.4	0.86	316	17	32.7	12	N	-	hypothetical protein
Pdlla_10428	0	0	3	0.0	1.00	+	1.00	+	1.00	346	9	38.3	12	N	-	-
Pdlla_10510	0	0	7	0.0	1.00	+	1.00	+	1.00	113	7	11.6	12	E	-	-

Table 20 continued …

Gene ID	Early	Medium	Late	Early vs Medium		Early vs Late		Medium vs Late		Cysteine count	Cysteine count	Molecular weight (kDa)	dN/dS	Rank	EffectorP	Pfam ID	PFAM Description	Top BLAST hit description
				Log2(Δ)	Q value	Log2(Δ)	Q value	Log2(Δ)	Q value									
Pdlla_10865	0	410	626	+	0.00	+	0.00	0.6	0.70	533	14	57.8	0.8	12	N	PF00172	Fungal Zn(2)-Cys(6) binuclear cluster domain	hypothetical protein
Pdlla_10877	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	253	6	27.2	0.0	12	E	-	-	-
Pdlla_11019	29	0	3	-	1.00	-3.3	1.00	+	1.00	448	6	48.3	0.0	12	N	-	-	hypothetical protein
Pdlla_1267	0	23	3	+	1.00	+	1.00	-2.9	1.00	289	9	32.2	0.0	12	N	-	-	-
Pdlla_141	0	565	1714	+	0.03	+	0.00	1.6	0.60	157	5	16.8		12	E	-	-	predicted protein
Pdlla_198	0	12	19	+	1.00	+	1.00	0.6	1.00	288	21	30.2	0.4	12	E	-	-	-
Pdlla_2470	97	48	9	-1.0	1.00	-3.4	1.00	-2.3	1.00	176	6	19.2	0.0	12	E	-	-	hypothetical protein
Pdlla_2697	0	304	460	+	0.00	+	0.00	0.6	0.77	374	0	38.4	0.5	12	N	-	-	hypothetical protein
Pdlla_3327	0	59	6	+	1.00	+	1.00	-3.3	1.00	313	7	34.9	4.0	12	N	-	-	hypothetical protein
Pdlla_3605	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	270	9	29.2	0.0	12	N	-	-	hypothetical protein
Pdlla_4302	0	0	8	0.0	1.00	+	1.00	+	1.00	370	8	42.2	0.0	12	N	-	-	predicted protein
Pdlla_4743	149	17	15	-3.2	1.00	-3.3	1.00	-0.1	1.00	231	3	25.4	0.0	12	N	-	-	hypothetical protein
Pdlla_4787	257	395	984	0.6	0.83	1.9	0.57	1.3	0.60	197	5	20.8	0.0	12	N	-	-	hypothetical protein
Pdlla_4872	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	144	12	14.9	2.0	12	E	-	-	-
Pdlla_495	31	93	15	1.6	1.00	-1.0	1.00	-2.6	1.00	381	10	42.3	0.0	12	N	-	-	hypothetical protein
Pdlla_4994	394	264	44	-0.6	0.79	-3.2	0.60	-2.6	0.60	275	16	26.5	0.0	12	N	-	-	-
Pdlla_5152	0	0	35	0.0	1.00	+	1.00	+	1.00	168	6	17.9	3.0	12	E	-	-	-
Pdlla_5333	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	431	3	47.6	1.5	12	N	-	-	hypothetical protein
Pdlla_5397	0	0	4	0.0	1.00	+	1.00	+	1.00	300	3	31.2	0.5	12	N	-	-	predicted protein
Pdlla_5667	0	229	233	+	0.00	+	0.01	0.0	0.97	301	0	29.6	0.0	12	N	-	-	-
Pdlla_5809	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	213	16	23.0	0.0	12	E	-	-	predicted protein
Pdlla_6020	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	333	28	34.9	0.5	12	E	-	-	-
Pdlla_6152	0	13	0	+	1.00	0.0	1.00	-	1.00	258	6	27.6	0.0	12	E	-	-	hypothetical protein
Pdlla_6222	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	160	8	17.3	0.0	12	N	-	-	-

Table 20 continued …

Gene ID	Early	Medium	Late	Early vs Medium		Early vs Late		Medium vs Late		Cysteine count	Cysteine count	Molecular weight (kDa)	dN/dS	Rank	EffectorP[a]	Pfam ID	PFAM Description	Top BLAST hit description
				Log2(Δ)	Q value	Log2(Δ)	Q value	Log2(Δ)	Q value									
Pdlla_10865	0	410	626	+	0.00	+	0.00	0.6	0.70	533	14	57.8	0.8	12	N	PF00172	Fungal Zn(2)-Cys(6) binuclear cluster domain	hypothetical protein
Pdlla_713	0	0	4	0.0	1.00	+	1.00	+	1.00	159	5	16.7	0.0	12	N	-	-	hypothetical protein
Pdlla_7950	0	0	1	0.0	1.00	+	1.00	+	1.00	384	7	41.8	1.3	12	N	-	-	
Pdlla_8371	0	347	463	+	0.01	+	0.00	0.4	0.86	212	5	21.3		12	N	-	-	hypothetical protein
Pdlla_9023	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	192	7	21.5	3.0	12	E	-	-	
Pdlla_9025	0	0	46	0.0	1.00	+	1.00	+	1.00	157	5	15.2	0.0	12	E	-	-	predicted protein
Pdlla_954	0	0	2	0.0	1.00	+	1.00	+	1.00	275	8	30.2	3.0	12	N	-	-	hypothetical protein
Pdlla_958	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	419	3	41.2	0.7	12	N	-	-	
Pdlla_9649	34	0	0	-	1.00	-	1.00	0.0	1.00	364	3	40.0	0.0	12	N	-	-	hypothetical protein
Pdlla_9745	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	207	6	23.2	1.3	12	E	-	-	
Pdlla_9757	0	0	1	0.0	1.00	+	1.00	+	1.00	427	8	40.5	0.0	12	N	-	-	
Pdlla_9847	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	138	6	15.5	0.0	12	E	-	-	hypothetical protein
Pdlla_992	0	10	1	+	1.00	+	1.00	-2.9	1.00	321	6	35.0	0.0	12	N	-	-	predicted protein
Pdlla_1003	0	0	141	0.0	1.00	+	1.00	+	1.00	224	5	23.4	0.5	11	N	PF03443	Glycosyl hydrolase 61	hypothetical protein
Pdlla_10051	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	202	19	21.6	2.5	11	E	PF00879	Defensin propeptide	hypothetical protein
Pdlla_10253	14	0	3	-	1.00	-2.4	1.00	+	1.00	426	16	46.7	0.8	11	N	PF09362	Domain of unknown function (DUF1996)	hypothetical protein
Pdlla_10339	0	0	0	0.0	1.00	0.0	1.00	0.0	1.00	384	9	38.8	0.3	11	N	PF00295	Glycosyl hydrolases 28	endopolygalacturonase
Pdlla_10419	0	933	3040	+	0.00	+	0.00	1.7	0.37	373	6	37.7		11	N	PF00089	Trypsin	hypothetical protein
Pdlla_10449	0	0	2	0.0	1.00	+	1.00	+	1.00	665	12	73.9	2.7	11	N	-	-	hypothetical protein

a E and N indicate genes predicted as effector and non-effector by EffectorP, respectively. b indicate top blast result at NCBI (nr) database

6.3.4 *In planta* expression pattern of effector like candidate genes in *Ph. koolunga*

There were no significant differences observed among the three-time point considered in levels of gene expressed during *Ph. koolunga* infection and disease development. However, about 34 genes showed high expression of greater than 1.5 fold changes compared to medium time post infection (Figure 29, Table 21, and Supplementary Table 10). Among 34 genes, only one gene (*Phkol_189*) is associated to protein of known domain, tannase and feruloyl esterase [Pfam: PF07519]. Not all the rest contain functionally characterized protein domain. On the other hand, 20 genes showed greater than 2 fold change in medium stage of infection relative to early time point. Four of these genes viz *Phkol_1305*, *Phkol_2036*, *Phkol_2123*, and *Phkol_2127* contain CFEM domain [PF05730], NADP oxidoreductase coenzyme F420-dependen [PF03807], and Alpha/beta hydrolase family [PF12697] of protein domains, respectively.

Twenty-six of the effector like candidate genes showed consistent expression pattern with log2 fold change values between -1 and +1 when early and medium time points were compared. Only seven of the genes *Phkol_2412*, *Phkol_9859*, *Phkol_1368*, *Phkol_3432*, *Phkol_4205*, *Phkol_8300* and *Phkol_9447* contain GH18 [Pfam: PF00704], pectinesterase [Pfam: PF01095], aldose 1-epimerase [Pfam: PF01263], germination protease [Pfam: PF03418], NAD(P)-binding rossmann-like domain [Pfam: PF13450], CFEM [Pfam: PF05730], and peroxisomal membrane anchor protein [Pfam: PF04695], respectively (Figure 20, Table 21, Supplementary Table 10). Some of the predicted candidate genes were not expressed during *in planta* infection and disease development. This may not rule out their contribution to infection and disease development, but may not be induced under current experimental condition.

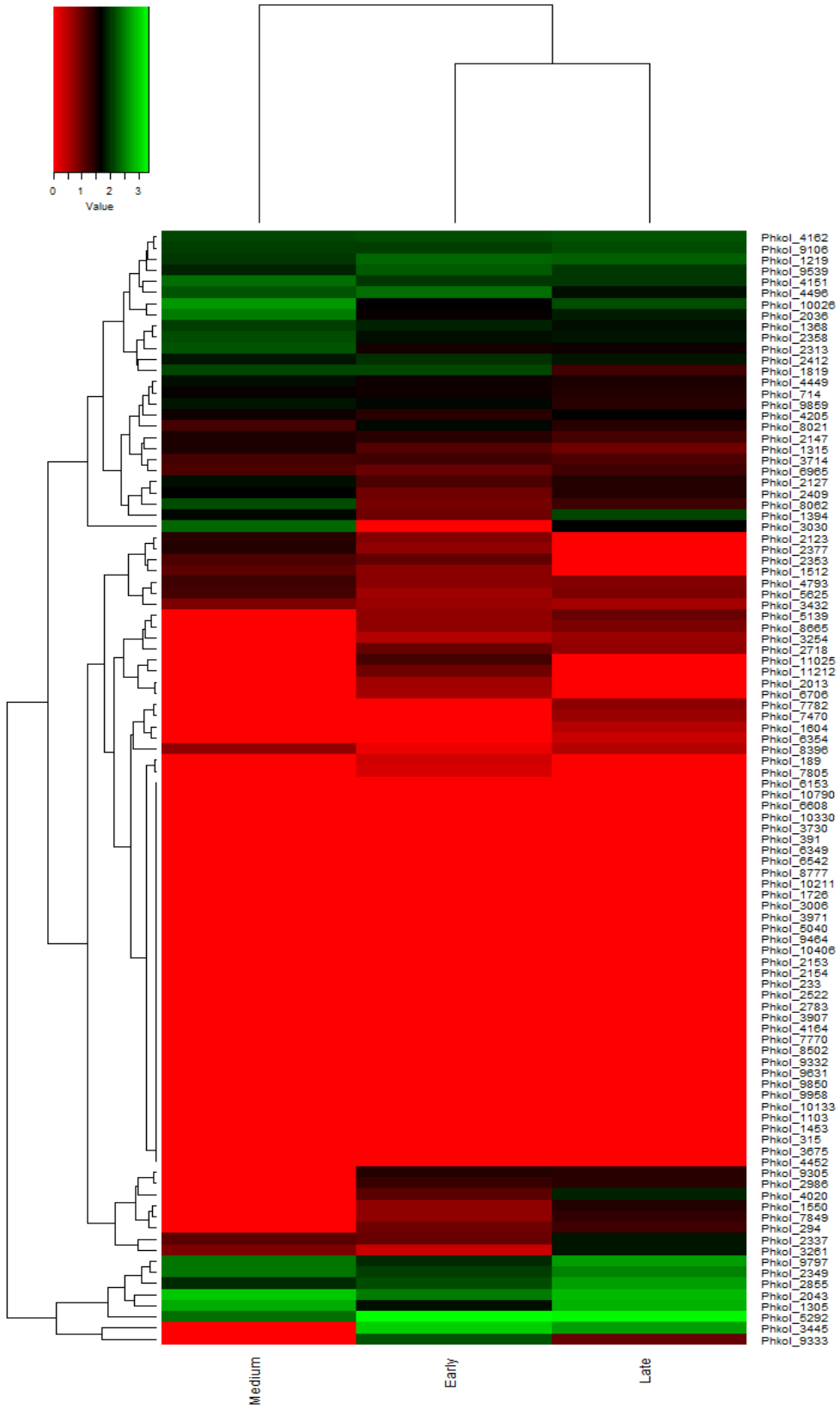


Figure 29 Heat map of relative expression levels of candidate effector-like genes in *Ph. koolunga* isolate FT04040.

Early = 0 – 9 hour post infection (HPI); Medium = 9 – 33 HPI; HPI; 48 - 96 HPI.

Table 21 Top 60 effector-like candidate genes in *Ph koolunga* isolate FT04040 as predicted by the evidence-based ranked score method and their associated supporting data.

Gene ID	Early	Medium	Late	Early Vs Medium		Early vs Late		Medium vs Late		Sequence length	Cystiene count	MWkDA	Rank	EffectorP prediction[a]	EffectorP *P value*	Pfam ID	Pfam Description	Top BLAST hit
				Log2(Δ)	Q value	Log2(Δ)	Q value	Log2(Δ)	Q value									
Phkol_10026	45.8	430.5	140.8	3.2	1.0	1.6	1.0	-1.6	1.0	176	10	18.83	11	E	1.0	-	-	hypothetical protein 9
Phkol_10133	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	60	2	6.7	11	N	0.9	-	-	
Phkol_10211	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	75	0	8.39	12	N	1.0	-	-	hypothetical protein
Phkol_10330	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	69	0	7.4	12	E	0.7	-	-	
Phkol_10406	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	86	2	9.51	11	E	1.0	-	-	
Phkol_10790	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	161	9	17.17	13	E	1.0	-	-	telomere silencing protein
Phkol_11025	16.6	0.0	0.0	-	1.0	-	1.0	0.0	1.0	69	0	7.36	12	N	1.0	-	-	
Phkol_1103	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	321	10	35.51	11	N	0.9	-	-	hypothetical protein
Phkol_11212	7.6	0.0	0.0	-	1.0	-	1.0	0.0	1.0	83	0	8.57	11	N	0.9	-	-	
Phkol_1219	208.1	99.9	185.4	-1.1	1.0	-0.2	1.0	0.9	1.0	237	1	25.55	12	N	0.5	-	-	plasma membrane fusion protein prm1
Phkol_1305	54.1	612.2	665.5	3.5	1.0	3.6	1.0	0.1	1.0	198	10	18.03	11	N	1.0	PF05730	CFEM domain	cfem domain
Phkol_1315	12.4	29.5	7.8	1.3	1.0	-0.7	1.0	-1.9	1.0	202	5	21.37	12	N	1.0			
Phkol_1368	72.9	118.9	59.1	0.7	1.0	-0.3	1.0	-1.0	1.0	396	5	43.37	11	N	1.0	PF01263	Aldose 1-epimerase	aldose 1-epimerase
Phkol_1394	7.9	49.0	130.2	2.6	1.0	4.0	1.0	1.4	1.0	233	6	23.62	11	N	0.9	-	-	hypothetical protein
Phkol_1453	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	312	9	34.2	11	N	1.0	-	-	hypothetical protein
Phkol_1512	4.7	11.1	0.0	1.2	1.0	-	1.0	-	1.0	279	6	28.31	11	N	1.0	-	-	-
Phkol_1550	4.6	0.0	27.9	-	1.0	2.6	1.0	+	1.0	177	7	18.82	11	E	1.0	-	-	hypothetical protein

Table 21 continued …

Gene ID	Early	Medium	Late	Early Vs Medium		Early Vs Late		Medium vs Late		Sequence length	Cystiene count	MWkDA	Rank	EffectorP Prediction	EffectorP P value	Pfam ID	Pfam description	Top BLAST hit
				Log2(Δ)	Q value	Log2(Δ)	Q value	Log2(Δ)	Q value									
Phkol_1604	0.0	0.0	2.3	0.0	1.0	+	1.0	+	1.0	300	9	30.99	12	N	1.0	-	-	-
Phkol_1726	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	107	3	12.29	12	N	1.0	PF08432	AAA-ATPase Vps4-associated protein	
Phkol_1819	125.5	127.3	16.9	0.0	1.0	-2.9	1.0	-2.9	1.0	121	6	12.57	11	E	1.0	-	-	hypothetical protein
Phkol_189	1.1	0.0	0.0	-	1.0	-	1.0	0.0	1.0	293	9	31.46	11	N	1.0	PF07519	Tannase and feruloyl esterase	feruloyl esterase b precursor protein
Phkol_2013	3.4	0.0	0.0	-	1	-	1	0	1	237	5	24.92	14	N	1.0		-	-
Phkol_2036	42.2	294.0	69.7	2.8	1.0	0.7	1.0	-2.1	1.0	119	2	12.55	11	E	1.0	PF03807	NADP oxidoreductase coenzyme	pyrroline-5-carboxylate reductase
Phkol_2043	286.7	932.5	740.1	1.7	1.0	1.4	1.0	-0.3	1.0	206	5	22.61	11	E	0.9	-	-	hypothetical protein
Phkol_2123	5.6	25.1	0.0	2.2	1.0	-	1.0	-	1.0	416	6	45.36	11	N	1.0	PF12697	Alpha/beta hydrolase family	abhydrolase domain-containing protein 12b
Phkol_2127	13.4	57.7	26.4	2.1	1.0	1.0	1.0	-1.1	1.0	360	7	40.34	11	N	1.0	PF00856	SET domain	set domain-containing protein 5
Phkol_2147	24.1	30.6	16.6	0.3	1.0	-0.5	1.0	-0.9	1.0	438	6	49.46	11	N	1.0	-	-	glycosyltransferase 25
Phkol_2153	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	89	10	9.82	11	E	1.0	-	-	-
Phkol_2154	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	94	10	10.3	11	E	1.0	-	-	-
Phkol_2313	33.0	162.5	36.1	2.3	1.0	0.1	1.0	-2.2	1.0	213	0	23.13	11	N	1.0	PF10177	Uncharacterised (DUF2371)	predicted protein

