AN INVESTIGATION OF THE TRANSCRIPTIONAL DYNAMICS DURING THE *PSEUDOPERONOSPORA CUBENSIS* – *CUCUMIS SATIVUS* INTERACTION

By

Alyssa Kay Burkhardt

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Cell and Molecular Biology – Doctor of Philosophy

2015

ABSTRACT

AN INVESTIGATION OF THE TRANSCRIPTIONAL DYNAMICS DURING THE *PSEUDOPERONOSPORA CUBENSIS* – *CUCUMIS SATIVUS* INTERACTION

By

Alyssa Kay Burkhardt

Downy mildew of cucumber (*Cucumis sativus*) is caused by the obligate oomycete, *Pseudoperonospora cubensis,* and the research described in the dissertation provides new insight on the transcriptional regulation within the pathogen through the mechanism of alternative splicing and investigates the transcriptional changes of the host genes within a resistant and susceptible interaction. Previously, the genome of cucumber and *P. cubensis* as well transcriptome of a compatible interaction between the plant and the pathogen were sequenced. In addition, one gene from *P. cubensis* was known to be alternatively spliced, but the breadth of alternative splicing across the transcriptome was unknown. Through the work described in this thesis, we investigated the breadth of alternative splicing across the entire transcriptome of *P. cubensis* over the time course of infection. We found *P. cubensis* genes are frequently spliced and have intron retention as the most common mechanism of alternative splicing with some evidence for the retention of the 5' or 3' end of the exon but no evidence for exon skipping. Furthermore, we found that alternative splicing occurred in genes encoding several types of proteins, including the effectors, which impact pathogenicity and virulence. In some cases, the frequency of alternative splicing was found to correlate with developmental stages of the pathogen and thus alternative splicing might play a role in regulating transcript abundance and availability during development.

Advances in sequencing and bioinformatics also contributed to our work in advancing the knowledge of the transcriptomic changes in a resistant (PI 197088) or susceptible (Vlaspik) plant during an infection time course. Our work shows that while *P. cubensis* is able to enter the resistant plant leaf, it is not able to sporulate; in contrast, the pathogen is able to grow and sporulate in the susceptible host, Vlaspik. To investigate the transcriptional changes underlying resistance, we identified the differentially expressed genes between the resistant and susceptible plant lines over the time course of infection. We found that the resistant plant responded earlier to the pathogen, as demonstrated by a higher number of differentially expressed genes at earlier time points compared to the susceptible plant. In addition, we found changes in genes encoding proteins with functions in hormone-related processes, nutrition, and transportation that might indicate a role for some of these genes in initiating or responding to the resistance response in cucumber. Beyond identifying differentially expressed transcripts, we identified small RNAs in the host and pathogen, as small RNAs have a role in modifying gene expression. We found novel miRNAs in the pathogen and known and novel miRNA in the host and predicted potential targets for each miRNA within the cucumber transcriptome. Some of these miRNAs may have a role in mediating the response of the plant to the pathogen. In the future, work will be done to validate the roles of candidate resistance-associated genes and to validate the presence and role of miRNA in both the host and the pathogen. Some of this future work will involve incorporating other "omics" methods including metabolomics and proteomics in order to get a more complete understanding of the molecular changes in the plant during infection. Finally, strong candidates for resistance could be validated using the proposed *in planta* methods, which includes the development of transgenic cucumbers.

I would like to dedicate this dissertation to my loving and supportive husband, Kannan Raghunathan and to my parents, John and Patty Burkhardt, who have always encouraged my education and have supported me in everything.

ACKNOWLEDGMENTS

I would like to acknowledge my advisor, Dr. Brad Day for his support and guidance throughout the last five years.

I would also like to thank my committee members, Dr. Robin Buell, Dr. Rebecca Grumet, Dr. Susanne Hoffmann-Benning, and Dr. Eran Andrechek for their support and advice throughout my dissertation research. I would also like to thank Dr. Shin-Han Shiu for teaching me bioinformatics and Dr. Kevin Childs for bioinformatics support and collaboration. Finally, I would like to thank Dr. Diane Ebert-May, Dr. Eva Farre, Dr. Rique Campa, and the Future Academic Scholars in Teaching program for their support of my teaching development.

I would like to thank Michigan State University for supporting me through a University Distinguished Fellowship and for research support through Project GREEEN. I would also like to thank the Pickle Packers International and the United States Department of Agriculture for funding this research.

I would like to thank all of my lab mates who have supported me throughout the years including Caleb Knepper, Elizabeth Savory, Katie Porter, Patricia Santos, Masaki Shimono, Alex Corrion, Yi-Ju Lu, Mieder Palm-Forster, and Phil Colgan. Finally and most importantly, I would like to thank God and my church family at Martin Luther Chapel in East Lansing, all of whom have provided unwavering support and emotional stability.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

A genomics perspective on cucurbit-oomycete interactions

This review was previously published, in part, in:

Burkhardt, A and Day, B. 2013. A genomics perspective on cucurbit-oomycete interactions. Plant Biotechnology 30:265-271.

Abstract

Pseudoperonospora cubensis is an obligate plant pathogenic oomycete that is the causative agent of cucumber downy mildew. Like *Phytophthora capsici,* a hemibiotrophic oomycete infecting cucumber, this pathogen is a severe threat to cucurbit cultivation because of its global distribution, broad host range among the *Cucurbitaceae* family, and its ability to overcome susceptibilities to host, environment, and chemical management. Like other downy mildews, *P. cubensis* is an obligate pathogen and gains all of its nutrition from its host without killing it, and as a result, is a challenging pathogen to study due to its recalcitrance to artificial culturing. In recent years, advancements have been made in investigating sources of resistance to these oomycetes, and several mechanisms of resistance have been identified in different host-pathogen systems that range in inheritance patterns from a single gene to quantitative trait loci. However, the specific genes and pathways involved in cucumber resistance to *P. cubensis* have not been identified. Aiding in this research is the recent advent of next-generation sequencing, which has enabled genome, transcriptome, and small RNA sequencing of several non-model hosts and pathogens including the genomes and transcriptomes of *Cucumis sativus* (cucumber) and *P. cubensis*. Through this chapter, previous research on the genomics of this host-pathogen system including known factors and processes regulating pathogen virulence and host response will be discussed. Work included in the remaining chapters will discuss research that has uncovered mechanisms of *P. cubensis* transcriptome regulation and identified differentially expressed transcripts and expressed small RNAs that have a role in regulating the resistance response in *C. sativus* to *P. cubensis.*

Pathogen impact and biology: Oomycetes of Cucurbits and downy mildews

History and taxonomy of *Pseudoperonospora cubensis*

Pseudoperonospora cubensis is an obligate biotroph and the causative oomycete pathogen of cucurbit downy mildew, a foliar disease characterized by the development of angular lesions on the leaf surface and by the production of sporangiophores on the underside of the leaf (Lebeda and Cohen 2010; Savory et al., 2010). *P. cubensis* was first discovered by Berkeley and Curtis (1868) from herbarium specimens originating in Cuba, thus the species name *cubensis*. However, it was not identified on live plants until 1903 in Moscow by Rostovzev (Lebeda and Cohen 2010). Although the taxonomy of this pathogen has varied, it is currently classified in Kingdom Stramenopila, Phylum Oomycota, Class Oomycetes, Order Peronosporales, Family Peronosporaceae, Genus *Pseudoperonospora*, Species *Pseudoperonospora cubensis* (Göker et al., 2007; Savory et al., 2010). Like other oomycetes, *P. cubensis* is thiamine-dependent, forms haustoria, and is an obligate, meaning that it requires a host for survival (Göker et al., 2007). Similar to other species in *Peronospora* and *Pseudoperonospora*, it has colored conidia, which may aid in the longevity of sporangia, as they are better protected from solar radiation during wind dispersal (Göker et al., 2007). *Pseudoperonospora* is distinguished from *Peronospora* because it has porous spore walls, germinates to infective zoospores, and has sporangiophore branches that are at acute angles (Igarashi et al., 2009; Palti and Cohen, 1980).

Host range and economic value of *P. cubensis*

Since the discovery of *P. cubensis*, several isolates have been identified, and recent data suggests that the pathogen is continually evolving; this is evidenced by different genotypes of *P. cubensis* observed in various geographic regions of the United States (Quesada-Ocampo et al., 2012; Lebeda and Widrlechner, 2003). At present, the host range for *P. cubensis* includes over 40 species within 20 genera of the Cucurbitaceae family, an important group of crops, which includes *Cucumis sativus* (cucumber), *Cucumis melo* (melon), *Cucurbita pepo* (squash), *Cucurbita maxima* (pumpkin), and *Citrullus lanatus* (watermelon) (Lebeda and Widrlechner 2003). In the United States, for example, the cucurbit industry produces $1.45 billion through 109 million metric tons of cucurbits produced on 229,000 hectares (Cantliffe, 2007). Cucumber downy mildew is a major problem for cucumber growers in the United States and abroad. In Europe, yield losses as high as 80% have been observed, and outbreaks since the 1980's have been managed by intensive fungicide treatments (Lebeda and Urban, 2007). In the United States, the threat of *P. cubensis* to cucumber production was successfully mitigated by host resistance until 2004; however, this pathogen has since overcome host resistance and is now a severe threat to growers, especially as it also evolves fungicide resistance (Holmes et al., 2006; Savory et al., 2010).

P. cubensis life cycle and infection stages

To evaluate the threat of *P. cubensis* and to develop prevention and treatment plans, it is important to study its life cycle. *P. cubensis* is a threatening pathogen because of its rapid polycyclic life cycle, which means that the pathogen is able to complete it's life cycle (as little as 6 days) multiple times in a growing season (Savory et al., 2010). The cycle begins when an airborne sporangia lands on the surface of a leaf and germinates into zoospores (Savory et al., 2010). Several studies have shown the effects of temperature and leaf wetness on germination and infection, and have found that *P. cubensis* prefers slightly cool temperatures (20 °C) and higher humidity (Neufeld and Ojiambo, 2012; Arauz et al., 2010). After germination, a zoospore encysts on stomata and forms a germ tube which develops into an appressorium, followed by a penetration hypha that begins colonization of the intercellular spaces of the mesophyll and palisade tissues (Savory et al., 2010). Hyphae are formed throughout infection and are accompanied by haustoria, which are specialized structures at the plant-pathogen interface that allow for the exchange of materials (Savory et al., 2010). For example, haustoria are the secretion sites for pathogen-associated proteins called effectors, which are involved in manipulating the plant-pathogen interaction in both the apoplast and the plant cytoplasm (Kamoun, 2006). Once *P. cubensis* has colonized the plant, sporangiophores are formed, and sporangia develop at the end of the branches. These sporangia are then released into the wind, and the cycle of infection continues.

Host range, taxonomy, and symptoms of *Phytophthora capsici*

Like *P. cubensis, P. capsici* is a diploid oomycete infects a wide range of cucurbits, including cucumbers, but its host range also extends to pepper, tomato, snap and lima beans, eggplant, and many others (Quesada-Ocampo et al., 2011; Granke et al., 2012; Lamour et al., 2011). *P. capsici* is compared to *P. cubensis* in Table 1. *P. capsici* is classified in the Kingdom Chromista, Phylum Oomycota, Class Oomycetes, Order Peronosporales, Family Peronosporaceae, Genus *Phytophthora*, Species *capsici* (Lamour et al., 2011). The pathogen was first described in 1922 on chili pepper collected in 1918 from a New Mexico Agriculture Experiment Station (Lamour et al., 2011; Leonian, 1922). Since then, *P. capsici* has been identified on several continents including North and South America, Asia, Europe, and Africa (Granke et al., 2012). In parts of the United States and Africa, sexual reproduction and oospore formation has been observed; however asexual reproduction is far more common in South America (Lamour et al., 2011). The different climate of the continents on which *P. capsici* is found also affects the symptoms and severity of the disease, in which a greater range of symptoms affecting the whole plant, including leaf infection, damping off of the root, and stunted plant growth are observed in moister areas (Granke et al., 2012; Lamour et al., 2011). In addition, symptoms of *P. capsici* include infected fruit, identified by the appearance of sporangia in late infection stages, and black/brown lesions on the plant (Lamour et al., 2011). In total, understanding how other oomycetes infect cucumber and how other downy mildews infect other plants can provide insight into the biology and infection mechanisms of *P. cubensis* on its cucumber host.