Table 21 continued …

Gene ID	Early	Medium	Late	Early Vs Medium		Early vs Late		Medium vs Late		Sequence length	Cysteine count	Molecular weight (kDa)	Rank	EffectorP Prediction	EffectorP *P value*	Pfam ID	Pfam description	Top BLAST hit
				Log2(Δ)	Q value	Log2(Δ)	Q value	Log2(Δ)	Q value									
Phkol_233	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	196	8	21.12	11	E	1.0	-	-	hypersensitive response inducing protein 1
Phkol_2337	9.0	11.7	59.7	0.4	1.0	2.7	1.0	2.4	1.0	433	16	46.45	11	N	1.0	-	-	gpi transamidase component pig-
Phkol_2349	114.7	271.7	328.0	1.2	1.0	1.5	1.0	0.3	1.0	380	6	39.78	11	N	1.0	PF00544	Pectate lyase	pectin lyase a precursor
Phkol_2353	9.8	14.6	0.0	0.6	1.0	-	1.0	-	1.0	208	5	22.26	11	N	0.9	-	-	related to heatshock protein hsp150
Phkol_2358	57.0	141.0	61.6	1.3	1.0	0.1	1.0	-1.2	1.0	203	5	20.36	11	N	1.0	-	-	-
Phkol_2377	4.5	26.5	0.0	2.6	1.0	-	1.0	-	1.0	219	6	24.37	11	E	0.8	-	-	hypothetical protein
Phkol_2409	8.0	46.9	26.0	2.5	1.0	1.7	1.0	-0.8	1.0	253	7	28.05	11	N	0.7	-	-	hypothetical protein SETTUDRAFT_44298
Phkol_2412	95.8	60.0	61.6	-0.7	1.0	-0.6	1.0	0.0	1.0	1792	11	192.44	12	N	1.0	PF00704	Glycosyl hydrolases 18	glycoside hydrolase 18
Phkol_2522	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	120	11	12.64	11	E	1.0	-	-	-
Phkol_2718	8.9	0.0	4.4	-	1.0	-1.0	1.0	+	1.0	315	7	34.65	11	N	1.0	-	-	tat pathway signal sequence
Phkol_2783	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	200	9	21.72	11	E	0.7	-	-	hypothetical protein
Phkol_2855	143.2	88.1	478.1	-0.7	1.0	1.7	1.0	2.4	1.0	113	8	11.37	11	E	0.5	-	-	endo- -beta-xylanase d precursor
Phkol_294	7.7	0.0	17.4	-	1.0	1.2	1.0	+	1.0	119	6	12.22	11	N	1.0	-	-	hypothetical protein
Phkol_2986	20.4	0.0	23.0	-	1.0	0.2	1.0	+	1.0	355	6	37.9	11	N	1.0	-	-	hypothetical protein
Phkol_3006	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	170	1	21.51	12	N	1.0	-	-	-
Phkol_3030	0.0	218.7	46.2	+	1.0	+	1.0	-2.2	1.0	76	0	7.2	12	E	1.0	-	-	-
Phkol_315	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	307	7	32.6	11	N	1.0	-	-	endo- -beta-

Table 21 continued …

Gene ID	Early	Medium	Late	Early Vs Medium		Early vs Late		Medium vs Late		Sequence length	Cystiene count	MWkDA	Rank	EffectorP prediction	EffectorP p value	Pfam ID	Pfam description	Top BLAST hit
				Log2(Δ)	Q value	Log2(Δ)	Q value	Log2(Δ)	Q value									
Phkol_233	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	196	8	21.12	11	E	1.0	-	-	hypersensitive response inducing protein 1
Phkol_3254	2.4	0.0	4.0	-	1.0	0.8	1.0	+	1.0	344	7	39.34	11	N	1.0	-	-	hypothetical protein
Phkol_3261	1.5	6.7	62.0	2.2	1.0	5.4	1.0	3.2	1.0	407	7	44.32	11	N	1.0	-	-	c2h2 transcription
Phkol_3432	4.1	6.2	2.8	0.6	1.0	-0.5	1.0	-1.1	1.0	462	2	47.65	11	N	1.0	PF03418	Germination protease	
Phkol_3445	1023.0	0.0	507.6	-	0.6	-1.0	1.0	+	0.9	185	13	19.55	12	E	1.0			cell wall glycoprotein
Phkol_3675	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	276	0	29.06	11	N	1.0	PF11972	HTH DNA binding domain	
Phkol_3714	16.6	15.6	13.8	-0.1	1.0	-0.3	1.0	-0.2	1.0	473	5	49.9	14	N	1.0	-	-	ribosomal protein s17 protein
Phkol_3730	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	135	7	14.85	12	E	1.0	-	-	
Phkol_3907	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	191	7	20.33	11	E	0.9	-	-	hypothetical protein
Phkol_391	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	81	3	8.82	12	E	0.8	PF07632	Unknown domain (DUF1593)	cellulose-binding protein
Phkol_3971	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	164	5	18.8	12	N	1.0	-	-	
Phkol_4020	10.9	0.0	77.9	-	1.0	2.8	1.0	+	1.0	366	6	40.13	11	N	1.0	-	-	glycosyltransferase 25 protein

Results of effector prediction across the three fungal pathogens of field pea indicate that parts of effector like genes activated during infection and disease development were CAZymes directed towards depolymerzation of plant cell wall integrity. According to Pfam domain analysis, GH3, GH5, GH18, GH28 and GH61 in *P. pinodes*, GH43, GH10, GH76 and GH61 in *P. pinodella*, and, GH18 in *Ph. koolunga* were predicted as putative effector like proteins. The GH5 (cellulase) and GH61 act on plant cellulose while GH43 and GH10 are active against hemicellulose degradation. The GH88 (d-4, 5-unsaturated β-glucuronyl hydrolase) are involved in the breakdown of the products from pectate lyases. Proteins containing tannase and feruloyl esterase in *Aspergillus* are involved in degradation of pectin and xylan [64]. Gene encoding alpha/beta hydrolase family observed in all the three pathogens function as hydrolases, lyases, transferases, hormone precursors or transporters, chaperones or routers of other proteins [187, 188].

On the other hand, GH18 is directed towards fungal cell wall degradation mainly chitin. It seems that sequestering of fungal fragments is important to lessen fungal recognition by host. Perception of fungal chitin by plant leads to activation of host defense signaling pathway [152, 328]. Alternatively, *P. pinodes*, *P. pinodella* and *Ph. koolunga* may also produce chitinases to overcome competition for niche and resources against each other and possibly other pathogens adapted to same host or against saprophytic organism. Davidson *et al.* [58] observed existence of antagonism among *P. pinodes*, *P. pinodella* and *Ph. koolunga* using paired co-inoculation techniques. Further experimental evidence would be required to determine mechanism of antagonism. Langner and Gohre [178] reviewed fungal chitinase and suggested the possibilities that it could contribute to overall pathogen fitness in intricate conditions.

The number of predicted effector like PLs were high in *P. pinodes* (5) while only one PL families predicted as putative effector candidate in *P. pinodella* and *Ph. koolunga* , all of which were expressed during infection. Pectate lyases are secreted proteins involved in maceration and soft rotting of plant tissue [344]. Pectinesterase is also involved in alteration of plant cell wall integrity [89]. This suggest that *P. pinodes*, *P. pinodella* and *Ph. koolunga* could use PLs families that portray different mechanism

of pectin digestion. Rapid disease development observed in *P. pinodes* compared to *Ph. koolunga* and *P. pinodella* could be associated to the higher number of secreted PLs involved in *P. pinodes* (dose dependent) instead of functional redundancy.

6.3.5 Function of some predicted candidate effector–like genes are similar in three pathogens

Interestingly, putative cytochrome P450 pisatin demethylase (*Ppdes_5.300*: GO: 0005506*)* and cytochrome P450 alkane (*Ppdes_5.48*; GO: 0005506) were also predicted as putative effector like candidate *P. pinodes*. Prediction of the P450 pisatin demethylase and P450 alkane as putative effector candidate are imperative in ascochyta blight pathogens of field pea (ABPFP) as both likely related to pathogenicity. Previous work has shown that cytochrome P450 monooxygenase is involved in pisatin detoxification process in *N. haematococca* and suggested to play similar role in *P. pinodes*, *P. pinodella* and *A. pisi* [71, 105]. Nevertheless, cytochrome P450 genes are numerous, highly divergent in filamentous ascomycetes [72] and have broad roles across various cellular activities. On the other hand, fungal P450 alkane belong to CYP52 are useful for utilization of hydrocarbons as nutrition [194]. In addition to satisfying nutritional demand, P450 alkane might have also contribute to fungal pathogenicity by facilitating attachment and penetration of the hydrophobic pea leaf surface that contain high levels of alkanes.

Likewise, necrosis inducing protein (NPP1, Pfam: PF05630) also known as necrosis- and ethylene - inducing peptide – like protein (NLPs) was predicted as potential effector-like candidates in *P. pinodella Pdlla_1551*. Although not predicted as an effector, homologs of the protein containing NLP exist in all the three pathogens (*P. pinodes*, *P. pinodella*, *Ph. koolunga*). The NLPs contribute to disease development *via* induction of cell cytotoxicity in a number of fungal pathogens including *C. coccodes*, *A. theophrasti*, and *R. commune* [9, 83, 165]. As observed in other pathosystem and discussed in Chapter 3, NLPs were conserved in ABPFP and demonstrated to induce

necrosis when purified NLP from *P. pinodes* was infiltrated on intact pea and other dicot plants.

Cyanovirin-N (CVNH) [Pfam: PF08881] and LysM [Pfam: PF01476] domain (*Ppdes_95.2*) was predicted as putative candidate effector-like protein in *P. pinodes.* As opposed to NLP, secreted fungal LysM domain-containing proteins protect the invasive fungal hyphae and prevent chitin-triggered immunity to enhance establishment of infection and subsequent disease development. The genes *Ecp6 Cladosporium fulvum* (de Jonge et al, 2010), *Mg3LysM* from *M. graminicola* (Marshall et al. 2011*)* and *Slp1* from *M. oryzae (Mentlak et al. 2012)* encode similar chitin-binding effector molecules that enhance virulence by protecting the fungal cell exterior from host-defences.

The early expression preference of the pathogenesis related genes such as cutinase may suggest its contribution to infection and disease development. In *M. grisea*, a cutinase (*CUT2*) is required for full virulence [289]. Similarly, cutinases have been shown to contribute to pathogenicity of *Pyrenopeziza brassicae* on *Brassica napus* [59].

Protein sequences represented by PF05730 in the Pfam database were reported as cysteine rich fungal-specific extracellular membrane proteins (CFEM) [173] [353]. In *M. grisea*, CFEM contribute to diseases development through activation of appressorium differentiation in response to host surfaces [73]. Similarly, one of the CFEM genes (*CFEM1)* identified from *F. graminearum* showed expression *in planta* during wheat infection and is compulsory for full virulence on wheat coleoptiles [352]. Microscopic observation indicated that ABPFP can penetrate their host either through natural opening or penetrate the cuticle directly through appressorium formation (Supporting Figure 4). Therefore, the *in planta* expression of putative CFEM in three of the ABPFP may reflect its importance during infection and disease development.

A number of genes involved in protein metabolism were also predicted as putative effector candidate genes in ABPFP. These include subtilase, peptidase, aspartyl protease, amidase and trypsin in *P. pinodes*, peptidase, asparaginase, arginase, peroxidase and tripsin were identified in *P. pinodella* while only germination protease (aspartyl protease) was predicted in *Ph. koolunga.*

Subtilisin-like secreted proteases have been shown to increase the virulence of the entomopathogenic fungus *Beauveria bassiana* [80]. Similarly, a serine protease from *P. infestans* is one of the cytoplasmic effectors which inhibit tomato-pathogenesis related proteins (PR) during infection [312]. On the other hand, a gene encoding subtilase (*prt1)* in *F. oxysporum* f.sp. *lycopersici* has no impact on its virulence [244].

A gene related to aspartyl protease is commonly expressed in three of the ABPFP. It was also expressed *in planta* in necrotrophic fungal pathogens *B. cinerea* and *S. sclerotiorum* while its role in disease development remains to be resolved [247, 310]. However, aspartyl protease reported to contribute to infection process by degrading host surface molecules in human pathogen, *Candida albicans* [281]. Recent investigation indicated that *F. oxysporum* f.sp *lycopersici* metalloprotease and serine protease can break chitin binding tomato chitinase [143].

Asparaginase is secreted protein reported from various organisms and used in blood cancer chemotherapy to treat lymphoblastic leukemia as well as food processing industries [18]. According to recent review by Batool *et al.* [18], only 9 (5 of them from *Aspergillus*) fungal species were reported to produce asparaginase while there were a number of bacterial species reported to produce it. The contribution of asparaginase to disease development is not known, test for production potential of asparaginase by ABPFP would be useful as alternative source of asparaginase production and commercialisation. Since its production was suggested to be regulated by nitrogen [273], possibilities would exist for mass production using pea plants as substrates for chemotherapy.

6.3.6 Some effector like candidate genes are under positive selection

Large numbers of predicted effector like candidate genes in ABPFP were observed to be under purifying selection (*dN/dS* <1), however only a few genes showed signs of positive selection (Figure 29). In *P. pinodes*, only two genes (*Ppdes_6.215* and *Ppdes_26.38)* showed *dN/dS* values of greater than ≥ 2. The former belonged to the PAN domain protein family [Pfam: PF14295] while the latter had no predicted protein domains. In contrast to *P. pinodes*, 24 of predicted effector like candidate genes showed sign of positive selection (*dN/dS* ≥ 2) in *P. pinodella* (Figure 30). Some of these genes include pathogenicity related genes such as cutinase (*Pdlla*_7944, *Pdlla*_9293; Pfam: PF01083), cellulase (GH5) (*Pdlla*_572; Pfam: PF00150) and FAD binding domain (*Pdlla*_986; Pfam: PF01494) while all the rest were under purifying selection (conserved). Although *Ph. koolunga* secreted proteome consisted of 1518 protein sequences, none of the predicted effector-like candidate genes exhibited positive selection (*dN/dS* >1). These results suggest that most of the effector-like protein sequences are evolving under purifying selection pressure in *P. pinodes* and *Ph. koolunga* than *P. pinodella*.

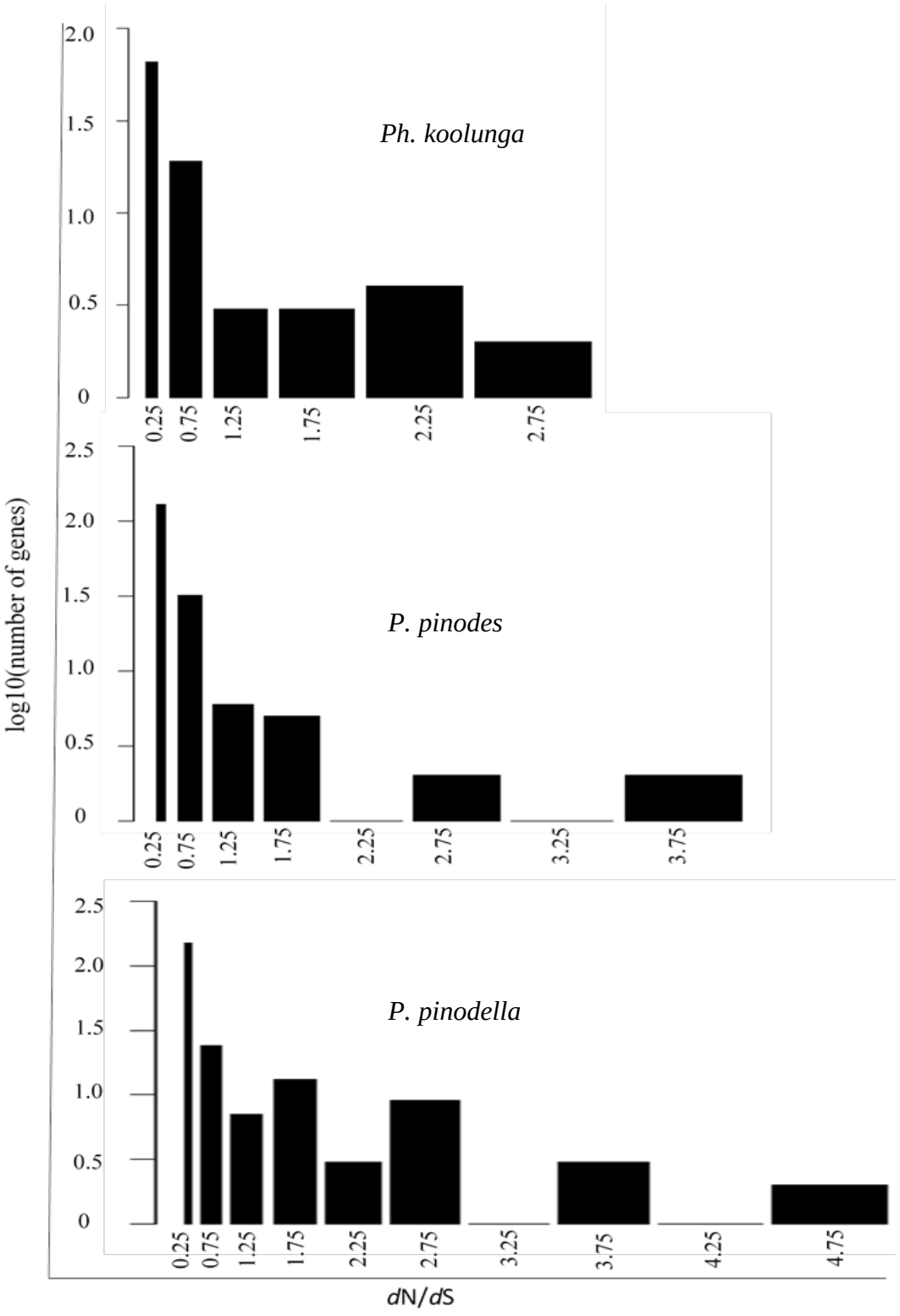


Figure 30 Distribution of dN/dS values of the secretome of *P. pinodes, P. pinodella* and *Ph. koolunga.*

Analysis of BLASTP reciprocal best BLAST hits (RBH) of the proteins sequences from *P. pinodes, P. pinodella* and *Ph. koolunga* against 2296 protein sequences obtained from the Pathogen Host Interaction Database [334] (PHI-base version 3.4 proteins) were locally performed. The result indicated that only 2 genes from *P. pinodes*, 2 genes from *Ph. koolunga* and 7 genes from *P. pinodella* which were predicted as effector-like candidate genes had RBH to functionally characterized proteins in PHI-base database (Table 22). *Peyronellaea pinodes* gene *Ppdes_6.42*1 and *P. pinodella* gene *Pdlla*_352 shared highest similarity with their RBH, which was 74% to the *N. haematococca gene PELD* (PHI-base: PHI: 180) and 73% to *M. oryzae* endo-1, 4-beta-xylanase (PHI-base: PHI: 2204), respectively. Previous experimental gene disruption of *N. haematococca PelD* revealed its importance as a virulence factor on pea [263]. Knock-down analysis of multiple genes coding for xylanase using RNAi severely reduced pathogenicity of *M. oryzae* on wheat [226]. The *P. pinodella Pdlla_2072* gene has also 70% similarity to the *M. oryzae GAS1* gene, a secreted protein belong to the mitogen-activated kinase group. Deletion of *GAS1* from *M. oryzae* severely compromised its virulence. Except *Pdlla_2072*, all were conserved across the three species suggesting a common origin for these effector-like genes in closely related pathogens adapted to the same or similar hosts.