Table 1. *Lifecycle, host range, and genomics of cucurbit oomycete pathogens.*

	P. cubensis	***P. capsici***
Plant families infected	Cucurbitaceae	Cucurbitaceae, Solanaceae, Fabaceae, and many others
Part of plant affected	Leaves only	Leaves, fruit, stem, root
Dispersal	Wind	Water, soil
Overwintering method	Greenhouses/warm climates in USA; oospores in other countries	Warm climates; irrigation water; oospores in soil common is USA
Life style	Obligate biotroph	Hemibiotroph
Time to complete asexual life cycle	6-10 days	2-3 days
Infection mechanism	Germinated zoospores encyst on stomata and form appressorium	Germinated oospore, germinated sporangia, or zoospore penetrates plant cuticle
Genome	67.9 Mb; 23,000 predicted genes	64 Mb; 17,123 predicted genes

P. capsici life cycle and infection stages

Unlike the obligate biotroph *P. cubensis, P. capsici* is a hemibiotroph, meaning that it has both a biotrophic and a nectrophic phase and can also be cultured on artificial media in a laboratory. In addition, unlike *P. cubensis, P. capsici* regularly forms oospores in the United States, which are a result of sexual reproduction (Lamour et al., 2011). Oospores are formed when two different mating types of *P. capsici* are stimulated to mate and form male and female gametangia, which form haploid gametes that fuse and develop into an oospore. These oospores are able to overwinter and allow this pathogen to be a persistent, difficult problem to manage. Asexually, *P. capsici* reproduces through the formation and propagation of sporangia. The infection cycle begins with penetration of the plant cuticle *via* a germinated oospore, directly germinated sporangia, or a zoospore resulting from germinated sporangia. Once inside the plant, *P. capsici* enters a biotrophic phase in which it colonizes with hyphae and forms haustoria, or feeding structures, but it does not kill the plant tissue. Later, the pathogen switches to a necrotrophic phase in which it does kill the plant tissue, resulting in tissue collapse. Under optimal conditions, these two stages will be completed and the pathogen will sporulate in 2-3 days.

Downy mildews of non-cucurbit hosts

Although *P. capsici* is an oomycete that shares the cucumber host with *P. cubensis,* it does not share the same life cycle; thus, in order to better understand the biology of *P. cubensis* and its obligate interactions with its host, it is informative to examine other downy mildews. A number of different obligate oomycete pathogens are the causative agents of

downy mildew disease on a wide range of hosts, including several vegetables, grapes, and even the model plant *Arabidopsis thaliana*. For vegetables, the most widely studied downy mildew is lettuce downy mildew, *Bremia lactucae,* which has been extensively studied because of its severe economic impact and for its genetic interactions with its host (Michelmore and Wong, 2008). *B. lactucae* is member of the Peronosporales and differs biologically from *P. cubensis* in that the spores of *B. lactucae* germinate directly (do not form zoospores) and are able to penetrate the plant cuticle directly rather than entering through the stomata. *B. lactucae* research has provided valuable insights into the downy mildews in the last several decades by elucidating obligate genome structure, avirulence genes, and factors involved in mating and genetic variation.

In addition to causing damage to vegetables, other downy mildew species cause severe economic damage to plants involved in making popular alcoholic beverages, including beer and wine. Grapevine downy mildew, *Plasmopara viticola,* has caused severe damage to grapes in North America and over the past several decades and has been ranked as one of the top ten oomycetes (Kamoun et al., 2014). Because of the economic importance of this crop to the wine industry, several studies have focused on how to control this devastating pathogen with fungicide; however, researchers have found that *P. viticola* is especially good at overcoming fungicides and is also specialized in the infection of specific hosts (Gessler et al., 2011; Kamoun et al., 2014). A lesser studied downy mildew, but one that is quickly gaining economic significance, is hop downy mildew, *Pseudoperonospora humuli*, which is a close relative of *P. cubensis* and infects hops in the Pacific Northwest and in other areas with a damp climate (Mitchell et al., 2011; Gent and Ocamb, 2009). In addition, basil downy

mildew (*Peronospora belbahrii)* is another emerging pathogen for which scientists have been quickly working to develop detection techniques for this airborne and seedborne pathogen (Koroch et al., 2013; Djalali Farahani-Kofoet et al., 2012).

Among the best studied of the oomycetes is *Hyaloperonospora arabidopsidis,* which infects Arabidopsis and has been used for decades to research the cellular and molecular interactions between a leaf host and an obligate downy mildew (Kamoun et al., 2014). Knowledge of *H. arabidopsidis* is pertinent to the study of *P. cubensis* because the lifecycle and infection process are very similar between these two obligate leaf pathogens with the exception that *H. arabidopsidis* typically enters between epidermal cells instead of through the stomata (Coates and Beynon, 2010). Several isolates of *H. arabidopsidis* and several mutant lines of Arabidopsis exist to study this host-pathogen interaction, and through this research, several effectors in *H. arabidopsidis* and several R genes in Arabidopsis have been discovered (Coates and Beynon, 2010; Fabro et al., 2011). This model system has been especially useful in the functional study of effectors in oomycetes, as many of the *H. arabidopsidis* effectors have been shown to contain the signature the RXLR motif, which facilitates entry into the host cell (Coates and Beynon 2010; Win et al., 2007; Tyler et al., 2013; Yaeno et al., 2011). Additionally, on the plant side of the interaction, multiple mutants have been identified that are resistant to downy mildew, which provides further insight into the way in which the pathogen is able to manipulate the host and or evade the defense response (van Damme et al., 2008; Zeilmaker et al., 2014; Brewer et al., 2014; Gao et al., 2014; Krasileva et al., 2011).

Pathogen biology summary

In total, understanding the biology of other cucumber-infecting oomycetes and downy mildews that infect other hosts will help us to better understand the biology, infection mechanisms, and obligate needs of *P. cubensis.* By studying *P. capsici*, we can gain a better understanding of the biology of an oomycete infecting cucumber; however, transferable knowledge is limited from studying this pathogen because it is a hemibiotroph and largely infects fruits instead of leaves. Alternatively, research from other downy mildew pathogens, despite infecting different hosts, can shed light on the obligate-specific infection mechanisms used by *P. cubensis* to overcome the host defense response while maintaining live plant tissue.