Table 22 Predicted effector candidate genes with reciprocal blast hit to pathogen host interaction database (PHI-base)

Gene Id	PHI accession	PHI-base Gene name	PHI-base Species	PHI-base KO Phenotype	Identity (%)	Similarity	E-value	
Ppdes_6.421	PHI:180	*PELD*	*N. haematococca*	Reduced virulence	62	74	2e-96	++
Ppdes_1.96	PHI:811	*MGG_10510*	*M. oryzae*	Reduced virulence	55.8	68	1e-115	++
Pdlla_4025	PHI:201	*AVR-Pita*	*M. oryzae*	Plant avirulence determinant	32.45	50	8E-21	++
Pdlla_352	PHI:2204	endo-1,4-beta-xylanase	*M. oryzae*	Reduced virulence	64.74	73	1E-142	++
Pdlla_10419	PHI:2384	*Ust1*	*Ustilago maydis*	Reduced virulence	41.79	52	4.4E-2	++
Pdlla_4447	PHI:2412	*lipA*	*Xanthomonas campestris*	Reduced virulence	27.85	43	2E-04	++
Pdlla_2072	PHI:256	*GAS1*	*M. oryzae*	Reduced virulence	55.72	70	9E-73	+-
Pdlla_1744	PHI:653	*GIP2*	*Phytophthora sojae*	Plant avirulence determinant	27.27	41	9E-15	++
Phkol_3132	PHI:320	*SSN6*	*Candida albicans*	Reduced virulence	57.6	69	6E-138	++
Phkol_8405	PHI:364	*MIT1*	*Candida albicans*	Reduced virulence	55	67	4E-90	++

++ Orthologous in three of the species, +- no orthologues in *Ph. koolunga*

6.4 Conclusion

The diverse virulence related genes of three necrotrophic fungal pathogen *P. pinodes*, *P. pinodella* and *Ph. koolunga* have been systematically explored. Both computational and *in planta* expression analysis were employed to identify and shortlist key pathogenicity-related candidate genes expressed during infection and disease development. Subsequently, 1681, 1451, and 1518 genes were predicted to be secreted in *P. pinodes, P. pinodella* and *Ph. koolunga,* respectively. Among these 223 genes in *P. pinodes,* 212 in *P. pinodella* and 140 genes in *Ph. koolunga* predicted as putative effector like candidate proteins (CELPs). All the information can be found at Curtin University, Centre of Crop and Disease Management (CCDM), Pulse pathology and genetics research group.

Peyronellaea pinodes showed the highest number of proteins in its predicted secretome as well as CELPs lacking functional annotations, relative to *P. pinodella* and *Ph. koolunga.* Species-specific CELPs are higher in *Ph. koolunga* than both *Peyronellaea* species. Bioinformatics analysis of functional annotations indicated the potential involvement of numerous cell wall degrading enzymes and secreted proteins involved in the interaction between ascochyta blight pathogens of field pea (ABPFP) and *Pisum sativum*. The discovery of functionally active necrosis and ethylene inducing peptide like proteins (NLPs) that cause necrosis on pea and other dicot plants (see chapter 3) can be used as a positive control in further fishing effector discovery experiments for ABPFP.

To further improve the ABPFP pathogenomic resources moving forward, further functional analysis using gene knockout or *via* gene silencing among CELPs will be invaluable to determine their role during infection and disease development. Targeting single-copy candidate genes could be invaluable with this regard to eliminate the potential of functional redundancy in genes that could exist in multiple copy. Further validation of the CELPs help in the development of novel screening techniques for disease resistance as well as development of targeted breeding techniques. Knowledge

of effectors can be useful to exploit disease susceptibility gene in host given these effectors trigger diseases susceptibility [103, 196, 324].

Recently, an International Consortium for the Pea Genome Sequencing (ICPG) has been established with the aim to of providing quality reference pea genome to the broader community. Progress of this project indicated that first draft assembly of the genome has been completed (https://pag.confex.com/pag/xxiii/webprogram/Paper14072.html). As Curtin University is one of the collaborator of ICPG, we have provided the generated transcriptomic data to the consortium. These *in planta* transcriptomics data are significant contribution to ICPG, which help in correcting gene calling when included in evidence-based annotation. Further analysis of the generated *in planta* transcriptomic data will help us to pea defence responses and map metabolic pathways involved in pea defence reactions in response to infection by *P. pinodes*, *P. pinodella* and *Ph. koolunga* . This will contribute to the improvement and development of pea genotypes to rescue pea production and subsequently enhance natural resource management through pea integration into crop rotation.

7. General Discussion and Conclusion

7.1 General discussion

Ascochyta blight of field pea caused by fungal pathogens including *Peyronellaea pinodes, Peyronellaea pinodella, Ascochyta pisi* and *Phoma koolunga* [31, 57, 240], continues to be a prime pea production constraint worldwide. Limited availability of molecular data in public sequence databases is one of the major bottlenecks towards the advancement of molecular studies in the interaction of Ascochyta Blight Pathogens of Field Pea (ABPFP) with their host - pea. Prior to this thesis project, there no published genome sequence has been available for the pathogen species in Didymellaceae family. In this thesis, whole genome and *in planta* transcriptome sequencing were employed to expand our understanding of ABPFP – pea interaction at molecular level. The project tackled the assembly, annotation and comparative evaluation of the genomes of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* (Chapter 2), identified conserved pathogenicity related genes (Chapter 3 and 4), and predicted effector-like candidate genes in ABPFP *via* comparative analysis of *in planta* expressed genes in these pathogens (Chapter 5).

The first draft *de novo* genome assembly of *P. pinodes* was generated from short read (75 bp) by the Illumina Solexa sequencing platform in 2012 [158]. This fragmented genome was used to generate peptide databases, which allowed for the proteomics-based prediction of 16 strong potential candidate effector-like genes in *P. pinodes* [158]. In the current project, substantial improvements to the genome assembly of *P. pinodes* (M074) was achieved through the sequencing of multiple libraries of different insert size and re-assembling these sequences using various updated techniques (Chapter 2).

One of the main challenges of this PhD project was a lack of reference genome sequence data from closely related species. The sequencing and assembly strategies employed in this project allowed us to generate a good quality *P. pinodes* (M074)

reference genome assembly, which will serve as a reference for the pathogens in the Didymellaceae. This reference species was assembled to 33.3 mega base pairs (Mbp) with N50 and L50 of 15 and 817 kb, respectively. The genome sequences of *P. pinodella, A. pis*i, *Ph. koolunga* and another *P. pinodes* isolates were assembled to equivalent size varying between 29.4 – 34 Mbp. Equilibrated GC content of between 51 -53 %GC was observed in these species. The average total length of repetitive DNA regions, which comprise the majority of non-equilibrated (AT-rich) GC content, varied between 2.3% in *Ph. koolunga* (FT04040*)* to 3% in *P. pinodes* (M074). The total number of predicted open reading frames (ORFs) of ≥ 50 amino acids in length varied from 9, 679 in *A. pisi* (Georgia-12) to 12, 482 in *P. pinodes* (MP1). The availability of new genome data contributes to previously scarce molecular data available for studying the host – ABPFP interaction. It now represents the majority of sequenced species of necrotrophic pathogens of legume crops in the Didymellaceae family. The data will also indirectly aid the discovery and characterization of novel pathogenicity related genes in other species through comparative genomics.

In this thesis, the genome sequences have aided further comparative genome analysis, including: repeat induced point mutation analysis, global genome alignment and synteny analyses, and large scale orthology analysis (Chapter 2). RIP-like mutation analysis indicated a bias towards more frequent mutation of CpT ↔ TpT and CpA ↔ TpA in both long terminal repeat (LTR) and DNA transposons. RIP-like activities were stronger in *P. pinodella* and *Ph. koolunga* than in *P. pinodes* and *A. pisi,* indicating the former 2 species possess better genome defenses against repeat invasion. Unusually high RIP-like mutations was observed in DNA transposons of *P. pinodes* (MP1), *Ph. koolunga* and *A. pisi* (Georgia-7 and Georgia-12) which is likely related to their relatively high proportion of DNA transposons.

The synteny relationship between *P. pinodes* and closely related pathogens from Didymellaceae family, *Dothideomycetes* and Sordariomycetes is congruent to the established evolutionary relationship of the fungal pathogens used in this study. Very closely related pathogens showed macrosynteny, more distantly related species showed a gradual breakdown of macrosynteny through large-scale rearrangements,

frequently inversions of the same sequence (mesosynteny). Comparison between closely related species, *P. pinodes* and *P. pinodella* indicated that about 91- 94 % of *P. pinodella* genome could be aligned to the *P. pinodes* (M074) genome. Moreover, about 71 - 80% of the unassembled reads from *P. pinodella* could be mapped to the *P. pinodes* reference isolate M074. The data suggested that only a small fraction (<10%) of the genome accounted for the observed sequence differences between *P. pinodes* and *P. pinodella*. These fractions were classified as 'species-specific regions' and were also found to harbor isolate-specific genes. Most of the sequences present in one species and absent the other are short, which is likely due to a proximity to flanking repetitive DNA that hampers un-fragmented genome assembly in these regions. These isolate/species-specific regions could be exploited as targets for designing species-specific markers to aid rapid discrimination between *P. pinodes* and *P. pinodella* in species complexes infecting field pea.

The new genome sequences also allowed us to identify ABPFP-specific protein families via orthology analysis. Among 297, 042 proteome sequences from 24 fungal genomes classified into 86,702 protein families, 4,555 protein families were found core proteins of *P. pinodes*, *P. pinodella*, *Ph. koolunga* and *A. pisi* while 109 protein families were found only in ABPFP. Most of those proteins with known functional domains are associated to oxidoreductases, hydrolases, transferases, peptidases, kinases, transcription factors, toxin biosynthesis or had conserved protein domains of unknown function. These proteins common to AB pathogens may determine characteristics that are common to their niche adaptation on field pea, while those proteins specific to either of *P. pinodes*, *P. pinodella*, *Ph. koolunga* or *A. pisi* may have roles in determining host/cultivar specificity.

Genome analysis also shed light on the possibilities of ABPFP being able to produce secondary metabolites that are encoded by a cluster of genes (Chapter 2). A number of putative type I polyketide synthase gene cluster (PKS) and non-ribosomal peptide synthase (NRPS) were predicted form the genome assemblies. The product of these

genes (and their flanking genes within the cluster) and their possible roles in ABPFP remains to be determined.

Genome sequencing and comparative analysis also uncovered common strategies of cell wall de-polymerization to breach the physical defences and structure of the host cell. The *P. pinodes, P. pinodella, Ph. koolunga* and *A. pisi* contained 562 – 581, 592 – 600, 511– 520 and 503 – 509 carbohydrate active enzymes (CAZyme), respectively (Chapter 5). Pea-infecting species had exceptionally high proportions of pectin degrading enzymes (≥31), including the non-ABPFP species *Nectria haematococca* (34), compared to other pathogens of other dicot plants such as *Alternaria brassicicola (23), Leptosphaeria maculans* (18) and *Botrytis cinerea* (11). High pectinase gene content in dicot-infecting species appears to be correlated with the high pectin content of dicot plants, with Pea pathogens exhibiting the highest frequencies. Pectinase was implicated as virulence factor in pea pathogen *N. haematococca* [263], *Colletotrichum gloesporioides* [341] and *Colletotrichum lindemuthianum* [179].

One of the major achievements of this PhD thesis project was the identification of a well-defined class of pathogenicity genes, the necrosis- and ethylene-inducing peptide like proteins (NLPs). NLP1 and NLP2 are effector genes identified from the whole genome sequence, employing comparative and phylogenetic analysis. In collaboration with other researchers, we have demonstrated the necrotic activities of NLP2 on pea and other dicot plants (Chapter 3). The NLP1 and NLP2 are likely duplicated genes, but only NLP2 has activity to induce necrosis. Interestingly, NLP2 was not detected in at least in four pathogens infecting cereals, suggesting niche adaptive evolution of Didymellaceae pathogens to dicotyledonous hosts. Because NLP1 is expressed during infection and did not show any necrotic activity, NLP1 may have a role other than cytolysis such as triggering host immunity leading to programmed cell death and nutrient release.

Beyond the NLPs, additional functionally conserved pathogenicity genes could be predicted using motif signature searches complemented with evolutionary analysis. There were three main lysine motif domain-containing proteins (LysM) identified in *P. pinodes*, *P. pinodella*, *Ph. koolunga, A. pisi.* The first is a repeat of LysM domain

with signal peptide. The second group consisted of LysM repeats without signal peptide. The third group consisted of LysM domain with a Cyanovirin-N homology (CVNH) sandwiched within LysM (Chapter 4). In addition, a number of putative pisatin demethylase (PDA) encoding genes (3 in each *Ph. koolunga* and *A. pisi* and 10 – 12 in *P. pinodes* and *P. pinodella*) were predicted in ABPFP. No known report available for the existence and functional relevance of LysM in ABPFP. Available literature indicates the broad relevance of LysM as a virulence factor in a number of fungal pathogens [62, 63, 210, 216, 219]. These functionally known LysM domain containing proteins often share common characteristics, such as extracellular secretion signals and repeats of the LysM domain within the sequence. Similarly, some report suggested the potential role of PDA in *P. pinodes*, *P. pinodella* and *A. pisi*.

In addition to the conserved effector genes detected in ABPFP, genome sequences of *P. pinodes*, *P. pinodella* and *A. pisi* also allowed us to predict novel candidate effector-like genes (Chapter 5). To this end, 1681, 1451 and 1518 protein sequences were predicted to have extracellular secretion signals in *P. pinodes*, *P. pinodella* and *Ph. koolunga*, respectively (Chapter 5). Among these, 223 in *P. pinodes*, 226 in *P. pinodella* and 140 in *Ph. koolunga* of the secreted proteins were designated as putative effector candidate genes. Furthermore, 8.9 % (21) of these shortlisted candidates in *P. pinodes*, 3.7% (8) in *P. pinodella* and 47% (67) in *Ph. koolunga* were also observed to be species-specific, as identified by the previous cross-species orthology analysis (Chapter 2, Chapter 5).

Effectors genes that are likely to be expressed during host infection have been at the forefront of many host-pathogen interaction studies [167, 228, 352]. Expression analysis of RNA-seq sequence indicated that 20% of predicted effectors in *P. pinodes* were significantly upregulated at early time points *in planta* relative to at least one *in vitro* condition (Chapter 5). Similarly, the expression of 16 and 34 of the effector like candidate genes was high at early time points in *P. pinodella* and *Ph. koolunga*, respectively. The early activation of these effector like candidate genes could support their possible role during infection and disease development in ABPFP.

Combining various types of supporting evidence, an overall ranking of effector candidates from most to least probable was performed (Chapter 5). A recent software tool for the *de novo* effector prediction of effectors – EffectorP [291] was also applied to the candidates. These bioinformatics techniques complemented the previous analysis of known protein domains (and sequence comparison) (Chapter 3 and Chapter 4) which had enabled the identification of common effector like candidate genes with high confidence. For instance, the necrosis- and ethylene inducing - peptide like protein (NLPs, Pfam: PF05630), CyanoVirin-N (CVNH) [Pfam: PF08881] and LysM [Pfam: PF01476] domain containing protein and putative cytochrome P450 pisatin demethylase (*Ppdes_5.300:* GO: 0005506) was predicted as putative candidate effector-like protein in at least one of the ABPFP using either the ranking method or EffectorP. In addition, the *de novo* effector prediction in ABPFP reported genes of similar function across the three pathogens. Genes encoding cellulase, hemicellulase, pectinase, cutinase, and fungal-specific extracellular membrane proteins (CFEM, Pfam: PF05730) were part of a common set of highly ranked pathogenesis-related genes that were expressed at early infection stage. This may suggest their possible role during infection and disease development (Chapter 5).

A number of genes involved in protein metabolism were common to the three fungal pathogens and predicted as putative effector candidate genes. These include subtilisin-like protein, an aspartyl protease, and an asparaginase. While subtilisin and aspartyl protease are reported in past studies as associated with fungal pathogenicity [282] asparaginase is commonly reported as a chemotherapy agent used to treat blood cancer in humans [18]. According to Batool *et al.* [18], only 9 (5 of them are *Aspergillus*) fungal species have previously been reported to produce asparaginase.