Host resistance: Cucurbit oomycetes and downy mildews on different hosts

Downy mildew resistance in cucurbits

Previous research investigated a susceptible interaction between *P. cubensis* and *C. sativus;* however, the genetic interactions between *P. cubensis* and resistant *C. sativus* are still unknown (Adhikari et al., 2012). Until 2004, downy mildew was controlled in the U.S. by host resistance derived from the plant introduction line PI 197087; however, over the last decade, this resistance has broken down and new sources of host resistance in cucumber have been sought (Call et al., 2012a; 2012b; Holmes et al., 2006). Screening of several cucumber cultigens for resistance to downy mildew has revealed that plant introduction

line PI 197088 is resistant to downy mildew and has thus been used for breeding and quantitative trait loci (QTL) mapping analyses (Call et al., 2012a; 2012b). Historically, downy mildew resistance has been thought to be recessively inherited through two or more loci, including the *dm-1* loci observed in PI 197087 (Call et al., 2012b). However, several recent studies have mapped downy mildew resistance using QTL lines derived from PI 197088, and the number of loci and the mode of inheritance is not consistent with the previous hypothesis (Caldwell et al., 2014; Yoshioka et al., 2014). Indeed, a patent by Caldwell et al., indicates that QTL mapping using PI 197088 as a resistant parent resulted in 3 QTLs with locations on Chromosome 5, 4, and 2, each of which were approximately 50 cM in length, with the QTLs on Chromosomes 4 and 5 being detected in multiple genetic backgrounds and in multiple environmental locations and each contributing approximately 20% to the observed resistance phenotype. This supports our working hypothesis as presented in Chapter 3 that multiple genes contribute to resistance.

In a recent study, Yoshioka et al., identified eleven QTLs in cucumber using a monospore isolate of *P. cubensis* and were successful in confirming resistance effects of four of the QTLs using residual heterozygous lines (RHLs). Of the identified QTL regions, they found that resistance associated with the PI 197088-derived parent was recessive at the *dm5.1* locus, dominant at the *dm4.1* locus, and partially dominant at the *dm3.1* and *dm5.3* loci. Among these QTLs, *dm5.1* and *dm5.3* contributed the most to resistance. As a final note, based on several studies, PI 197088 appears to have different QTL loci for different isolates, genetic backgrounds, and environments, indicating some degree of isolate-specific resistance, background effect, and or environmental influence. Nevertheless, PI 197088 has

been show to be resistant to numerous isolates of *P. cubensis* in multiple different settings, suggesting that resistance is broad-spectrum and somewhat robust.

Beyond using PI 197088 as a resistant parent, a suppression subtractive hybridization study was done using inoculated and non-inoculated samples from the resistant IL-57 line to identify genes up-regulated during infection (Li et al., 2011a). This method identified multiple genes, some of which were transcription factors or were related to reactive oxygen species (ROS), but not all of which were involved in defense-associated responses (Li et al., 2011a). The same group has identified a heat shock protein that may be involved in downy mildew resistance as well as abiotic stresses, including increased temperature (Li et al., 2012). Recently, Zhang et al., has completed a QTL map for downy mildew resistance from two Chinese inbred cucumber lines and has found 5 QTLs associated with resistance to downy mildew using simple sequence repeat (SSR) markers (Zhang et al., 2013). These QTLs are located on chromosomes 1, 6, and 5 and provide broad regions associated with downy mildew resistance, leaving further work to determine the specific genes conferring resistance (Zhang et al., 2013).

In contrast to our sparse understanding of downy mildew resistance in cucumber, specific genes conferring downy mildew resistance in melon are known. Taler et al., found that glyoxylate aminotransferases *At1* and *At2* confer enzymatic resistance genes to downy mildew (Taler et al., 2004). These genes were identified in the wild melon PI 124111 and resistance is observed as chlorotic lesions, a massive amount of callose, and phenolics and lignin-like substances in the cytoplasm. When the *At1* and *At2* genes are transformed into a

susceptible melon, they are able to confer resistance and *P. cubensis* sporulation is not observed. The proposed mechanism of resistance involves increased activation of glycolate oxidase, which increases the amount of hydrogen peroxide that is being produced. The type of resistance conferred by these two genes is unique in that the resistance-associated genes are not the typical R genes, but instead are constitutively highly expressed enzymes.

Resistance to *P. capsici*

Like *P. cubensis,* host resistance to *P. capsici* is not common in cucurbits, and the pathogen is largely managed through cultural practices, like proper irrigation and sanitation, and fungicide usage (Granke et al., 2012). In addition, rapid evolution to fungicides and a great diversity of *P. capsici* populations make managing and breeding for resistance to this pathogen difficult. In 2014, Colle et al., found through a screen of plant introduction lines that young fruits of PI 109483, PI 178884, and PI 214049 all had low disease ratings to *P. capsici* and that this resistance was likely heritable given a study of PI 109483 progeny (Colle et al., 2014). In addition, another study of cucumber found that L-type lectin receptor kinases differentially respond to *P.* capsici between a resistant (JSH) and susceptible (B80) cucumber cultivar (Wu et al., 2014). In seedlings of *Cucumis melo*, accession CNPH-093 and L040 were found to be resistant to five different isolates of *P. capsici* which were collected from different host sources (Pontes et al., 2014). Some sources of complete resistance have been identified in the Solanaceae family in peppers and tomatoes, but thus far the only complete resistance in the Cucurbitaceae family has been identified in squash (Padley et al., 2009; Quesada-Ocampo and Hausbeck, 2010;

Thabuis et al., 2003; 2004) . In squash, 3 dominant alleles identified through breeding were found to confer resistance to *P. capsici* as observed by an absence of the crown rot syndrome (Padley et al., 2009).

In other cucurbits, including cucumber, quantitative, but not qualitative resistance has been observed (Lee et al., 2001; Ando et al., 2009; Gevens et al., 2006). A detached fruit assay was developed to survey the susceptibility of 480 cucumber cultigens and found that while no cultigens displayed a complete absence of infection, some did have reduced sporulation, supporting the concept of quantitative resistance (Gevens et al., 2006). Quantitative resistance to nine isolates of *P. capsici* was also observed using a wide range of pumpkin cultivars from Korea and Japan (Lee et al., 2001). In cucumber and other cucurbits, Ando et al., demonstrated that the age of the fruit has a role in determining its susceptibility to *P. capsici,* with younger fruits, especially cucumbers, being the most susceptible until they reach full fruit length at about 10-12 days after pollination (Ando et al., 2009). In addition, other cucurbits, zucchini and summer squash, were the most susceptible fruits to *P. capsici.* Cucumbers, butternut squash, and zucchini also displayed differences in susceptibility along the length of the fruit in addition to the age-related variation in susceptibility (Ando et al., 2009). Further studies are needed to determine the exact genes involved in mediating *P. capsici* resistance in cucurbits including cucumber, squash, and pumpkin.

Host resistance to other downy mildews

Previous studies of resistance to downy mildew in other species, including in the model system, *Arabidopsis thaliana*, has revealed multiple mechanisms of host resistance. In some cases, resistance to Arabidopsis downy mildew, *H. arabidopsidis*, is mediated by R genes like RPP1, RPP2, RPP5, RPP7, RPP8, and RPP13, which are often specific to individual plant lines against specific races of *H. arabidopsidis* (Nemri et al., 2010; Krasileva et al., 2011; Coates and Beynon, 2010; Mohr et al., 2010). In addition, broad spectrum resistance to *H. arabidopsidis* was found to be conferred through a combination of quantitative and qualitative loci in Arabidopsis C24 (Lapin et al., 2012). In a separate study, a broad-spectrum transcription factor was found to confer resistance to *H. arabidopsidis* and other pathogens by regulating hundreds of genes, including those involved in defense hormone signaling pathways (Gao et al., 2014). Intermediate levels of resistance were uncovered through a global study of 400 interactions between five strains of *H. arabidopsidis* and several Arabidopsis accessions (Krasileva et al., 2011). These intermediate levels of resistance were present in several interactions and were influenced by the developmental stage of the host. Interestingly, despite a gene-for-gene relationship between and Avr and R genes, the interaction between the plant and the pathogen could also be influenced by the pathogen, thus the authors called this relationship a "genome-for-genome" interaction (Krasileva et al., 2011). Other works supports this hypothesis, as it was found that EDM2 (enhanced downy mildew 2) is required for the function of the resistance gene RPP7, and other genes in Arabidopsis also contribute to immunity (Tsuchiya and Eulgem, 2011). Beyond the typical disease resistance-associated response genes and pathways, several studies have identified susceptibility genes and other genes that are able to confer resistance in mutant form;

among these are the homoserine kinase DMR1 (van Damme et al., 2009), the 2-oxoglutarate (2OG)-Fe(II) oxygenase DMR6 (van Damme et al., 2008), the malectin-like receptor kinase IOS1 (Hok et al., 2011; 2014), and the negative regulator of plant defense response GSL5/PMR4 (Wawrzynska et al., 2010).

Like in Arabidopsis, resistance to lettuce downy mildew has been studied for several years and several different mechanisms of resistance have been identified. A cluster of nucleotide binding site leucine-rich repeat (NBS-LRR) genes known as the *RGC2* gene cluster has been associated with resistance and was shown to confer resistance through the multiple *dm* loci within the cluster (Wroblewski et al., 2007). In addition to known genes conferring resistance, multiple QTL mapping studies with different parent lines, have identified QTLs which can also contribute to the resistance phenotype (Simko et al., 2013; Boer et al., 2013). However, it has been found that in some cases, stacked quantitative QTL do not "add up" and resistance from these multiple loci do not have the desired additive effect due to epistasis (Boer et al., 2014). Finally, in lettuce, biotechnology has been used to create resistant lettuce using host-induced gene silencing using plant lines that expressed an siRNA which targeted highly expressed genes in *B. lactucae* (Govindarajulu et al., 2014).

Resistance to grapevine downy mildew (*Plasmopara viticola)* has also been studied on a molecular level, and transcription factors VvWRKY33 and VvWKRY1 are involved in regulating defense (Marchive et al., 2013; Merz et al., 2014). Another WRKY, MrWRKY30 from *Muscadinia rotundifolia* was also found to be involved in grapevine resistance and accumulated in response to *P. viticola* and in plants treated with defense hormones

(SA/JA/ET) (Jiang et al., 2015). Support for this WRKY as a resistance-associated gene was demonstrated by its ability to confer resistance to *Peronospora parasitica* when ectopically expressed in Arabidopsis. Beyond WRKYs, research has also shown that thiamine is capable of increasing phenolic compounds in grapevine cv. Chardonnay which leads to induced resistance to *P. viticola* (Boubakri et al., 2013). In support of this, another group found that a resistant grape cultivar accumulated a distinct set of phenolic compounds upon downy mildew infection compared to its susceptible counterpart (Ali et al., 2012). A general study of resistant grapevines has shown through microscopy-based analyses that *P. viticola* develop fewer haustoria and hyphae on several cultivars of resistant grape (Yu et al., 2012). In total, several studies have investigated resistance to *P. viticola* in grapevine, and sources of resistance and hints of mechanisms of resistance have been discovered; yet this fast-evolving pathogen continues to be a problem.

Resistance summary

Resistance to oomycetes of cucurbits and downy mildews has been well studied, yet much work still needs to be done in order to identify durable sources of resistance and uncover the mechanism of resistance. In several systems, R genes have been identified which confer resistance against many different isolates of oomycetes; however R gene-mediated resistance is easily overcome by mutations in the pathogens. As a result, additional research in oomycete resistance has identified non-R-gene genes, which confer resistance, and several of these have encoded genes previously not thought to play a role in defense response. Still, other more characteristic defense-associated genes, such as genes relating

to hormones or secondary metabolism are also key players in oomycete resistance. Finally, several studies in multiple systems have identified quantitative trait loci that contribute to resistance, yet the genes underlying these QTLs continue to be unknown.

Genomic studies in *P. cubensis, P. capsici,* and *H. arabidopsidis*

Over the last several years, sequencing technology has greatly improved; the quality and quantity of reads has increased while the cost of sequencing has decreased (van Dijk et al., 2014). In a little over a decade, the standard for sequencing has shifted from Sanger sequencing to next generation sequencing technologies like Roche 454 or Illumina which incorporate sequencing by synthesis with the use of pyrosequencing and immersed beads (454) or reversible terminator, fluorescently tagged nucleotides and bridge amplification (Illumina) (van Dijk et al., 2014). For genome assembly projects in which long reads are preferred, new technologies – third generation sequencing platforms like Pacific Biosciences – now allows the sequencing of a single molecule in real time with > 10 kb read length (Buermans and Dunnen, 2014). For transcriptomics studies, RNA-Seq is quickly replacing microarray studies for measuring transcript abundance, and the methods used to analyze the next-generation sequencing data for this purpose are also rapidly advancing (Wang et al., 2009; Love et al., 2014; Buermans and Dunnen, 2014). The advancement of these technologies has enabled whole genome sequencing, transcriptome sequencing, and even small RNA sequencing to become accessible to researchers working on any organism. As a result, it is now possible to study the genomes and transcriptomes of non-model

systems and to use corresponding advancements in bioinformatics to analyze the sequencing data.