7.2 Conclusion and future direction

This research project has laid foundations for pea-pathogen interaction studies at a molecular level through the generation of whole genome and *in planta* transcriptome sequence resources. High quality reference genome (*P. pinodes* isolate M074) for the pathogens in the class Didymellaceae has been established. With that in mind, the

genome assemblies of *P. pinodella*, *Ph. koolunga* and *A. pisi* were generated from short read Illumina sequencing platform, and more advanced long-read platforms are now readily available. Continuous improvement of the genome sequences through additional sequencing and re-assembly and/or the inclusion of transcriptomic data for the correction of gene annotations are elements that could be considered in the near future. Meanwhile, release of currently available genomic data would enhance scientific community engagement and their contribution toward tackling these economically devastating pathogens.

Pea cell wall de-polymerization is likely a common strategies employed by ABPFP to breach physical host defences and cell structure. Consistent with dicot plant cell wall composition, CAZymes targeting pectin degradation were expanded in the ABPFP indicating metabolic adaptation to their common host. With this regard, *in planta* expressed pectin related genes in *P. pinodes*, *P. pinodella* or *Ph. koolunga* would be a good candidate for further functional analysis and determination of their possible role in pathogenesis.

Comparative sequence and phylogenetic analysis are powerful techniques to effectively predict functionally active effector candidate genes. Further analysis revealed that NLP1 is closely related to pathogens from monocot and NLP2 is more closely related to pathogens from dicot and absent from three pathogens infecting monocot. Purified protein infiltration assay on field pea stipules and other dicot plants proved that only NLP2 induce necrosis when purified proteins were infiltrated. The identification of NLP2 is important as it can be used as positive control in future functional studies of effector candidate genes in ABPFP. The lack of necrosis activities in NLP1 would simplify determination of NLP2 contribution to virulence via gene knockout/knockdown. Further studies would also be important to uncover the role of expressed NLP1 in ABPFP.

Conserved fungal effectors like LysM and pisatin demethylase (PDA) were also predicted in the project and presumed to play role during disease development. Many novel effector like candidate genes were predicted from these genomic resources. A

total of 223 genes in *P. pinodes*, 212 in *P. pinodella* and 140 genes in *Ph. koolunga* were predicted as putative effector like candidate proteins. These early expressed candidate genes are likely to be putative effector candidate genes. Further functional analysis using gene knock-out or gene silencing could help to determine the role of predicted effectors during infection and disease development. Validation of the CELPs will ultimately help the development of rapid screening techniques for disease resistance across pea cultivars, as well as in the development of targeted breeding techniques. Knowledge of effectors can be useful to exploit disease susceptibility gene in host given these effectors trigger disease susceptibility [103, 196, 324]. This may be promising in future improvements to commercial pea cultivars, as there is currently no resistance available within the commercial pea germplasm.

The availability of new genome data contributes to previously scarce molecular data available for studying the host – ABPFP interaction. It now represents the majority of sequenced species of necrotrophic pathogens of legume crops in the Didymellaceae family. These resources are a valuable tool for the development of markers for the study of pathogen population structure and for genotyping, as well as for the prediction of candidate effector genes and advancing the understanding of pathogenicity mechanisms in ABPFP. The data will also indirectly aid the discovery and characterization of novel pathogenicity related genes in other species through comparative genomics.

The recently established of International Consortium for the Pea Genome Sequencing (ICPG) (https://pag.confex.com/pag/xxiii/webprogram/Paper14072.html) has been encouraging pathologists and breeders to collaborate on field pea research projects. The *in planta* transcriptomic data generated during this thesis project is also major contribution towards the pea genome sequencing initiatives. These transcriptomic data will be further used to study pea defence response mechanisms and to map metabolic the pathways involved in pea defence reactions in response to infection by *P. pinodes*, *P. pinodella* and *Ph. koolunga.* This gives complete understanding of the interplay among ABPFP and *Pisum sativum* at molecular level. This will lead to the improvement and development of pea genotypes through establishment of novel

breeding techniques. Development and use of improved germplasm will increase production and productivity leading to increased supply of pea for food/feed at either at individual or industry level. It also enhance natural resource management through pea integration into crop rotation.

To sum up, concerted effort has been made to expand our level of understand why ABPFP continues to be threat to pea production worldwide. Sequencing and analysis of the genomes of fungal pathogens *P. pinodes, P. pinodella, Ph. koolunga* and *A. pisi* has highlighted the genome features of each species. First, comparative analysis of the genome sequences indicated expansion of some CAZyme families in these four fungal pathogens. Secondly, sequencing uncovered the presence of conserved pathogenicity related genes such as NLPs, LysM and PDA. Thirdly, genome sequencing has revealed the potential of ABPFP pathogens to produce secondary metabolites by a cluster of genes. Fourth, diverse putative effector like candidate genes have been predicted, which might play a decisive role in disease development and will inform future experimental validations. Finally, incorporation of *in planta* transcriptomic data has greatly increased the reliability of genes and effector – like candidate genes predictions made in these studies, and will have flow-on benefits to the genome project for the pea host.

Release of the new genome data to the public will serve to aid further discovery and characterization of novel pathogenicity related genes in ABPFP and other species by the wider scientific community. Once the pea genome available, it is perceived that genomic and previous pea genetic (such as quantitative trait loci) and single nucleotide polymorphism) data [75, 156, 287, 313, 314] would be integrated and able to display the interactome at molecule level. This will lead to the improvement and development of pea genotypes and rescue pea production currently under threat. Use of improved varieties increase production leading to increased pea supply for both human food and animal feed. Furthermore, improved Pea production is of broad benefit to agriculture as it can enhance the management of natural soil fertility in cereal based farming system and boost subsequent cereal crop production.

8. References

1. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research*. 2010;38(suppl 2):W7-W13.
2. Akamatsu HO, Chilvers MI, Kaiser WJ, Peever TL. Karyotype polymorphism and chromosomal rearrangement in populations of the phytopathogenic fungus, *Ascochyta rabiei*. *Fungal Biology*. 2012;116(11):1119-33.
3. Akcapinar G, Kappel L, Sezerman O, Seidl-Seiboth V. Molecular diversity of LysM carbohydrate-binding motifs in fungi. *Current Genetics*. 2015;61(2):103-13.
4. Ali S, Nitschke L, Dube A, Krause M, Cameron B. Selection of pea lines for resistance to pathotypes of *Ascochyta pinodes*, *A. pisi* and *Phoma medicaginis* var. *pinodella*. *Australian Journal of Agricultural Research*. 1978;29(4):841-9.
5. Ali SM, Sharma B, Ambrose MJ. Current status and future strategy in breeding pea to improve resistance to biotic and abiotic stresses. *Euphytica*. 1993;73(1-2):115-26.
6. Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman D. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*. 1997;25:3389 - 402.
7. Amselem J, Cuomo CA, van Kan JA, Viaud M, Benito EP, Couloux A, Coutinho PM, de Vries RP, Dyer PS, Fillinger S, Fournier E, Gout L, Hahn M, Kohn L, Lapalu N, Plummer KM, Pradier JM, Quevillon E, Sharon A, Simon A, ten Have A, Tudzynski B, Tudzynski P, Wincker P, Andrew M, Anthouard V, Beever RE, Beffa R, Benoit I, Bouzid O, Brault B, Chen Z, Choquer M, Collemare J, Cotton P, Danchin EG, Da Silva C, Gautier A, Giraud C, Giraud T, Gonzalez C, Grossetete S, Guldener U, Henrissat B, Howlett BJ, Kodira C, Kretschmer M, Lappartient A, Leroch M, Levis C, Mauceli E, Neuveglise C, Oeser B, Pearson M, Poulain J, Poussereau N, Quesneville H, Rascle C, Schumacher J, Segurens B, Sexton A, Silva E, Sirven C, Soanes DM, Talbot NJ, Templeton M, Yandava C, Yarden O, Zeng Q, Rollins JA, Lebrun MH,

Dickman M. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genetics*. 2011;7(8):e1002230.

8. Amselem J, Lebrun MH, Quesneville H. Whole genome comparative analysis of transposable elements provides new insight into mechanisms of their inactivation in fungal genomes. *BMC Genomics*. 2015;16(1).
9. Amsellem Z, Cohen BA, Gressel J. Engineering hypervirulence in a mycoherbicidal fungus for efficient weed control. *Nature Biotechnology*. 2002;20(10):1035-9.
10. Anderson AJ, Powelson ML. Production of plant cell wall degrading enzymes by *Phoma medicagnis f.sp. pinodella*. *Phytopathology*. 1979;69:372-5.
11. Aveskamp MM, de Gruyter J, Woudenberg JHC, Verkley GJM, Crous PW. Highlights of the Didymellaceae: A polyphasic approach to characterise Phoma and related pleosporalean genera. *Studies in Mycology*. 2010;65:1-60.
12. Bae H, Kim MS, Sicher RC, Bae HJ, Bailey BA. Necrosis- and ethylene-inducing peptide from *Fusarium oxysporum* induces a complex cascade of transcripts associated with signal transduction and cell death in Arabidopsis. *Plant Physiology*. 2006;141(3):1056-67.
13. Bailey BA. Purification of a protein from culture filtrates of *Fusarium oxysporum* that induces ethylene and necrosis in leaves of *Erythroxylum coca*. *Phytopathology*. 1995;85(10):1250-5.
14. Bailey BA, Apel-Birkhold PC, Luster DG. Expression of NEP1 by *Fusarium oxysporum* f. sp. *erythroxyli* after gene replacement and overexpression using polyethylene glycol-mediated transformation. *Phytopathology*. 2002;92(8):833-41.
15. Baker SE, Kroken S, Inderbitzin P, Asvarak T, Li B-Y, Shi L, Yoder OC, Turgeon BG. Two polyketide synthase-encoding genes are required for biosynthesis of the polyketide virulence factor, t-toxin, by *Cochliobolus heterostrophus Molecular Plant-Microbe Interactions*. 2006;19(2):139-49.
16. Bao Z, Eddy SR. Automated *de novo* identification of repeat sequence families in sequenced genomes. *Genome Research*. 2002;12(8):1269-76.

17. Barrett LG, Kniskern JM, Bodenhausen N, Zhang W, Bergelson J. Continua of specificity and virulence in plant host-pathogen interactions: causes and consequences. *New Phytologist*. 2009;183(3):513-29.
18. Batool T, Makky EA, Jalal M, Yusoff MM. A Comprehensive Review on l-Asparaginase and Its Applications. *Applied Biochemistry and Biotechnology*. 2015;10.1007/s12010-015-1917-3.
19. Beed FD, Sue RE, Strange RN. Variation in the production of ascochitine by *Ascochyta fabae*. *Mycological Research*. 1994;98(9):1069-76.
20. Berto P, Comménil P, Belingheri L, Dehorter B. Occurrence of a lipase in spores of *Alternaria brassicicola* with a crucial role in the infection of cauliflower leaves. *FEMS Microbiology Letters*. 1999;180(2):183-9.
21. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Research*. 2013;41(W1):W204-W12.
22. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*. 2011;27(4):578-9.
23. Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. *Genome Biology*. 2012;13(6):R56.
24. Bohm H, Albert I, Oome S, Raaymakers TM, Van den Ackerveken G, Nurnberger T. A conserved peptide pattern from a widespread microbial virulence factor triggers pattern-induced immunity in Arabidopsis. *PLoS Pathogens*. 2014;10(11):e1004491.
25. Boller T. Chemoperception of microbial signals in plant cells. *Annual Review of Plant Physiology and Plant Molecular Biology*. 1995;46:189-214.
26. Bolton MD, van Esse HP, Vossen JH, de Jonge R, Stergiopoulos I, Stulemeijer IJ, van den Berg GC, Borras-Hidalgo O, Dekker HL, de Koster CG, de Wit PJ, Joosten MH, Thomma BP. The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. *Molecular Microbiology*. 2008;69(1):119-36.
27. Borodovsky M, Lomsadze A. Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES. *Current Protocols in Bioinformatics*. 2011;35:4.6.1-4.6.10.

28. Boros L, Marcinkowska J. Assessment of selected pea genotypes reaction to ascochyta blight under field conditions and the impact of disease severity on yield components. *Journal of Agricultural Science*. 2010;2(3).
29. Bowen JK, Lewis BG, Matthews P. Discovery of the teleomorph of *Phoma medicaginis* var. *pinodella* in culture. *Mycological Research*. 1997;101(1):80-4.
30. Braumann I, van den Berg M, Kempken F. Repeat induced point mutation in two asexual fungi, Aspergillus niger and Penicillium chrysogenum. *Current Genetics*. 2008;53(5):287-97.
31. Bretag TW, Keane PJ, Price TV. The epidemiology and control of ascochyta blight in field peas: A review. *Australian Journal of Agricultural Research*. 2006;57(8):883-902.
32. Buist G, Steen A, Kok J, Kuipers OP. LysM, a widely distributed protein motif for binding to (peptido)glycans. *Molecular Microbiology*. 2008;68(4):838-47.
33. Byrne KP, Wolfe KH. Consistent patterns of rate asymmetry and gene loss indicate widespread neofunctionalization of yeast genes after whole-genome duplication. *Genetics*. 2007;175(3):1341-50.
34. Cabral A, Oome S, Sander N, Kufner I, Nurnberger T, Van den Ackerveken G. Nontoxic Nep1-like proteins of the downy mildew pathogen *Hyaloperonospora arabidopsidis*: repression of necrosis-inducing activity by a surface-exposed region. *Molecular Plant-Microbe Interactions*. 2012;25(5):697-708.
35. Calmes B, Morel-Rouhier M, Bataillé-Simoneau N, Gelhaye E, Guillemette T, Simoneau P. Characterization of glutathione transferases involved in the pathogenicity of *Alternaria brassicicola*. *BMC Microbiology*. 2015;15(1).
36. Chen LH, Lin CH, Chung KR. A nonribosomal peptide synthetase mediates siderophore production and virulence in the citrus fungal pathogen *Alternaria alternata*. *Molecular Plant Pathology*. 2013;14(5):497-505.
37. Chen X-L, Shi T, Yang J, Shi W, Gao X, Chen D, Xu X, Xu J-R, Talbot NJ, Peng Y-L. N-Glycosylation of Effector Proteins by an α-1,3-Mannosyltransferase Is Required for the Rice Blast Fungus to Evade Host Innate Immunity. *The Plant Cell*. 2014;26(3):1360-76.

38. Chilvers MI, Rogers JD, Dugan FM, Stewart JE, Chen W, Peever TL. *Didymella pisi* sp. nov., the teleomorph of *Ascochyta pisi*. *Mycological Research*. 2009;113(3):391-400.

39. Cho Y, Jang M, Srivastava A, Jang J-H, Soung N-K, Ko S-K, Kang D-O, Ahn JS, Kim BY. A pectate lyase-coding gene abundantly expressed during early stages of infection is required for full virulence in *Alternaria brassicicola*. *PLoS ONE*. 2015;10(5):e0127140.

40. Chooi YH, Muria-Gonzalez MJ, Solomon PS. A genome-wide survey of the secondary metabolite biosynthesis genes in the wheat pathogen. *Mycology*. 2014;5(3):192-206.

41. Chu CG, Faris JD, Xu SS, Friesen TL. Genetic analysis of disease susceptibility contributed by the compatible Tsn1-SnToxA and Snn1-SnTox1 interactions in the wheat-*Stagonospora nodorum* pathosystem. *Theoretical and Applied Genetics*. 2010;120(7):1451-9.

42. Cimmino A, Andolfi A, Fondevilla S, Abouzeid MA, Rubiales D, Evidente A. Pinolide, a new nonenolide produced by *Didymella pinodes*, the causal agent of ascochyta blight on *Pisum sativum Journal of Agricultural and Food Chemistry*. 2012;60(21):5273-8.

43. Ciuffetti L, Manning V, Pandelova I, Betts M, Martinez J. Host-selective toxins, Ptr ToxA and Ptr ToxB, as necrotrophic effectors in the *Pyrenophora tritici-repentis*-wheat interaction. *New Phytologist*. 2010;187:911 - 9.

44. Clulow SA, Lewis BG, Parker ML, Matthews P. Infection of pea epicotyls by *Mycosphaerella pinodes*. *Mycological Research*. 1991;95(7):817-20.

45. Clutterbuck A. Genomic evidence of repeat-induced point mutation (RIP) in filamentous ascomycetes. *Fungal Genetics and Biology*. 2011;48:306 - 26.

46. Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S, Schwartz DC, Freitag M, Ma LJ, Danchin EG, Henrissat B, Coutinho PM, Nelson DR, Straney D, Napoli CA, Barker BM, Gribskov M, Rep M, Kroken S, Molnar I, Rensing C, Kennell JC, Zamora J, Farman ML, Selker EU, Salamov A, Shapiro H, Pangilinan J, Lindquist E, Lamers C, Grigoriev IV, Geiser DM, Covert SF, Temporini E, Vanetten HD. The genome of *Nectria haematococca*: contribution of

supernumerary chromosomes to gene expansion. *PLoS Genetics*. 2009;5(8):e1000618.

47. Coleman JJ, White GJ, Rodriguez-Carres M, Vanetten HD. An ABC transporter and a cytochrome P450 of *Nectria haematococca* MPVI are virulence factors on pea and are the major tolerance mechanisms to the phytoalexin pisatin. *Molecular Plant-Microbe Interactions*. 2011;24(3):368-76.
48. Condon BJ, Leng Y, Wu D, Bushley KE, Ohm RA, Otillar R, Martin J, Schackwitz W, Grimwood J, MohdZainudin N, Xue C, Wang R, Manning VA, Dhillon B, Tu ZJ, Steffenson BJ, Salamov A, Sun H, Lowry S, LaButti K, Han J, Copeland A, Lindquist E, Barry K, Schmutz J, Baker SE, Ciuffetti LM, Grigoriev IV, Zhong S, Turgeon BG. Comparative genome structure, secondary metabolite, and effector coding capacity across *Cochliobolus* pathogens. *PLoS Genetics*. 2013;9(1):e1003233.
49. Condon BJ, Oide S, Gibson DM, Krasnoff SB, Turgeon BG. Reductive iron assimilation and intracellular siderophores assist extracellular siderophore-driven iron homeostasis and virulence. *Molecular Plant-Microbe Interactions*. 2014;27(8):793-808.
50. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21(18):3674-6.
51. Cuesta Arenas Y, Kalkman ERIC, Schouten A, Dieho M, Vredenbregt P, Uwumukiza B, Osés Ruiz M, van Kan JAL. Functional analysis and mode of action of phytotoxic Nep1-like proteins of *Botrytis cinerea*. *Physiological and Molecular Plant Pathology*. 2010;74(5-6):376-86.
52. Cuomo C, Guldener U, Xu J, Trail F, Turgeon B, Di Pietro A, Walton J, Ma L, Baker S, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang Y, Decaprio D, Gale L, Gnerre S, Goswami R, Hammond-Kosack K, Harris L, Hilburn K, Kennell J, Kroken S, Magnuson J, Mannhaupt G, Mauceli E, Mewes H, Mitterbauer R, Muehlbauer G. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science*. 2007;317:1400 - 2.

53. Dallal Bashi Z, Hegedus DD, Buchwaldt L, Rimmer SR, Borhan MH. Expression and regulation of *Sclerotinia sclerotiorum* necrosis and ethylene-inducing peptides (NEPs). *Molecular Plant Pathology*. 2010;11(1):43-53.

54. Dalmais B, Schumacher J, Moraga J, P LEP, Tudzynski B, Collado IG, Viaud M. The *Botrytis cinerea* phytotoxin botcinic acid requires two polyketide synthases for production and has a redundant role in virulence with botrydial. *Molecular Plant Pathology*. 2011;12(6):564-79.

55. Dang H, Pryor B, Peever T, Lawrence C. The *Alternaria* genomes database: a comprehensive resource for a fungal genus comprised of saprophytes, plant pathogens, and allergenic species. *BMC Genomics*. 2015;16(1):239.

56. Darby P, Lewis BG, Matthews P. Diversity of virulence within *Ascochyta pisi* and resistance in the genus *Pisum*. *Plant Pathology*. 1986;35(2):214-23.

57. Davidson JA, Hartley D, Priest M, Krysinska-Kaczmarek M, Herdina, McKay A, Scott ES. A new species of *Phoma* causes ascochyta blight symptoms on field peas (*Pisum sativum*) in South Australia. *Mycologia*. 2009;101(1):120-8.

58. Davidson JA, Krysinska-Kaczmarek M, Herdina, McKay A, Scott ES. Comparison of cultural growth and *in planta* quantification of *Didymella pinodes*, *Phoma koolunga* and *Phoma medicaginis* var. *pinodella*, causal agents of ascochyta blight on field pea (*Pisum sativum*). *Mycologia*. 2012;104(1):93-101.

59. Davies KA, De Lorono I, Foster SJ, Li D, Johnstone K, Ashby AM. Evidence for a role of cutinase in pathogenicity of *Pyrenopeziza brassicae* on brassicas. *Physiological and Molecular Plant Pathology*. 2000;57(2):63-75.

60. de Gruyter J, Aveskamp MM, Woudenberg JHC, Verkley GJM, Groenewald JZ, Crous PW. Molecular phylogeny of *Phoma* and allied anamorph genera: Towards a reclassification of the *Phoma* complex. *Mycological Research*. 2009;113(4):508-19.

61. de Jonge R. In silico identification and characterization of effector catalogs. *Methods in Molecular Biology*. 2012;835:415-25.

62. de Jonge R, Thomma B. Fungal LysM effectors: extinguishers of host immunity? *Trends Microbiology*. 2009;17(4):151 - 7.

63. de Jonge R, van Esse H, Kombrink A, Shinya T, Desaki Y, Bours R, van der Krol S, Shibuya N, Joosten M, Thomma B. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*. 2010;329(5994):953 - 5.

64. de Vries RP, vanKuyk PA, Kester HC, Visser J. The *Aspergillus niger* faeB gene encodes a second feruloyl esterase involved in pectin and xylan degradation and is specifically induced in the presence of aromatic compounds. *Biochemical Journal*. 2002;363(Pt 2):377-86.

65. De Wit PJ, Mehrabi R, Van den Burg HA, Stergiopoulos I. Fungal effector proteins: past, present and future. *Molecular Plant Pathology*. 2009;10(6):735-47.

66. de Wit PJ, van der Burgt A, Okmen B, Stergiopoulos I, Abd-Elsalam KA, Aerts AL, Bahkali AH, Beenen HG, Chettri P, Cox MP, Datema E, de Vries RP, Dhillon B, Ganley AR, Griffiths SA, Guo Y, Hamelin RC, Henrissat B, Kabir MS, Jashni MK, Kema G, Klaubauf S, Lapidus A, Levasseur A, Lindquist E, Mehrabi R, Ohm RA, Owen TJ, Salamov A, Schwelm A, Schijlen E, Sun H, van den Burg HA, van Ham RC, Zhang S, Goodwin SB, Grigoriev IV, Collemare J, Bradshaw RE. The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLoS Genetics*. 2012;8(11):e1003088.

67. Dean R, Talbot N, Ebbole D, Farman M, Mitchell T, Orbach M, Thon M, Kulkarni R, Xu J, Pan H, Read N, Lee Y, Carbone I, Brown D, Oh Y, Donofrio N, Jeong J, Soanes D, Djonovic S, Kolomiets E, Rehmeyer C, Li W, Harding M, Kim S, Lebrun M, Bohnert H, Coughlan S, Butler J, Calvo S, Ma L. The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature*. 2005;434:980 - 6.

68. Dean RA, Lichens-Park A, Kole C. Genomics of plant-associated fungi and oomycetes: Dicot pathogens. 2014. 239 p.

69. Dean RA, Lichens-Park A, Kole C. Genomics of plant-associated fungi: Monocot pathogens. 2014. 201 p.

70. Dell A, Galadari A, Sastre F, Hitchen P. Similarities and Differences in the Glycosylation Mechanisms in Prokaryotes and Eukaryotes. *International Journal of Microbiology*. 2010;2010:14.

71. Delserone LM, McCluskey K, Matthews DE, Vanetten HD. Pisatin demethylation by fungal pathogens and nonpathogens of pea: Association with pisatin tolerance and virulence. *Physiological and Molecular Plant Pathology*. 1999;55(6):317-26.

72. Deng J, Carbone I, Dean R. The evolutionary history of Cytochrome P450 genes in four filamentous Ascomycetes. *BMC Evolutionary Biology*. 2007;7:30.

73. DeZwaan TM, Carroll AM, Valent B, Sweigard JA. *Magnaporthe grisea* pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive substrate cues. *The Plant Cell*. 1999;11(10):2013-30.

74. Dong S, Kong G, Qutob D, Yu X, Tang J, Kang J, Dai T, Wang H, Gijzen M, Wang Y. The NLP toxin family in *Phytophthora sojae* includes rapidly evolving groups that lack necrosis-inducing activity. *Molecular Plant-Microbe Interactions*. 2012;25(7):896-909.

75. Duarte J, Rivière N, Baranger A, Aubert G, Burstin J, Cornet L, Lavaud C, Lejeune-Hénaut I, Martinant JP, Pichon JP, Pilet-Nayel ML, Boutet G. Transcriptome sequencing for high throughput SNP development and genetic mapping in Pea. *BMC Genomics*. 2014;15(1).

76. Duyvesteijn RG, van Wijk R, Boer Y, Rep M, Cornelissen BJ, Haring MA. Frp1 is a *Fusarium oxysporum* F-box protein required for pathogenicity on tomato. *Molecular Microbiology*. 2005;57(4):1051-63.

77. Eddy S. Accelerated Profile HMM Searches. *PLoS Computational Biology*. 2011;7(10):e1002195.

78. Ellwood S, Liu Z, Syme R, Lai Z, Hane J, Keiper F, Moffat C, Oliver R, Friesen T. A first genome assembly of the barley fungal pathogen *Pyrenophora teres* f. *teres*. *Genome Biology*. 2010;11(11):R109.

79. Evidente A, Lanzetta R, Capasso R, Vurro M, Botralico A. Pinolidoxin, a phytotoxic nonenolide from *Ascochyta pinodes*. *Phytochemistry*. 1993;34(4):999-1003.

80. Fang W, Feng J, Fan Y, Zhang Y, Bidochka MJ, Leger RJS, Pei Y. Expressing a fusion protein with protease and chitinase activities increases the virulence of the insect pathogen *Beauveria bassiana*. *Journal of Invertebrate Pathology*. 2009;102(2):155-9.

81. FAO. FAO Statistical Pocketbook. 2015:236.

82. Fellbrich G, Romanski A, Varet A, Blume B, Brunner F, Engelhardt S, Felix G, Kemmerling B, Krzymowska M, Nurnberger T. NPP1, a *Phytophthora*-associated trigger of plant defense in parsley and *Arabidopsis*. *The Plant Journal*. 2002;32(3):375-90.

83. Feng BZ, Zhu XP, Fu L, Lv RF, Storey D, Tooley P, Zhang XG. Characterization of necrosis-inducing NLP proteins in *Phytophthora capsici*. *BMC Plant Biology*. 2014;14:126.

84. Fernández-Aparicio M, Amri M, Kharrat M, Rubiales D. Intercropping reduces *Mycosphaerella pinodes* severity and delays upward progress on the pea plant. *Crop Protection*. 2010;29(7):744-50.

85. Fitzpatrick DA. Horizontal gene transfer in fungi. *FEMS Microbiology Letters*. 2012;329(1):1-8.

86. Fondevilla S, Küster H, Krajinski F, Cubero JI, Rubiales D. Identification of genes differentially expressed in a resistant reaction to *Mycosphaerella pinodes* in pea using microarray technology. *BMC Genomics*. 2011;12.

87. Fondevilla S, Satovic Z, Rubiales D, Moreno MT, Torres AM. Mapping of quantitative trait loci for resistance to *Mycosphaerella pinodes* in *Pisum sativum* subsp. *syriacum*. *Molecular Breeding*. 2008;21(4):439-54.

88. Foremska E, Marcinkowska J, Chelkowski J. Formation of ascochitine by plant pathogens of the genus *Ascochyta*. *Mycotoxin Research*. 1990;6(2):93-7.

89. Fries M, Ihrig J, Brocklehurst K, Shevchik VE, Pickersgill RW. Molecular basis of the activity of the phytopathogen pectin methylesterase. *EMBO Journal*. 2007;26(17):3879-87.

90. Friesen T, Chu C, Liu Z, Xu S, Halley S, Faris J. Host-selective toxins produced by *Stagonospora nodorum* confer disease susceptibility in adult wheat plants under field conditions. *Theoretical Applied Genetics*. 2009;118:1489 - 97.

91. Friesen T, Meinhardt S, Faris J. The *Stagonospora nodorum* - wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. *The Plant Journal*. 2007;51:681 - 92.

92. Friesen T, Stukenbrock E, Liu Z, Meinhardt S, Ling H, Faris J, Rasmussen J, Solomon P, McDonald B, Oliver R. Emergence of a new disease as a result of interspecific virulence gene transfer. *Nature Genetics*. 2006;38:953 - 6.

93. Friesen TL, Faris JD, Solomon PS, Oliver RP. Host-specific toxins: Effectors of necrotrophic pathogenicity. *Cellular Microbiology*. 2008;10(7):1421-8.

94. Fudal I, Ross S, Brun H, Besnard A, Ermel M, Kuhn M, Balesdent M, Rouxel T. Repeat-induced point mutation (RIP) as an alternative mechanism of evolution toward virulence in *Leptosphaeria maculans*. *Molecular Plant-Microbe Interactions*. 2009;22:932 - 41.

95. Gabaldon T, Koonin EV. Functional and evolutionary implications of gene orthology. *Nature Reviews Genetics*. 2013;14(5):360-6.

96. Galagan J, Calvo S, Borkovich K, Selker E, Read N, Jaffe D, FitzHugh W, Ma L, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen C, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson M, Werner-Washburne M, Selitrennikoff C, Kinsey J, Braun E, Zelter A, Schulte U, Kothe G, Jedd G, Mewes W. The genome sequence of the filamentous fungus *Neurospora crassa Nature*. 2003;422:859 - 68.

97. Galagan J, Selker E. RIP: the evolutionary cost of genome defense. *Trends in Genetics*. 2004;20:417 - 23.

98. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma L-J, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CPC, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford

A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B. The genome sequence of the filamentous fungus *Neurospora crassa Nature*. 2003;422(6934):859-68.

99. Galagan JE, Henn MR, Ma LJ, Cuomo CA, Birren B. Genomics of the fungal kingdom: Insights into eukaryotic biology. *Genome Research*. 2005;15(12):1620-31.
100. Gan PHP, Rafiqi M, Hardham AR, Dodds PN. Effectors of biotrophic fungal plant pathogens. *Functional Plant Biology*. 2010;37(10):913-8.
101. Gan Y, Hamel C, O'Donovan JT, Cutforth H, Zentner RP, Campbell CA, Niu Y, Poppy L. Diversifying crop rotations with pulses enhances system productivity. *Scientific Reports*. 2015;5(14625):14625.
102. Gao Y, Faris JD, Liu Z, Kim YM, Syme RA, Oliver RP, Xu SS, Friesen TL. Identification and Characterization of the SnTox6-Snn6 Interaction in the *Parastagonospora nodorum*–Wheat Pathosystem. *Molecular Plant-Microbe Interactions*. 2015;28(5):615-25.
103. Gawehns F, Cornelissen BJ, Takken FL. The potential of effector-target genes in breeding for plant innate immunity. *Microbial Biotechnolology*. 2013;6(3):223-9.
104. George HL, Hirschi KD, VanEtten HD. Biochemical properties of the products of cytochrome P450 genes (PDA) encoding pisatin demethylase activity in *Nectria haematococca*. *Archives of Microbiology*. 1998;170(3):147-54.
105. George HL, VanEtten HD. Characterization of pisatin-inducible cytochrome p450s in fungal pathogens of pea that detoxify the pea phytoalexin pisatin. *Fungal Genetics and Biology*. 2001;33(1):37-48.
106. Gijzen M, Nurnberger T. Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. *Phytochemistry*. 2006;67(16):1800-7.
107. Glemin S, Bataillon T. A comparative view of the evolution of grasses under domestication. *New Phytologist*. 2009;183(2):273-90.
108. Gniwotta F, Vogg G, Gartmann V, Carver TLW, Riederer M, Jetter R. What Do Microbes Encounter at the Plant Surface? Chemical Composition of Pea Leaf Cuticular Waxes. *Plant Physiology*. 2005;139(1):519-30.

109. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG. Life with 6000 genes. *Science*. 1996;274(5287):546, 63-7.

110. Gomez BL, Nosanchuk JD. Melanin and fungi. *Current Opinion in Infectious Diseases*. 2003;16(2):91-6.

111. Goodwin SB, M'Barek S B, Dhillon B, Wittenberg AH, Crane CF, Hane JK, Foster AJ, Van der Lee TA, Grimwood J, Aerts A, Antoniw J, Bailey A, Bluhm B, Bowler J, Bristow J, van der Burgt A, Canto-Canche B, Churchill AC, Conde-Ferraez L, Cools HJ, Coutinho PM, Csukai M, Dehal P, De Wit P, Donzelli B, van de Geest HC, van Ham RC, Hammond-Kosack KE, Henrissat B, Kilian A, Kobayashi AK, Koopmann E, Kourmpetis Y, Kuzniar A, Lindquist E, Lombard V, Maliepaard C, Martins N, Mehrabi R, Nap JP, Ponomarenko A, Rudd JJ, Salamov A, Schmutz J, Schouten HJ, Shapiro H, Stergiopoulos I, Torriani SF, Tu H, de Vries RP, Waalwijk C, Ware SB, Wiebenga A, Zwiers LH, Oliver RP, Grigoriev IV, Kema GH. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genetics*. 2011;7(6):e1002070.

112. Gorfu G. Yield loss of field pea due to ascochyta blight in central Ethiopia. *Pest Management Journal of Ethiopia*. 2000;4:89-5.

113. Gossmann TI, Ziegler M. Sequence divergence and diversity suggests ongoing functional diversification of vertebrate NAD metabolism. *DNA Repair*. 2014;23:39-48.

114. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*. 2011;29(7):644-52.

115. Graïa F, Lespinet O, Rimbault B, Dequard-Chablat M, Coppin E, Picard M. Genome quality control: RIP (repeat-induced point mutation) comes to Podospora. *Molecular Microbiology*. 2001;40(3):586-95.

116. Grandaubert J, Lowe R, Soyer J, Schoch C, Van de Wouw A, Fudal I, Robbertse B, Lapalu N, Links M, Ollivier B, Linglin J, Barbe V, Mangenot S, Cruaud C, Borhan H, Howlett B, Balesdent M-H, Rouxel T. Transposable element-assisted evolution and adaptation to host plant within the *Leptosphaeria maculans* - *Leptosphaeria biglobosa* species complex of fungal pathogens. *BMC Genomics*. 2014;15(1):891.

117. Greenshields DL, Liu G, Wei Y. Roles of Iron in Plant Defence and Fungal Virulence. *Plant Signaling & Behavior*. 2007;2(4):300-2.

118. Gugel R, Bailey K, Gossen B, Morall R, Dostaler D. Diseases of Field Crops in Canada. 2004. p. 119.

119. Guyon K, Balagué C, Roby D, Raffaele S. Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. *BMC Genomics*. 2014;15(1).

120. H D VanEtten, D E Matthews a, Matthews PS. Phytoalexin detoxification: Importance for pathogenicity and practical implications. *Annual Review of Phytopathology*. 1989;27(1):143-64.

121. Haas B, Salzberg S, Zhu W, Pertea M, Allen J, Orvis J, White O, Buell CR, Wortman J. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*. 2008;9(1):R7.

122. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*. 2013;8(8):1494-512.