P. cubensis genome

The genome of *P. cubensis* was sequenced in 2011 by our lab using the single-spore isolate MSU-1 which was collected from Homerville, Ohio in 2007 (Tian et al., 2011). Genomic DNA was sequenced using the Roche 454 platform, and over 600,000 reads were assembled into 42,799 contigs using the Roche gsAssembler (Tian et al., 2011). This sequencing accounts for approximately 14% of the genome. Over 3,000 of the contigs contained over 500 nucleotides and the longest contig was over 9,000 nucleotides. From this assembly, 604 genes encoding predicted secreted proteins, including 32 putative RXLR effector proteins and 29 QXLR proteins were identified (Tian et al., 2011). Secreted proteins are important because they often encode effectors, which are proteins that aid in pathogenicity and or virulence. These effectors can either function in the host apoplast or can enter the host cytoplasm like Crinklers (CRN) or proteins containing and RXLR motif near the N-terminal of the protein (Kamoun, 2006) The *P. cubensis* genome study also demonstrated that the QXLR motif could replace the functionality of the RXLR motif in *P. cubensis* (Tian et al., 2011). Additional studies performed by our lab discussed in later in this Chapter and in Chapter 2 have worked to improve the genome annotation of *P. cubensis* through additional sequencing using the Illumina platform.

P. capsici genome

Previously, genetic data from *P. capsici* has been used to evaluate population structure using amplified fragment length polymorphisms (AFLP) or single-nucleotide polymorphism (SNP) markers of a select group of common genes (Lamour and Hausbeck, 2001; Quesada-Ocampo et al., 2011). This type of work allowed for regions of the *P. capsici* genome to be sequenced; however, it the full genome was only recently published in 2012 (Lamour et al., 2012). In order to sequence this genetically diverse pathogen, a partial inbred line, LT 1534 was developed. Arachne was used to assemble the 64 Mb genome using reads from both the 454 and Sanger platforms. Following the removal of genes with homology to transposable elements, 17,123 genes were predicted, which is similar to the number of predicted genes in other sequenced oomycete genomes. In addition, the structure of the *P. capsici* genome was similar to other oomycete genomes in that gene-rich regions were conserved in blocks, while gene-poor regions contained mostly repetitive sequences and unique sequences which are potentially related to pathogenesis (Lamour et al., 2012; Jiang and Tyler, 2012). Among these genes related to pathogenesis are effectors, including the well-characterized RXLR and Crinklers (CRN) (Lamour et al., 2012).

In addition, restriction-site-associated DNA (RAD) sequencing and single nucleotide variant (SNV) density were used to genotype 65 different isolates of *P. capsici* to generate a genetic map and examine the genotype of each isolate (Lamour et al., 2012). Through this work, a loss of heterozygosity (LOH) was observed in many of the isolates. LOH is hypothesized to

have significant consequences to *P. capsici,* including changing mating types of an isolate and loss of pathogenicity.

H. arabidopsidis genome

The genome of *H. arabidopsidis* was sequenced in 2010 from the Emoy2 isolate using Sanger shotgun sequencing with at 9.5-fold coverage and Illumina sequencing at a 46-fold coverage (Baxter et al., 2010b). The final combined sequenced genome was 81.6 Mb and had a composition of 42% repetitive elements, which might have caused an under-representation of the genome size due to compression of tandem repeats in the assembly. From the genome assembly, 14,543 genes were predicted, including 134 RXLR effector genes and 20 Crinklers, which were fewer effectors than were predicted in the hemibiotrophic oomycetes *Phytophthora sojae* and *Phytophthora ramorum*. In addition, it was found that the obligate *H. arabidopsidis* is different from non-obligate oomycetes in that it had reduced cell wall degrading enzymes and necrosis and ethylene-inducing proteins, which could be a genomic characteristic of an obligate pathogen compared to a hemibiotrophic pathogen (Baxter et al., 2010b).

P. cubensis transcriptome and effector analyses in *C. sativus*

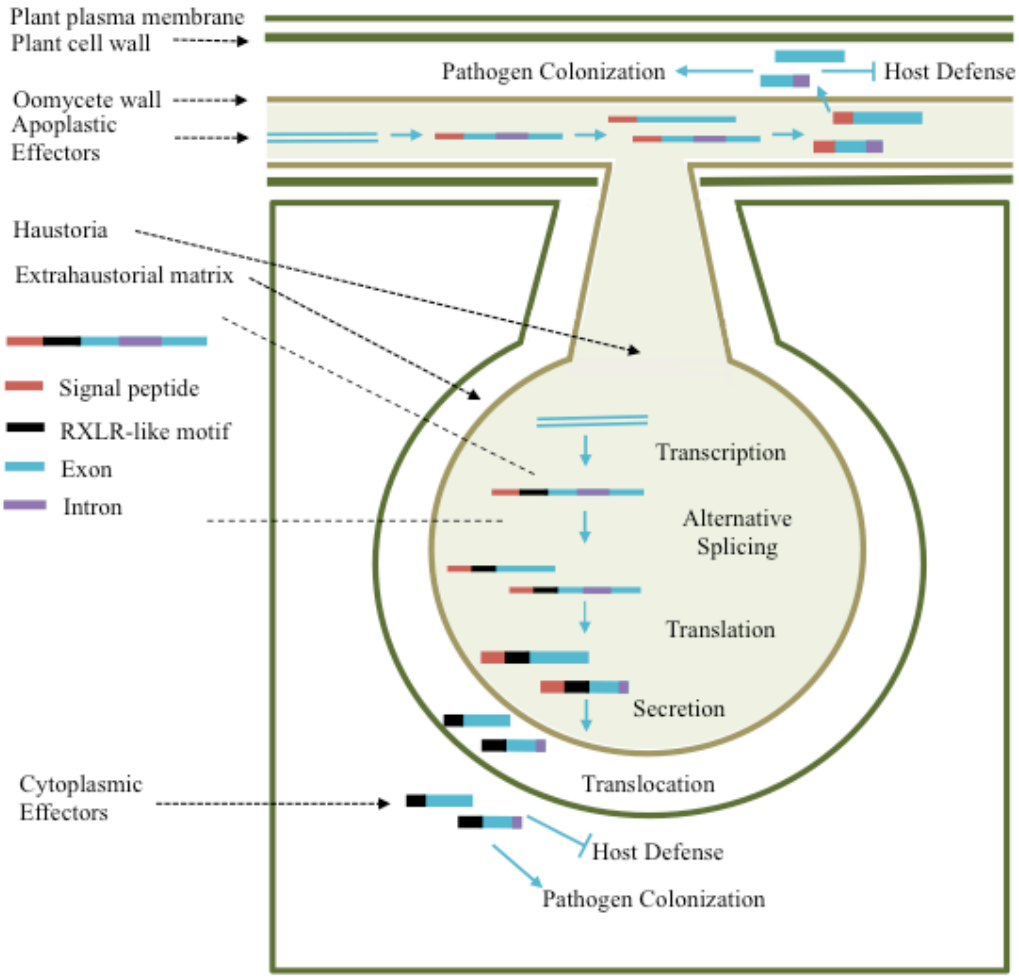
To better understand the pathogen and its life cycle at a molecular level, previous research has characterized the infection stages using a combination of microscopy and next-generation sequencing (Savory et al., 2012a). To this end, the transcriptomes of *P. cubensis*