123. Habibi A, Peever TL, Kim W, Chilvers MI, Chen W, Kaiser WJ, Muehlbauer FJ. First report of ascochyta blight of spotted locoweed (Astragalus lentiginosus) caused by Ascochyta sp. in Idaho. *Plant Disease*. 2015;99(10).

124. Han Y, Liu X, Benny U, Corby Kistler H, VanEtten HD. Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the

fungal plant pathogen *Nectria haematococca*. *Plant Journal*. 2001;25(3):305-14.

125. Hane J, Oliver R. RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences. *BMC Bioinformatics*. 2008;9:478.

126. Hane J, Williams A, Taranto A, Solomon P, Oliver R. Repeat-induced point mutation: A fungal-specific, endogenous mutagenesis process. In: van den Berg MA, Maruthachalam K, editors. Genetic Transformation Systems in Fungi, Volume 2. Fungal Biology;10.1007/978-3-319-10503-1_4: Springer International Publishing; 2015. p. 55-68.

127. Hane JK, Anderson JP, Williams AH, Sperschneider J, Singh KB. Genome sequencing and comparative genomics of the broad host-range pathogen *Rhizoctonia solani* ag8. *PLoS Genetics*. 2014;10(5):e1004281.

128. Hane JK, Lowe RGT, Solomon PS, Tan K-C, Schoch CL, Spatafora JW, Crous PW, Kodira C, Birren BW, Galagan JE, Torriani SFF, McDonald BA, Oliver RP. Dothideomycete - Plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *Plant Cell*. 2007;19(11):3347-68.

129. Hane JK, Rouxel T, Howlett BJ, Kema GH, Goodwin SB, Oliver RP. A novel mode of chromosomal evolution peculiar to filamentous Ascomycete fungi. *Genome Biology*. 2011;12(5):R45.

130. Hartl L, Zach S, Seidl-Seiboth V. Fungal chitinases: diversity, mechanistic properties and biotechnological potential. *Applied Microbiology and Biotechnology*. 2012;93(2):533-43.

131. Hartman T, Weinrick B, Vilchèze C, Berney M, Tufariello J, Cook GM, Jacobs WR. Succinate dehydrogenase is the regulator of respiration in *Mycobacterium tuberculosis*. *PLoS Pathogens*. 2014;10(11).

132. Hayashi K, Schoonbeek HJ, De Waard MA. Bcmfs1, a novel major facilitator superfamily transporter from *Botrytis cinerea*, provides tolerance towards the natural toxic compounds camptothecin and cercosporin and towards fungicides. *Applied and Environmental Microbiology*. 2002;68(10):4996-5004.

133. He X. Rapid Subfunctionalization Accompanied by Prolonged and Substantial Neofunctionalization in Duplicate Gene Evolution. *Genetics*. 2005;169(2):1157-64.

134. HEATH MC, WOOD RKS. Role of Cell-wall-degrading Enzymes in the Development of Leaf Spots Caused by *Ascochyta pisi* and *Mycosphaerella pinodes*. *Annals of Botany*. 1971;35(2):451-74.

135. Hernandez-Bello MA, Chilvers MI, Akamatsu H, Peever TL. Host specificity of *Ascochyta spp*. infecting legumes of the viciae and cicerae tribes and pathogenicity of an interspecific hybrid. *Phytopathology*. 2006;96(10):1148-56.

136. Herrero E, Ros J, Tamarit J, Bellí G. Glutaredoxins in fungi. *Photosynthesis Research*. 2006;89(2):127-40.

137. Hogenhout SA, Van Der Hoorn RAL, Terauchi R, Kamoun S. Emerging concepts in effector biology of plant-associated organisms. *Molecular Plant-Microbe Interactions*. 2009;22(2):115-22.

138. Horns F, Petit E, Yockteng R, Hood ME. Patterns of repeat-induced point mutation in transposable elements of basidiomycete fungi. *Genome Biology and Evolution*. 2012;4(3):240-7.

139. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. *Nucleic Acids Research*. 2007;35(Web Server issue):W585-7.

140. Howlett BJ, Lowe RG, Marcroft SJ, van de Wouw AP. Evolution of virulence in fungal plant pathogens: exploiting fungal genomics to control plant disease. *Mycologia*. 2015;107(3):441-51.

141. Huang X, Adams M, Zhou H, Kerlavage A. A tool for analyzing and annotating genomic sequences. *Genomics*. 1997;46:37 - 45.

142. Inderbitzin P, Asvarak T, Turgeon BG. Six new genes required for production of T-toxin, a polyketide determinant of high virulence of *Cochliobolus heterostrophus* to maize. *Molecular Plant-Microbe Interactions*. 2010;23(4):458-72.

143. Jashni MK, Dols IH, Iida Y, Boeren S, Beenen HG, Mehrabi R, Collemare J, de Wit PJ. Synergistic Action of a Metalloprotease and a Serine Protease from *Fusarium oxysporum* f. sp. *lycopersici* Cleaves Chitin-Binding Tomato

Chitinases, Reduces Their Antifungal Activity, and Enhances Fungal Virulence. *Molecular Plant-Microbe Interactions*. 2015;28(9):996-1008.

144. Jiang C, He B, Huang R, Huang M, Xu L. Identification and functional analysis of LysM effectors from *Marssonina brunnea*. *Australasian Plant Pathology*. 2014;43(6):615-22.

145. Johnson L. Iron and siderophores in fungal-host interactions. *Mycological Research*. 2008;112(Pt 2):170-83.

146. Jones JDG, Dangl JL. The plant immune system. *Nature*. 2006;444(7117):323-9.

147. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30(9):1236-40.

148. Joubert A, Bataille-Simoneau N, Campion C, Guillemette T, Hudhomme P, Iacomi-Vasilescu B, Leroy T, Pochon S, Poupard P, Simoneau P. Cell wall integrity and high osmolarity glycerol pathways are required for adaptation of *Alternaria brassicicola* to cell wall stress caused by brassicaceous indolic phytoalexins. *Cellular Microbiology*. 2011;13(1):62-80.

149. Kachroo AH, Laurent JM, Yellman CM, Meyer AG, Wilke CO, Marcotte EM. Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. *Science*. 2015;348(6237):921-5.

150. Kaiser WJ. Inter- and intranational spread of ascochyta pathogens of chickpea, faba bean, and lentil. *Canadian Journal of Plant Pathology*. 1997;19(2):215-24.

151. Kaiser WJ, Viruega JR, Peever TL, Trapero A. First report of ascochyta blight outbreak of pea caused by *Ascochyta pisi* in Spain. *Plant Disease*. 2008;92(9):1365.

152. Kaku H, Nishizawa Y, Ishii-Minami N, Akimoto-Tomiyama C, Dohmae N, Takio K, Minami E, Shibuya N. Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences*. 2006;103(29):11086-91.

153. Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology*. 2004;338(5):1027-36.

154. Kamoun S. A catalogue of the effector secretome of plant pathogenic oomycetes. *Annual Review of Phytopathology*. 2006;44:41-60.

155. Kanneganti TD, Huitema E, Cakir C, Kamoun S. Synergistic interactions of the plant cell death pathways induced by *Phytophthora infestans* Nepl-like protein PiNPP1.1 and INF1 elicitin. *Molecular Plant-Microbe Interactions*. 2006;19(8):854-63.

156. Kaur S, Pembleton LW, Cogan NOI, Savin KW, Leonforte T, Paull J, Materne M, Forster JW. Transcriptome sequencing of field pea and faba bean for discovery and validation of SSR genetic markers. *BMC Genomics*. 2012;13(1).

157. Keller NP, Turner G, Bennett JW. Fungal secondary metabolism [mdash] from biochemistry to genomics. *Nature Reviews Microbiology*. 2005;3(12):937-47.

158. Kessie FK. Insights into the Genome and Secretome of *Didymella pinodes*, the causal agent of ascochyta blight of pea. *PhD Thesis submitted to Murdoch University*. 2013:182.

159. Khan R, Straney DC. Regulatory Signals Influencing Expression of the *PDA1* Gene of *Nectria haematococca* MPVI in Culture and During Pathogenesis of Pea. *Molecular Plant-Microbe Interactions*. 1999;12(8):733-42.

160. Khani M, Davidson JA, Sosnowski MR, Scott ES. Survival, transmission and control of Phoma koolunga in field pea seed and reaction of field pea genotypes to the pathogen. *Australasian Plant Pathology*. 2016;45(1):91-102.

161. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*. 2013;14(4):R36.

162. Kimura N, Tsuge T. Gene cluster involved in melanin biosynthesis of the filamentous fungus *Alternaria alternata*. *Journal of Bacteriology*. 1993;175(14):4427-35.

163. King BC, Waxman KD, Nenni NV, Walker LP, Bergstrom GC, Gibson DM. Arsenal of plant cell wall degrading enzymes reflects host preference among plant pathogenic fungi. *Biotechnology for Biofuels*. 2011;4:4.

164. Kinsey GC. *Phoma pinodella*. [Descriptions of Fungi and Bacteria]. *IMI Descriptions of Fungi and Bacteria*. 2002(151):Sheet 1505.

165. Kirsten S, Navarro-Quezada A, Penselin D, Wenzel C, Matern A, Leitner A, Baum T, Seiffert U, Knogge W. Necrosis-inducing proteins of *Rhynchosporium commune*, effectors in quantitative disease resistance. *Molecular Plant-Microbe Interactions*. 2012;25(10):1314-25.

166. Kirzinger MWB, Stavrinides J. Host specificity determinants as a genetic continuum. *Trends in Microbiology*. 2012;20(2):88-93.

167. Kleemann J, Rincon-Rivera LJ, Takahara H, Neumann U, Ver Loren van Themaat E, van der Does HC, Hacquard S, Stuber K, Will I, Schmalenbach W, Schmelzer E, O'Connell RJ. Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen *Colletotrichum higginsianum*. *PLoS Pathogens*. 2012;8(4):e1002643.

168. Klosterman SJ, Subbarao KV, Kang S, Veronese P, Gold SE, Thomma BP, Chen Z, Henrissat B, Lee YH, Park J, Garcia-Pedrajas MD, Barbara DJ, Anchieta A, de Jonge R, Santhanam P, Maruthachalam K, Atallah Z, Amyotte SG, Paz Z, Inderbitzin P, Hayes RJ, Heiman DI, Young S, Zeng Q, Engels R, Galagan J, Cuomo CA, Dobinson KF, Ma LJ. Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. *PLoS Pathogens*. 2011;7(7):e1002137.

169. Koharudin LMI, Viscomi AR, Montanini B, Kershaw MJ, Talbot NJ, Ottonello S, Gronenborn AM. Structure-function analysis of a CVNH-LysM lectin expressed during plant infection by the rice blast fungus *Magnaporthe oryzae*. *Structure*. 2011;19(5):662-74.

170. Koonin E. Orthologs, paralogs, and evolutionary genomics. *Annual Review of Genetics*. 2005;39:309 - 38.

171. Kraft JM, Dunne B, Goulden D, Armstrong S. A search for resistance in peas to *Mycosphaerella pinodes*. *Plant Disease*. 1998;82(2):251-3.

172. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology*. 2001;305(3):567-80.

173. Kulkarni RD, Kelkar HS, Dean RA. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends in Biochemical Sciences*. 2003;28(3):118-21.

174. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. Versatile and open software for comparing large genomes. *Genome Biology*. 2004;5(2).

175. Laluk K, Mengiste T. Necrotroph attacks on plants: wanton destruction or covert extortion? *The Arabidopsis Book / American Society of Plant Biologists*. 2010;8:e0136.

176. Langfelder K, Streibel M, Jahn B, Haase G, Brakhage AA. Biosynthesis of fungal melanins and their importance for human pathogenic fungi. *Fungal Genetics and Biology*. 2003;38(2):143-58.

177. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9(4):357-9.

178. Langner T, Gohre V. Fungal chitinases: function, regulation, and potential roles in plant/pathogen interactions. *Current Genetics*. 2015;10.1007/s00294-015-0530-x.

179. Lara-Márquez A, Zavala-Páramo MG, López-Romero E, Calderón-Cortés N, López-Gómez R, Conejo-Saucedo U, Cano-Camacho H. Cloning and characterization of a pectin lyase gene from *Colletotrichum lindemuthianum* and comparative phylogenetic/structural analyses with genes from phytopathogenic and saprophytic/opportunistic microorganisms. *BMC Microbiology*. 2011;11(1):1-15.

180. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947-8.

181. Lawrence DP, Kroken S, Pryor BM, Arnold AE. Interkingdom gene transfer of a hybrid NPS/PKS from bacteria to filamentous ascomycota. *PLoS ONE*. 2011;6(11).

182. Le May C, Guibert M, Baranger A, Tivoli B. A wide range of cultivated legume species act as alternative hosts for the pea aschochyta blight fungus, *Didymella pinodes*. *Plant Pathology*. 2013.

183. Le May C, Guibert M, Leclerc A, Andrivon D, Tivoli B. A single, plastic population of *Mycosphaerella pinodes* causes ascochyta blight on winter and spring peas (*Pisum sativum*) in France. *Applied and Environmental Microbiology*. 2012;78(23):8431-40.

184. Le May C, Potage G, Andrivon D, Tivoli B, Outreman Y. Plant disease complex: Antagonism and synergism between pathogens of the ascochyta blight complex on pea. *Journal of Phytopathology*. 2009;157(11-12):715-21.

185. Lechner M, FindeiSZ S, Steiner L, Marz M, Stadler P, Prohaska S. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics*. 2011;12(1):124.

186. Lee BN, Kroken S, Chou DY, Robbertse B, Yoder OC, Turgeon BG. Functional analysis of all nonribosomal peptide synthetases in *Cochliobolus heterostrophus* reveals a factor, NPS6, involved in virulence and resistance to oxidative stress. *Eukaryotic Cell*. 2005;4(3):545-55.

187. Lenfant N, Hotelier T, Bourne Y, Marchot P, Chatonnet A. Proteins with an alpha/beta hydrolase fold: Relationships between subfamilies in an ever-growing superfamily. *Chemico-Biological Interactions*. 2013;203(1):266-8.

188. Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A. ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. *Nucleic Acids Research*. 2013;41(Database issue):D423-D9.

189. Lepoivre P. Sensitivity of pea Cultivars to Ascochitine and the possible role of the toxin in the pathogenicity of *Ascochyta pisi* (Lib.). *Journal of Phytopathology*. 1982;103(1):25-34.

190. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078 - 9.

191. Li X, Yan-Qing S, Li-Na J, Xiao-Lei Z, Wen-Chao Y, Wei H, Guang-Fu Y. Succinate Dehydrogenase: An Ideal Target for Fungicide Discovery. Discovery and Synthesis of Crop Protection Products. ACS Symposium Series. 1204: American Chemical Society; 2015. p. 175-94.
192. Li YP, You MP, Finnegan PM, Khan TN, Lanoiselet V, Eyres N, Barbetti MJ. First Report of Black Spot Caused by *Boeremia exigua* var. *exigua* on Field Pea in Australia. *Plant Disease*. 2011;96(1):148-.
193. Li YP, You MP, Khan TN, Finnegan PM, Barbetti MJ. First Report of Phoma herbarum on Field Pea (*Pisum sativum*) in Australia. *Plant Disease*. 2011;95(12):1590-.
194. Lin L, Fang W, Liao X, Wang F, Wei D, St. Leger RJ. The MrCYP52 Cytochrome P450 monoxygenase gene of *Metarhizium robertsii* is important for utilizing insect epicuticular hydrocarbons. *PLoS ONE*. 2011;6(12):e28984.
195. Liu X, Inlow M, VanEtten HD. Expression profiles of pea pathogenicity (PEP) genes in vivo and in vitro, characterization of the flanking regions of the PEP cluster and evidence that the PEP cluster region resulted from horizontal gene transfer in the fungal pathogen *Nectria haematococca*. *Current Genetics*. 2003;44(2):95-103.
196. Liu Z, Faris JD, Oliver RP, Tan K-C, Solomon PS, McDonald MC, McDonald BA, Nunez A, Lu S, Rasmussen JB, Friesen TL. SnTox3 Acts in Effector Triggered Susceptibility to Induce Disease on Wheat Carrying the *Snn3* Gene. *PLoS Pathogens*. 2009;5(9):e1000581.
197. Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R. Fungal effectors and plant susceptibility. Annual Review of Plant Biology2015. p. 513-45.
198. Lorang JM, Sweat TA, Wolpert TJ. Plant disease susceptibility conferred by a “resistance” gene. *Proceedings of the National Academy of Sciences*. 2007;104(37):14861-6.
199. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam

TW, Wang J. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*. 2012;1(1):18.

200. Ma L-J, van der Does H, Borkovich K, Coleman J, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman P, Kang S, Shim W-B, Woloshuk C, Xie X, Xu J-R, Antoniw J, Baker S, Bluhm B, Breakspear A, Brown D, Butchko R, Chapman S, Coulson R, Coutinho P, Danchin E, Diener A, Gale L, Gardiner D, Goff S. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature*. 2010;464:367 - 73.

201. Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim WB, Woloshuk C, Xie X, Xu JR, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RA, Chapman S, Coulson R, Coutinho PM, Danchin EG, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee YH, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature*. 2010;464(7287):367-73.

202. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*. 2011;27(21):2957-63.

203. Maloney A, VanEtten H. A gene from the fungal plant pathogen *Nectria haematococca* that encodes the phytoalexin-detoxifying enzyme pisatin demethylase defines a new cytochrome P450 family. *Molecular Genetics and Genomics*. 1994;243:506 - 14.

204. Manning V, Pandelova I, Dhillon B, Wilhelm L, Goodwin S, Berlin A, Figueroa M, Freitag M, Hane J, Henrissat B. Comparative genomics of a plant-pathogenic fungus, Pyrenophora tritici-repentis, reveals transduplication and the impact of repeat elements on pathogenicity and population divergence. *G3 Genes| Genomes| Genetics*. 2013;3(1):41 - 63.

205. Manning VA, Pandelova I, Dhillon B, Wilhelm LJ, Goodwin SB, Berlin AM, Figueroa M, Freitag M, Hane JK, Henrissat B, Holman WH, Kodira CD,

Martin J, Oliver RP, Robbertse B, Schackwitz W, Schwartz DC, Spatafora JW, Turgeon BG, Yandava C, Young S, Zhou S, Zeng Q, Grigoriev IV, Ma LJ, Ciuffetti LM. Comparative genomics of a plant-pathogenic fungus, *Pyrenophora tritici-repentis*, reveals transduplication and the impact of repeat elements on pathogenicity and population divergence. *G3 (Bethesda)*. 2013;3(1):41-63.

206. Marcet-Houben M, Ballester A-R, de la Fuente B, Harries E, Marcos J, Gonzalez-Candelas L, Gabaldon T. Genome sequence of the necrotrophic fungus *Penicillium digitatum*, the main postharvest pathogen of citrus. *BMC Genomics*. 2012;13(1):646.

207. Marcet-Houben M, Gabaldón T. Beyond the whole-genome duplication: Phylogenetic evidence for an ancient interspecies hybridization in the baker's yeast lineage. *PLoS Biology*. 2015;13(8).

208. Marcinkowska J, Klos B, Shcherbakova A. Ascochitine production by fungi responsible for ascochyta diseases of pea. *Journal of Phytopathology*. 1991;131(3):253-8.

209. Mariani D, Mathias CJ, da Silva CG, Herdeiro RdS, Pereira R, Panek AD, Eleutherio ECA, Pereira MD. Involvement of glutathione transferases, Gtt1 and Gtt2, with oxidative stress response generated by H2O2 during growth of *Saccharomyces cerevisiae Redox Report*. 2008;13(6):246-54.

210. Marshall R, Kombrink A, Motteram J, Loza-Reyes E, Lucas J, Hammond-Kosack KE, Thomma BPHJ, Rudd JJ. Analysis of two in planta expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat. *Plant Physiology*. 2011;156(2):756-69.

211. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*. 2011;17(1).

212. Mathew FM, Goswami RS, Markell SG, Osborne L, Tande C, Ruden B. First report of ascochyta blight of field pea caused by *Ascochyta pisi* in South Dakota. *Plant Disease*. 2010;94(6):789.

213. Matthews DE, Van Etten HD. Detoxification of the phytoalexin pisatin by a fungal cytochrome P-450. *Archives of Biochemistry and Biophysics*. 1983;224(2):494-505.

214. Mayer WE, Schuster LN, Bartelmes G, Dieterich C, Sommer RJ. Horizontal gene transfer of microbial cellulases into nematode genomes is associated with functional assimilation and gene turnover. *BMC Evolutionary Biology*. 2011;11(1).

215. McPhee K. Dry pea production and breeding - a mini review. *Food Agriculture and Environment*. 2003;1(1).

216. Mentlak T, Kombrink A, Shinya T, Ryder L, Otomo I, Saitoh H, Terauchi R, Nishizawa Y, Shibuya N, Thomma B. Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease. *Plant Cell Online*. 2012;24(1):322 - 35.

217. Miguel-Rojas C, Hera C. The F-box protein Fbp1 functions in the invasive growth and cell wall integrity mitogen-activated protein kinase (MAPK) pathways in *Fusarium oxysporum*. *Molecular Plant Pathology*. 2016;17(1):55-64.

218. Milani NA, Lawrence DP, Arnold AE, VanEtten HD. Origin of pisatin demethylase (PDA) in the genus *Fusarium*. *Fungal Genetics and Biology*. 2012;49(11):933-42.

219. Mirzadi Gohari A, Ware SB, Wittenberg AHJ, Mehrabi R, Ben M'Barek S, Verstappen ECP, van der Lee TAJ, Robert O, Schouten HJ, de Wit PPJGM, Kema GHJ. Effector discovery in the fungal wheat pathogen Zymoseptoria tritici. *Molecular Plant Pathology*. 2015;16(9):931-45.

220. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK, Sigrist CJ, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD. The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Research*. 2015;43(Database issue):D213-21.

221. Morkunas I, Formela M, Marczak Ł, Stobiecki M, Bednarski W. The mobilization of defence mechanisms in the early stages of pea seed germination against *Ascochyta pisi*. *Protoplasma*. 2013;250(1):63-75.

222. Motteram J, Kufner I, Deller S, Brunner F, Hammond-Kosack KE, Nurnberger T, Rudd JJ. Molecular characterization and functional analysis of MgNLP, the

sole NPP1 domain-containing protein, from the fungal wheat leaf pathogen *Mycosphaerella graminicola*. *Molecular Plant-Microbe Interactions*. 2009;22(7):790-9.

223. Murray G, Brennan J. The current and potential costs from diseases of pulse crops in Australia. 2012:205.

224. Nafisi M, Stranne M, Zhang L, van Kan JA, Sakuragi Y. The endo-arabinanase BcAra1 is a novel host-specific virulence factor of the necrotic fungal phytopathogen *Botrytis cinerea*. *Molecular Plant-Microbe Interactions*. 2014;27(8):781-92.

225. Nehrt NL, Clark WT, Radivojac P, Hahn MW. Testing the ortholog conjecture with comparative functional genomic data from mammals. *PLoS Computational Biology*. 2011;7(6).

226. Nguyen QB, Itoh K, Van Vu B, Tosa Y, Nakayashiki H. Simultaneous silencing of endo-beta-1,4 xylanase genes reveals their roles in the virulence of *Magnaporthe oryzae*. *Molecular Microbiology*. 2011;81(4):1008-19.

227. O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein L, Alkan N, Altmuller J, Alvarado-Balderrama L, Bauser CA, Becker C, Birren BW, Chen Z, Choi J, Crouch JA, Duvick JP, Farman MA, Gan P, Heiman D, Henrissat B, Howard RJ, Kabbage M, Koch C, Kracher B, Kubo Y, Law AD, Lebrun M-H, Lee Y-H, Miyara I, Moore N, Neumann U, Nordstrom K, Panaccione DG, Panstruga R, Place M, Proctor RH, Prusky D, Rech G, Reinhardt R, Rollins JA, Rounsley S, Schardl CL, Schwartz DC, Shenoy N, Shirasu K, Sikhakolli UR, Stuber K, Sukno SA, Sweigard JA, Takano Y, Takahara H, Trail F, van der Does HC, Voll LM, Will I, Young S, Zeng Q, Zhang J, Zhou S, Dickman MB, Schulze-Lefert P, Ver Loren van Themaat E, Ma L-J, Vaillancourt LJ. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics*. 2012;44(9):1060-5.

228. O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein L, Alkan N, Altmüller J, Alvarado-Balderrama L, Bauser CA, Becker C, Birren BW, Chen Z, Choi J, Crouch JA, Duvick JP, Farman MA, Gan P, Heiman D, Henrissat B, Howard RJ, Kabbage M, Koch C, Kracher B, Kubo Y, Law AD, Lebrun MH, Lee YH, Miyara I, Moore N,

Neumann U, Nordström K, Panaccione DG, Panstruga R, Place M, Proctor RH, Prusky D, Rech G, Reinhardt R, Rollins JA, Rounsley S, Schardl CL, Schwartz DC, Shenoy N, Shirasu K, Sikhakolli UR, Stüber K, Sukno SA, Sweigard JA, Takano Y, Takahara H, Trail F, Van Der Does HC, Voll LM, Will I, Young S, Zeng Q, Zhang J, Zhou S, Dickman MB, Schulze-Lefert P, Ver Loren Van Themaat E, Ma LJ, Vaillancourt LJ. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics*. 2012;44(9):1060-5.

229. Ohm RA, Feau N, Henrissat B, Schoch CL, Horwitz BA, Barry KW, Condon BJ, Copeland AC, Dhillon B, Glaser F, Hesse CN, Kosti I, LaButti K, Lindquist EA, Lucas S, Salamov AA, Bradshaw RE, Ciuffetti L, Hamelin RC, Kema GH, Lawrence C, Scott JA, Spatafora JW, Turgeon BG, de Wit PJ, Zhong S, Goodwin SB, Grigoriev IV. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathogens*. 2012;8(12):e1003037.
230. Oide S, Moeder W, Krasnoff S, Gibson D, Haas H, Yoshioka K, Turgeon BG. NPS6, encoding a nonribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a conserved virulence determinant of plant pathogenic ascomycetes. *Plant Cell*. 2006;18(10):2836-53.
231. Oliver RP, Solomon Peter S PS. New developments in pathogenicity and virulence of necrotrophs. *Current Opinion in Plant Biology*. 2010;13(4):415-9.
232. Onfroy C, Tivoli B, Corbière R, Bouznad Z. Cultural, molecular and pathogenic variability of *Mycosphaerella pinodes* and *Phoma medicaginis* var. *pinodella* isolates from dried pea (*Pisum sativum*) in France. *Plant Pathology*. 1999;48(2):218-29.
233. Oome S, Van Den Ackerveken G. Comparative and functional analysis of the widely occurring family of Nep1-like proteins. *Molecular Plant-Microbe Interactions*. 2014;27(10):1081-94.
234. Ottmann C, Luberacki B, Kufner I, Koch W, Brunner F, Weyand M, Mattinen L, Pirhonen M, Anderluh G, Seitz HU, Nurnberger T, Oecking C. A common toxin fold mediates microbial attack and plant defense. *Proceedings of the National Academy for Sciences*. 2009;106(25):10359-64.

235. Parra G, Bradnam K, Korf I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*. 2007;23(9):1061-7.

236. Pedras M, Chumala P, Jin W, Islam M, Hauck D. The phytopathogenic fungus *Alternaria brassicicola*: Phytotoxin production and phytoalexin elicitation. *Phytochemistry*. 2009;70:394 - 402.

237. Pedras MS, Abdoli A. Metabolism of the phytoalexins camalexins, their bioisosteres and analogues in the plant pathogenic fungus *Alternaria brassicicola*. *Bioorganic and Medicinal Chemistry*. 2013;21(15):4541-9.

238. Pedras MS, Yaya EE, Glawischnig E. The phytoalexins from cultivated and wild crucifers: chemistry and biology. *Natural Product Report*. 2011;28(8):1381-405.

239. Pedras MSC, Minic Z. Differential protein expression in response to the phytoalexin brassinin allows the identification of molecular targets in the phytopathogenic fungus *Alternaria brassicicola*. *Molecular Plant Pathology*. 2012;13(5):483-93.

240. Peever TL, Barve MP, Stone LJ, Kaiser WJ. Evolutionary relationships among *Ascochyta* species infecting wild and cultivated hosts in the legume tribes Cicereae and Vicieae. *Mycologia*. 2007;99(1):59-77.

241. Pemberton CL, Salmond GP. The Nep1-like proteins-a growing family of microbial elicitors of plant necrosis. *Molecular Plant Pathology*. 2004;5(4):353-9.

242. Percudani R, Montanini B, Ottonello S. The anti-HIV cyanovirin-N domain is evolutionarily conserved and occurs as a protein module in eukaryotes. *Proteins: Structure, Function and Genetics*. 2005;60(4):670-8.

243. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*. 2011;8(10):785-6.

244. Pietro AD, Huertas-González MD, Gutierrez-Corona JF, Martínez-Cadena G, Méglecz E, Roncero MIG. Molecular Characterization of a Subtilase from the Vascular Wilt Fungus *Fusarium oxysporum*. *Molecular Plant-Microbe Interactions*. 2001;14(5):653-62.

245. Poppe S, Dorsheimer L, Happel P, Stukenbrock EH. Rapidly evolving genes are key players in host specialization and virulence of the fungal wheat

pathogen *Zymoseptoria tritici* (*Mycosphaerella graminicola*). *PLoS Pathogens*. 2015;11(7).

246. Poulin R, Mouillot D. Combining phylogenetic and ecological information into a new index of host specificity. *Journal of Parasitology*. 2005;91(3):511-4.

247. Poussereau N, Gente S, Rascle C, Billon-Grand G, Fevre M. aspS encoding an unusual aspartyl protease from *Sclerotinia sclerotiorum* is expressed during phytopathogenesis. *FEMS Microbiology Letters*. 2001;194(1):27-32.

248. Prioul S, Frankewitz A, Deniot G, Morin G, Baranger A. Mapping of quantitative trait loci for partial resistance to *Mycosphaerella pinodes* in pea (*Pisum sativum* L.), at the seedling and adult plant stages. *Theoretical and Applied Genetics*. 2004;108(7):1322-34.

249. Punithalingam E, Holliday P. *Mycosphaerella pinodes*. [Descriptions of Fungi and Bacteria]. *IMI Descriptions of Fungi and Bacteria*. 1972(34):Sheet 340.

250. Punta M, Coggill P, Eberhardt R, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer E, Eddy S, Bateman A, Finn R. The Pfam protein families database. *Nucleic Acids Research*. 2012;40(D1):D290 - D301.

251. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-2.

252. Qutob D, Kamoun S, Gijzen M. Expression of a *Phytophthora sojae* necrosis-inducing protein occurs during transition from biotrophy to necrotrophy. *Plant Journal*. 2002;32(3):361-73.

253. Qutob D, Kemmerling B, Brunner F, Kufner I, Engelhardt S, Gust AA, Luberacki B, Seitz HU, Stahl D, Rauhut T, Glawischnig E, Schween G, Lacombe B, Watanabe N, Lam E, Schlichting R, Scheel D, Nau K, Dodt G, Hubert D, Gijzen M, Nurnberger T. Phytotoxicity and innate immune responses induced by Nep1-like proteins. *Plant Cell*. 2006;18(12):3721-44.

254. Rafiqi M, Ellis J, Ludowici V, Hardham A, Dodds P. Challenges and progress towards understanding the role of effectors in plant-fungal interactions. *Current Opinon in Plant Biology*. 2012;15(4):477 - 82.

255. Rao Uppalapati S, Toyoda K, Yasuhiro I, Ichinose Y, Shiraishi T. Differential regulation of MBP kinases by a glycoproptein elicitor and a polypeptide suppressor from *Mycosphaerella pinodes* in pea. *Physiological and Molecular Plant Pathology*. 2004;64(1):17-25.

256. Reis H, Pfiff S, Hahn M. Molecular and functional characterization of a secreted lipase from *Botrytis cinerea*. *Molecular Plant Pathology*. 2005;6(3):257-67.

257. Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. DOG 1.0: illustrator of protein domain structures. *Cell Research*. 2009;19(2):271-3.

258. Rep M. Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbiology Letters*. 2005;253(1):19-27.

259. Richards TA, Soanes DM, Jones MDM, Vasieva O, Leonard G, Paszkiewicz K, Foster PG, Hall N, Talbot NJ. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(37):15258-63.

260. Rio DC, Ares M, Jr., Hannon GJ, Nilsen TW. Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb Protocols*. 2010;2010(6):pdb.prot5439.

261. Roger C, Tivoli B. Spatio-temporal development of pycnidia and perithecia and dissemination of spores of *Mycosphaerella pinodes* on pea (*Pisum sativum*). *Plant Pathology*. 1996;45(3):518-28.

262. Roger C, Tivoli B, Huber L. Effects of temperature and moisture on disease and fruit body development of *Mycosphaerella pinodes* on pea (*Pisum sativum*). *Plant Pathology*. 1999;48(1):1-9.

263. Rogers LM, Kim YK, Guo W, Gonzalez-Candelas L, Li D, Kolattukudy PE. Requirement for either a host- or pectin-induced pectate lyase for infection of *Pisum sativum* by Nectria hematococca. *Proceedings of the National Academy of Sciences*. 2000;97(17):9813-8.

264. Roohparvar R, De Waard MA, Kema GH, Zwiers LH. MgMfs1, a major facilitator superfamily transporter from the fungal wheat pathogen *Mycosphaerella graminicola*, is a strong protectant against natural toxic

compounds and fungicides. *Fungal Genetics and Biology*. 2007;44(5):378-88.

265. Roth Z, Yehezkel G, Khalaila I. Identification and quantification of protein glycosylation. *International Journal of Carbohydrate Chemistry*. 2012;2012:1-10.
266. Rouxel T, Grandaubert J, Hane J, Hoede C, van de Wouw A, Couloux A, Dominguez V, Anthouard V, Bally P, Bourras S, Cozijnsen A, Ciuffetti L, Degrave A, Dilmaghani A, Duret L, Fudal I, Goodwin S, Gout L, Glaser N, Linglin J, Kema G, Lapalu N, Lawrence C, May K, Meyer M, Ollivier B, Poulain J, Schoch C, Simon A, Spatafora J. Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point Mutations. *Nature Communication*. 2011;2:202.
267. Rovenich H, Boshoven JC, Thomma BPHJ. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. *Current Opinion in Plant Biology*. 2014;20:96-103.
268. Saha D, Fetzner R, Burkhardt B, Podlech J, Metzler M, Dang H, Lawrence C, Fischer R. Identification of a polyketide synthase required for alternariol (AOH) and alternariol-9-methyl ether (AME) formation in *Alternaria alternata*. *PLoS ONE*. 2012;7(7):e40564.
269. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution.* 1987;4(4):406-25.
270. Salam MU, MacLeod WJ, Salam KP, Maling T, Barbetti MJ. Impact of climate change in relation to ascochyta blight on field pea in Western Australia. *Australasian Plant Pathology*. 2011;40(4):397-406.
271. Sanders IR. Rapid disease emergence through horizontal gene transfer between eukaryotes. *Trends in Ecology & Evolution*. 2006;21(12):656-8.
272. Santhanam P, van Esse HP, Albert I, Faino L, Nurnberger T, Thomma BP. Evidence for functional diversification within a fungal NEP1-like protein family. *Molecular Plant-Microbe Interactions.* 2013;26(3):278-86.
273. Sarquis MI, Oliveira EM, Santos AS, Costa GL. Production of L-asparaginase by filamentous fungi. *Memorias do Instituto Oswaldo Cruz.* 2004;99(5):489-92.

274. Saunders D, Win J, Cano L, Szabo L, Kamoun S, Raffaele S. Using Hierarchical Clustering of Secreted Protein Families to Classify and Rank Candidate Effectors of Rust Fungi. *PLoS ONE*. 2012;7(1):e29847.