and 'Vlaspik', a susceptible *C. sativus* cultivar, were sequenced at 1, 2, 3, 4, 6, and 8 days post inoculation (dpi) using next generation Illumina technology (Adhikari et al., 2012; Savory et al., 2012a). These 35-42 bp single end reads were then mapped to the *P.* cubensis genome using Bowtie and TopHat and were processed and quantified using Cufflinks and Cuffdiff (Savory et al., 2012a). Out of the approximately 23,000 predicted genes in the 67.9 Mb genome of *P. cubensis*, 7,821 genes were expressed within the time course. The relationship between differentially expressed genes at each time point was evaluated using the Pearson correlation coefficient, and early, middle, and late phases of infection were determined. These phases correlated well with the expression of orthologous genes in another oomycete, *Phytophthora infestans,* especially at 4 dpi during which haustoria formation and pathogen colonization is occurring. In addition to using the Pearson correlation coefficient, the data was also analyzed using the Weighted Gene Correlation Network Analysis (WGCNA) in which genes with similar expression patterns over the time course were grouped into representative gene models called eigengenes.

Within the transcriptome, multiple effectors have been predicted in *P. cubensis;* Figure 1 diagrams a mechanism by which these effectors, both apoplastic and cytoplasmic, could be expressed and utilized to the benefit of the pathogen. For example, genes related to the virulence and pathogenicity of *P. cubensis* include enzymes that damage the host, including proteases, lipases, and carbohydrate active enzymes (Savory et al., 2012a). In addition, 271 putative pathogen effector genes were identified, many of which were found to have an RXLR-type motif with a varied R1 position (Savory et al., 2012a). The number of effectors predicted in *P. cubensis* is within the range of predicted effectors in other oomycetes,

including the obligate *Arabidopsis thaliana* oomycete, with 134 RXLR-like effectors and the well-characterized potato late blight pathogen, *Phytophthora infestans,* with 563 predicted RXLR-like effectors (Stassen and Van den Ackerveken 2011; Baxter et al., 2010a). The RXLR motif, an important component of cytoplasmic effectors, has been demonstrated to facilitate effector entry through interaction with phosphatidylinositol-3-phosphate (PI3P) on the plant cell (Jiang et al., 2008; Kale et al., 2010; Stassen and Van den Ackerveken, 2011; Whisson et al., 2007). Once delivered into the host cell, the pathogen molecule functions in large part to abrogate host defenses and/or facilitate pathogen survival. While the mechanism(s) and targeting of host defenses by pathogen effectors has made significant strides over the past decade, one area of research whose mechanism remains largely undefined is the process of host entry (Jiang and Tyler, 2012).

Figure 1. *Diagrammatic representation of oomycete infection at site of haustoria invasion of host cell*. The molecular mechanism by which oomycete genes are expressed, spliced, translated, and secreted (as denoted by the red signal peptide box) is highlighted. Cytoplasmic effectors, which enter the plant cell, often have an RXLR-like motif, as denoted by the black box. Apoplastic effectors, which are secreted from the oomycete but do not enter the plant cell, are also shown and follow the same pattern of transcription, splicing, translation, and secretion. Both types of effectors function to abrogate host defense responses and promote pathogen colonization.



In this regard, recent debate has challenged the role of the RXLR/PI3P-mediated effector entry, and specifically, has proposed the importance of a positively charged patch over the canonical RXLR (Yaeno et al., 2011). In support of this, research from our group has demonstrated that variation of the translocation-associated RXLR motif, in association with the previously characterized dEER motif, is important for effector translocation into the host (Tian et al., 2011). Mechanistically, it has been determined that the RXLR motif and surrounding amino acids are important in effector dimerization (Wawra et al., 2012). The same study also indicated that the C-terminus, not the RXLR, is able to bind to PI3P; however, this binding was only observed when the protein was denatured, suggesting that this observation is likely not biologically relevant (Wawra et al., 2012). Moving forward, a complete analysis and understanding of the role of the RXLR-motif in effector entry is needed to resolve current debate.

P. capsici transcriptome and effectors

The *P. capsici* transcriptome was recently published in 2013 using the Illumina platform to sequence RNA collected fro mycelia, zoospores, and germinating cysts; 13,901, 14,633, and 14,695 expressed genes were detected from each sample, respectively (Chen et al., 2013). From the sequenced transcriptome, 98 predicted effector genes were identified and effectors representing RXLRs, Crinklers, and apoplastic proteins were validated through transient expression in *Nicotiana benthamiana*. Most of the effectors showed an increase in gene expression in either the zoospore or the germinating cyst stage compared to the mycelia. In addition to a sequenced transcriptome which identified Crinkler expression, a

recent publication performed an in-depth analysis of this important component of the pathogen's transcriptome and identified and reannotated 84 CRN in the *P. capsici* genome (Stam et al., 2013). The CRN expansion within the genome might be associated with hemibiotropy and or necrotrophy (Stam et al., 2013).

H. arabidopsidis transcriptome

The *H. arabidopsidis* transcriptome has been sequenced using different techniques, and both have focused on the identification of effectors that are expressed during host infection (Cabral et al., 2011; Asai et al., 2014). Cabral et al., sequenced the expressed sequence tags (ESTs) from several *H. arabidopsidis* isolates and found 2,164 *H. arabidopsidis*-derived unigenes of which 75 were specific to *H. arabidopsidis* and 42 were putative effectors. Among the effectors, elicitins, Crinkler, and RXLR proteins were studied more in-depth and it was found that RXLR29 was capable of suppressing pathogen-induced callose deposition and enhanced disease susceptibility in Arabidopsis (Cabral et al., 2011). A recent study used transcriptome sequencing to monitor the transcript level changes in a compatible or incompatible *H. arabidopsidis* isolate and the transcript level changes in an Arabidopsis host simultaneously (Asai et al., 2014). Through this approach, the authors were able to identify that *H. arabidopsidis* suppresses SA-induced gene expression through the expression a highly induced effector; this transcriptome sequencing-based research could begin to answer questions about what makes a pathogen compatible on a plant in the absence of an R-gene mediated defense response (Asai et al., 2014).