275. Saupe SJ. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews*. 2000;64(3):489-502.

276. Schouten A, Van Baarlen P, Van Kan JAL. Phytotoxic Nep1-like proteins from the necrotrophic fungus *Botrytis cinerea* associate with membranes and the nucleus of plant cells. *New Phytologist*. 2008;177(2):493-505.

277. Schurch S, Linde CC, Knogge W, Jackson LF, McDonald BA. Molecular population genetic analysis differentiates two virulence mechanisms of the fungal avirulence gene NIP1. *Molecular Plant-Microbe Interactions*. 2004;17(10):1114-25.

278. Seidl V. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. *Fungal Biology Reviews*. 2008;22(1):36-42.

279. Selker EU. Repeat-Induced Gene Silencing in Fungi. *Advances in Genetics*. 2002;46:12.

280. Setti B, Bencheikh M, Henni J, Neema C. Comparative aggressiveness of *Mycosphaerella pinodes* on peas from different regions in western Algeria. *Phytopathologia Mediterranea*. 2009;48(2):195-204.

281. Sexton AC, Howlett BJ. Parallels in fungal pathogenesis on plant and animal hosts. *Eukaryotic Cell*. 2006;5(12):1941-9.

282. Shi L, Li R, Liao S, Bai L, Lu Q, Chen B. Prb1, a subtilisin-like protease, is required for virulence and phenotypical traits in the chestnut blight fungus. *FEMS Microbiology Letters*. 2014;359(1):26-33.

283. Shibuya N, Minami E. Oligosaccharide signalling for defence responses in plant. *Physiological and Molecular Plant Pathology*. 2001;59(5):223-33.

284. Shiraishi T. Suppression of defense response related to plant cell wall. *Japan Agricultural Research Quarterly*. 2013;47(1):21-7.

285. Shiraishi T, Saitoh K, Kim HM, Kato T, Tahara M, Oku H, Yamada T, Ichinose Y. Two suppressors, supprescins A and B, secreted by a pea pathogen, *Mycosphaerella pinodes*. *Plant and Cell Physiology*. 1992;33(5):663-7.

286. Shiraishi T, Yamada T, Ichinose Y, Kiba A, Toyoda K. The role of suppressors in determining host-parasite specificities in plant cells. In: Kwang WJ, editor. International Review of Cytology. Volume 172: Academic Press; 1997. p. 55-93.

287. Sindhu A, Ramsay L, Sanderson LA, Stonehouse R, Li R, Condie J, Shunmugam AS, Liu Y, Jha AB, Diapari M, Burstin J, Aubert G, Tar'an B, Bett KE, Warkentin TD, Sharpe AG. Gene-based SNP discovery and genetic mapping in pea. *Theoretical and Applied Genetics*. 2014;127(10):2225-41.

288. Skamnioti P, Furlong RF, Gurr SJ. Evolutionary history of the ancient cutinase family in five filamentous ascomycetes reveals differential gene duplications and losses and in *Magnaporthe grisea* shows evidence of sub- and neo-functionalization. *New Phytologist*. 2008;180(3):711-21.

289. Skamnioti P, Gurr SJ. *Magnaporthe grisea* cutinase2 mediates appressorium differentiation and host penetration and is required for full virulence. *Plant Cell*. 2007;19(8):2674-89.

290. Sperschneider J, Dodds PN, Gardiner DM, Manners JM, Singh KB, Taylor JM. Advances and challenges in computational prediction of effectors from plant pathogenic fungi. *PLoS Pathogens*. 2015;11(5):e1004806.

291. Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, Manners JM, Taylor JM. EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytologist*. 2015;10.1111/nph.13794.

292. Sperschneider J, Ying H, Dodds PN, Gardiner DM, Upadhyaya NM, Singh KB, Manners JM, Taylor JM. Diversifying selection in the wheat stem rust fungus acts predominantly on pathogen-associated gene families and reveals candidate effectors. *Frontiers in Plant Science*. 2014;5:372.

293. Staats M, P VANB, Schouten A, JA VANK. Functional analysis of NLP genes from *Botrytis elliptica*. *Molecular Plant Pathology*. 2007;8(2):209-14.

294. Stefanato FL, Abou-Mansour E, Buchala A, Kretschmer M, Mosbach A, Hahn M, Bochet CG, Metraux JP, Schoonbeek HJ. The ABC transporter BcatrB from *Botrytis cinerea* exports camalexin and is a virulence factor on Arabidopsis thaliana. *Plant Journal*. 2009;58(3):499-510.

295. Stergiopoulos I, de Wit PJ. Fungal effector proteins. *Annual Review of Phytopathology*. 2009;47:233-63.

296. Stergiopoulos I, Van Den Burg HA, Ökmen B, Beenen HG, Van Liere S, Kema GHJ, De Wit PJGM. Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;107(16):7610-5.

297. Stergiopoulos I, Zwiers LH, De Waard MA. The ABC transporter MgAtr4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of substomatal cavities in wheat leaves. *Molecular Plant-Microbe Interactions*. 2003;16(8):689-98.

298. Stukenbrock EH, Bataillon T. A population genomics perspective on the emergence and adaptation of new plant pathogens in agro-ecosystems. *PLoS Pathogens*. 2012;8(9):e1002893.

299. Stukenbrock EH, Jørgensen FG, Zala M, Hansen TT, McDonald BA, Schierup MH. Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen mycosphaerella graminicola. *PLoS Genetics*. 2010;6(12):1-13.

300. Su H, Hwang SF, Chang KF, Conner RL, Howard RJ, Turnbull GD. Variations in virulence and genome of *Mycosphaerella pinodes* isolates from field peas in central Alberta, Canada. *Journal of Plant Diseases and Protection*. 2006;113(2):53-60.

301. Sun BF, Xiao JH, He S, Liu L, Murphy RW, Huang DW. Multiple interkingdom horizontal gene transfers in *Pyrenophora* and closely related species and their contributions to phytopathogenic lifestyles. *PLoS ONE*. 2013;8(3).

302. Swanson WJ, Yang Z, Wolfner MF, Aquadro CF. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proceedings of the National Academy of Sciences*. 2001;98(5):2509-14.

303. Sweigard JA, Carroll AM, Farrall L, Chumley FG, Valent B. *Magnaporthe grisea* pathogenicity genes obtained through insertional mutagenesis. *Molecular Plant-Microbe Interactions*. 1998;11(5):404-12.

304. Sweigard JA, Matthews DE, Vanetten HD. Synthesis of the phytoalexin pisatin by a methyltransferase from pea. *Plant Physiology*. 1986;80(1):277-9.

305. Syme RA, Hane JK, Friesen TL, Oliver RP. Resequencing and Comparative Genomics of *Stagonospora nodorum*: Sectional Gene Absence and Effector Discovery. *G3*. 2013;3(6):959-69.

306. Taborda CP, da Silva MB, Nosanchuk JD, Travassos LR. Melanin as a virulence factor of *Paracoccidioides brasiliensis* and other dimorphic pathogenic fungi: a minireview. *Mycopathologia*. 2008;165(4-5):331-9.

307. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution.* 2013;30

308. Tan KC, Oliver RP, Solomon PS, Moffat CS. Proteinaceous necrotrophic effectors in fungal virulence. *Functional Plant Biology*. 2010;37(10):907-12.

309. Temporini ED, Vanetten HD. An analysis of the phylogenetic distribution of the pea pathogenicity genes of *Nectria haematococca* MPVI supports the hypothesis of their origin by horizontal transfer and uncovers a potentially new pathogen of garden pea: Neocosmospora boniensis. *Current Genetics*. 2004;46(1):29-36.

310. ten Have A, Dekkers E, Kay J, Phylip LH, van Kan JA. An aspartic proteinase gene family in the filamentous fungus *Botrytis cinerea* contains members with novel features. *Microbiology*. 2004;150(Pt 7):2475-89.

311. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research*. 2008;18(12):1979-90.

312. Tian M, Win J, Song J, Van Der Hoorn R, Van Der Knaap E, Kamoun S. A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease. *Plant Physiology*. 2007;143(1):364-77.

313. Timmerman-Vaughan GM, Frew TJ, Russell AC, Khan T, Butler R, Gilpin M, Murray S, Falloon K. QTL mapping of partial resistance to field epidemics of Ascochyta blight of pea. *Crop Science*. 2002;42(6):2100-11.

314. Timmerman-Vaughan GM, T. J. Frew, R. Butler, S. Murray, M. Gilpin, K. Falloon, P.Johnston, M. B. Lakeman, A.Russell and T. Khan. Validation of

quantitative trait loci for Ascochyta blight resistance in pea (*Pisum sativum* L.), using populations from two crosses. *Theoretical Applied Genetics*. 2004;109:1620-31.

315. Tivoli B, Banniza S. Comparison of the epidemiology of ascochyta blights on grain legumes. *European Journal of Plant Pathology*. 2007;119(1):59-76.
316. Tivoli B, Baranger A, Avila CM, Banniza S, Barbetti M, Chen W, Davidson J, Lindeck K, Kharrat M, Rubiales D, Sadiki M, Sillero JC, Sweetingham M, Muehlbauer FJ. Screening techniques and sources of resistance to foliar diseases caused by major necrotrophic fungi in grain legumes. *Euphytica*. 2006;147(1-2):223-53.
317. Tony Fischer, Derek Byerlee, Edmeades aG. Crop yields and global food security. 2014:634.
318. Toyoda K, Ikeda S, Morikawa J, Hirose M, Maeda A, Suzuki T, Inagaki Y, Ichinose Y, Shiraishi T. The *Medicago truncatula* - *Mycosphaerella pinodes* interaction: A new pathosystem for dissecting fungal-suppressor-mediated disease susceptibility in plants. *Journal of General Plant Pathology*. 2013;79(1):1-11.
319. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*. 2012;7(3):562-78.
320. Tsuge T, Harimoto Y, Akimitsu K, Ohtani K, Kodama M, Akagi Y, Egusa M, Yamamoto M, Otani H. Host-selective toxins produced by the plant pathogenic fungus *Alternaria alternata*. *FEMS Microbiology Reviews*. 2013;37(1):44-66.
321. Van de Wouw A, Cozijnsen A, Hane J, Brunner P, McDonald B, Oliver R, Howlett B. Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants. *PLoS Pathogens*. 2010;6:e1001180.
322. van Kan JAL. Licensed to kill: the lifestyle of a necrotrophic plant pathogen. *Trends in Plant Science*. 2006;11(5):247-53.
323. Varki A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*. 1993;3(2):97-130.

324. Vleeshouwers VG, Oliver RP. Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *Molecular Plant-Microbe Interactions*. 2014;27(3):196-206.

325. Vogel J. Unique aspects of the grass cell wall. *Current Opinion in Plant Biology*. 2008;11(3):301-7.

326. Voigt K, Cozijnsen A, Kroymann J, Poggeler S, Howlett B. Phylogenetic relationships between members of the crucifer pathogenic *Leptosphaeria maculans* species complex as shown by mating type (MAT1-2), actin, and beta-tubulin sequences. *Molecular Phylogenetics and Evolution*. 2005;37:541 - 57.

327. Wagener J, Weindl G, de Groot PWJ, de Boer AD, Kaesler S, Thavaraj S, Bader O, Mailänder-Sanchez D, Borelli C, Weig M, Biedermann T, Naglik JR, Korting HC, Schaller M. Glycosylation of <italic>Candida albicans</italic> Cell Wall Proteins Is Critical for Induction of Innate Immune Responses and Apoptosis of Epithelial Cells. *PLoS ONE*. 2012;7(11):e50518.

328. Wan J, Zhang X-C, Stacey G. Chitin signaling and plant disease resistance. *Plant Signaling & Behavior*. 2008;3(10):831-3.

329. Wang J, Chu S, Zhu Y, Cheng H, Yu D. Positive selection drives neofunctionalization of the UbiA prenyltransferase gene family. *Plant Molecular Biology* 2015;87(4-5):383-94.

330. Wang Y, Kim JA, Cheong YH, Koh YJ, Hur J-S. Isolation and characterization of a non-reducing polyketide synthase gene from the lichen-forming fungus *Usnea longissima*. *Mycological Progress*. 2011;11(1):75-83.

331. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Journal of Biotechnology*. 2001;18(5):691-9.

332. Wibberg D, Jelonek L, Rupp O, Hennig M, Eikmeyer F, Goesmann A, Hartmann A, Borriss R, Grosch R, Puhler A, Schluter A. Establishment and interpretation of the genome sequence of the phytopathogenic fungus *Rhizoctonia solani* AG1-IB isolate 7/3/14. *Nigerian Journal of Biotechnology*. 2013;167(2):142-55.

333. Wilson K. Preparation of genomic DNA from bacteria. Current Protocols in Molecular Biology;10.1002/0471142727.mb0204s56: John Wiley & Sons, Inc.; 1997.

334. Winnenburg R, Baldwin TK, Urban M, Rawlings C, Kohler J, Hammond-Kosack KE. PHI-base: a new database for pathogen host interactions. *Nucleic Acids Research*. 2006;34(Database issue):D459-64.

335. Woudenberg JH, De Gruyter J, Crous PW, Zwiers LH. Analysis of the mating-type loci of co-occurring and phylogenetically related species of *Ascochyta* and *Phoma*. *Molecular Plant Pathology*. 2012;13(4):350-62.

336. Wroth JM. Variation in pathogenicity among and within *Mycosphaerella pinodes* populations collected from field pea in Australia. *Canadian Journal of Botany*. 1998;76(11):1955-66.

337. Wu Q, VanEtten HD. Introduction of plant and fungal genes into pea (*Pisum sativum* L.) hairy roots reduces their ability to produce pisatin and affects their response to a fungal pathogen. *Molecular Plant-Microbe Interactions*. 2004;17(7):798-804.

338. Xue AG, Warkentin TD, Gossen BD, Burnett PA, Vandenberg A, Rashid KY. Pathogenic variation of western Canadian isolates of *Mycosphaerella pinodes* on selected *Pisum* genotypes. *Canadian Journal of Plant Pathology*. 1998;20(2):189-93.

339. Xue AG, Warkentin TD, Kenaschuk EO. Effects of timings of inoculation with *Mycosphaerella pinodes* on yield and seed infection of field pea. *Canadian Journal of Plant Science*. 1997;77(4):685-9.

340. Xue M, Yang J, Li Z, Hu S, Yao N, Dean RA, Zhao W, Shen M, Zhang H, Li C, Liu L, Cao L, Xu X, Xing Y, Hsiang T, Zhang Z, Xu J-R, Peng Y-L. Comparative Analysis of the Genomes of Two Field Isolates of the Rice Blast Fungus *Magnaporthe oryzae*. *PLoS Genetics*. 2012;8(8):e1002869.

341. Yakoby N, Beno-Moualem D, Keen NT, Dinoor A, Pines O, Prusky D. *Colletotrichum gloeosporioides* pelB is an important virulence factor in avocado fruit-fungus interaction. *Molecular Plant-Microbe Interactions*. 2001;14(8):988-95.

342. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*. 2007;24(8):1586-91.

343. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*. 2012;40(Web Server issue):W445-51.

344. Yoder MD, Keen NT, Jurnak F. New domain motif: the structure of pectate lyase C, a secreted plant virulence factor. *Science*. 1993;260(5113):1503-7.

345. Yun SH, Turgeon BG, Yoder OC. REMI-induced mutants of *Mycosphaerella zeae-maydis* lacking the polyketide PM-toxin are deficient in pathogenesis to corn. *Physiological and Molecular Plant Pathology*. 1998;52(1):53-66.

346. Zaparoli G, Barsottini MR, de Oliveira JF, Dyszy F, Teixeira PJ, Barau JG, Garcia O, Costa-Filho AJ, Ambrosio AL, Pereira GA, Dias SM. The crystal structure of necrosis- and ethylene-inducing protein 2 from the causal agent of cacao's Witches' Broom disease reveals key elements for its activity. *Biochemistry*. 2011;50(45):9901-10.

347. Zaparoli G, Cabrera OG, Medrano FJ, Tiburcio R, Lacerda G, Pereira GG. Identification of a second family of genes in *Moniliophthora perniciosa*, the causal agent of witches' broom disease in cacao, encoding necrosis-inducing proteins similar to cerato-platanins. *Mycological Research*. 2009;113(Pt 1):61-72.

348. Zhang H, Gao S, Lercher MJ, Hu S, Chen W-H. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Research*. 2012;10.1093/nar/gks576.

349. Zhang J, Lu H, Li X, Li Y, Cui H, Wen CK, Tang X, Su Z, Zhou JM. Effector-triggered and pathogen-associated molecular pattern-triggered immunity differentially contribute to basal resistance to Pseudomonas syringae. *Molecular Plant-Microbe Interactions*. 2010;23(7):940-8.

350. Zhang JX, Fernando WGD, Xue AG. Virulence and Genetic Variability among Isolates of *Mycosphaerella pinodes*. *Plant Disease*. 2003;87(11):1376-83.

351. Zhang R, Hwang SF, Chang KF, Gossen BD, Strelkov SE, Turnbull GD, Blade SF. Genetic resistance to *Mycosphaerella pinodes* in 558 field pea accessions. *Crop Science*. 2006;46(6):2409-14.

352. Zhang XW, Jia LJ, Zhang Y, Jiang G, Li X, Zhang D, Tang WH. *In planta* stage-specific fungal gene profiling elucidates the molecular strategies of

Fusarium graminearum growing inside wheat coleoptiles. *Plant Cell*. 2013;24(12):5159-76.

353. Zhang Z-N, Wu Q-Y, Zhang G-Z, Zhu Y-Y, Murphy RW, Liu Z, Zou C-G. Systematic analyses reveal uniqueness and origin of the CFEM domain in fungi. *Scientific Reports*. 2015;5:13032.

354. Zhou BJ, Jia PS, Gao F, Guo HS. Molecular characterization and functional analysis of a necrosis- and ethylene-inducing, protein-encoding gene family from Verticillium dahliae. *Molecular Plant-Microbe Interactions*. 2012;25(7):964-75.