Alternative splicing as a method of transcriptome regulation

Given that many sequenced transcriptomes now exist for oomycetes, the next step in genomics research is to understand how these transcriptomes are regulated and what implication this regulation has on the pathogenicity and virulence of the pathogen. As a result of increased transcriptomic data, alternative splicing, which was a previously understudied phenomenon, especially in non-model systems, can now be analyzed on a genome-wide scale. By definition, alternative splicing is a process in which a pre-mRNA transcript, typically transcribed from DNA by RNA polymerase II, is processed into multiple isoforms to include different combinations of introns and exons by a mechanism that is mediated by a complex of proteins called the spliceosome (Kelemen et al., 2013; Nilsen and Graveley, 2010). This transcript processing is more common across a genome than was previously thought and can lead to a more diverse transcriptome, and thus proteome, using four main splicing mechanisms, including intron retention, exon skipping, alternative 3' splice site selection, and alternative 5' splice site selection (Nilsen and Graveley, 2010). Alternative splicing has been observed in transcripts across kingdoms, including animals, fungi, protists, and plants (McGuire et al., 2008). From these cross-kingdom analyses, the general trend was observed that fungi and protists (including the oomycetes *P. infestans* and *P. sojae*) were far more likely to use intron retention as an alternative splicing mechanism compared to other types of splicing (McGuire et al., 2008).

Thus far, few comprehensive surveys have investigated the breadth, type, and or effect of alternative splicing on the secretome of oomycetes, and none have specifically looked at

alternative splicing of oomycete effector genes. The first splicing in an oomycete sequence was reported only among a family of 5-endoglucanases in the *Phytophthora* genus and was validated using reverse transcriptase (RT)-PCR (Costanzo et al., 2007). Previous work on a genome-wide scale has investigated spliced transcripts based on expressed sequence tag (EST) data in both *P. infestans* and *P. sojae*, and they had 4,762 and 2,125 spliced transcripts, respectively, with a majority of the splice variants being retained introns (McGuire et al., 2008) . Recently, Shen et al., focusing on *P. sojae* transcription using ESTs identified 4013 introns in the genome sequence and found that 122 of the genes were alternatively spliced and validated three genes using RT-PCR (Shen et al., 2011). Previous work in our lab, described below, has investigated alternative splicing within a single gene in *P. cubensis*; Chapter 2 provides a genome-wide analysis of alternative splicing in *P. cubensis,* which utilizes the aforementioned transcriptomic resources.

Alternative splicing of *P. cubensis* effectors

The transcriptome of *P. cubensis* (Savory et al., 2012a) provides insight into the potential roles of identified effectors, given their patterns and levels of gene expression. Models of effector genes have gained additional complexity through a recent study showing that effector genes from *P. cubensis* can be alternatively spliced, which would hypothetically broaden the potential effector proteome (Savory et al., 2012b). Savory et al., demonstrated that a full-length gene that was predicted to be a non-effector multi-drug transporter resulted in a functional effector as a result of an alternative splicing mechanism called intron retention in which a premature stop codon was introduced (Savory et al., 2012b).

Previous and ongoing work in the lab involves the characterization of alternatively spliced effectors. In Chapter 2, alternative splicing in *P. cubensis* is examined on a genome-wide scale and both effector and non-effector transcripts were analyzed.

Pathogen genomics summary

Next-generation sequencing technologies have enabled the genomes and transcriptomes of non-model systems to be sequenced and analyzed. The genomes now exist for several oomycetes, including *P. cubensis*, which enables comparative genomics between oomycetes. Furthermore, transcriptome studies allow researchers to identify differentially expressed genes that could be contributing to pathogenicity or virulence in the pathogen or depressed defense response genes within the host. The transcriptomics studies also reveal differences that occur in gene transcription between different life stages of the pathogen, which reveal key genes regulating different stages of pathogen infection. In addition, deeper analyses of expressed transcripts could enable research on the occurrence and timing of splicing events, which could a key mechanism underpinning pathogen transcriptome regulation.

Genomics of *Cucumis sativus* (cucumber)

C. sativus genome

In 2009, the cucumber genome of 'Chinese long' inbred line 9930 was sequenced using a combination of Sanger and Illumina sequencing to achieve and 72-fold coverage of the

243.5 Mb genome (Huang et al., 2009). From the assembled genome, 26,682 genes were predicted, including 171 miRNA genes. For pathogen-associated genes, 61 NBS-LRR were identified, and it was hypothesized that the expansion of a family of lipoxygenase genes in the cucumber genome might also be associated with resistance. In 2011, the cucumber genome (version 2) was improved through RNA-Seq data, which examined several tissues from different parts of the cucumber plant and identified 23,248 protein-coding genes in which 8,700 genes had structural modifications compared to the original genome version and 5,285 new genes were discovered (Li et al., 2011b). In addition to the first sequenced cucumber genome, another cucumber, the North-European Borszczagowski cultivar, was recently sequenced and significant chromosomal rearrangements were observed, and the distribution of ABA-related genes were studied (Wóycicki et al., 2011).

Cucumber transcriptome in response to *P. cubensis*

RNA-Seq data from the previously described experiment (Savory et al., 2012) also has yielded valuable information on the cucumber transcriptome response during infection, in which 14,476 genes were expressed, with 3,286 genes being differentially expressed between the time points collected 1-8 dpi (Adhikari et al., 2012). These genes were grouped into modules using the Weighted Gene Correlation Network Analysis (WGCNA), and groups of genes with similar expression patterns were identified and represented as eigengenes. This type of analysis allows genes that are highly coexpressed at specific time points to be quickly identified. For example, a group of defense-related cucumber genes, including lipoxygenases, cationic peroxidases, and cinnamate 4-hydrolases were down-

regulated at 2 dpi in this susceptible interaction. This is very interesting because it could indicate that the pathogen is actively mitigating host defense responses through diminishing the expression of host defense genes. Other genes involved in defense and stress signaling were expressed throughout the time course – including peroxidases, protease inhibitors, catalases, lipoxygenases, and galactinol synthase – indicating that a susceptible cucumber is capable of eliciting a response to the pathogen, but is likely not sufficient to stop or prevent infection.

Small RNAs in cucumber

Over the past several years, several studies have indicated that small RNAs (sRNAs) play a key role in regulating the transcriptome of an organism and that sRNAs also influence the plant-pathogen dynamic (Weiberg et al., 2014). The small RNA pathway is such an essential component of plant defense response that oomycete pathogens have evolved effectors that silence sRNA biogenesis, which lead to increased susceptibility (Qiao et al., 2013). In addition, *Botrytis cinerea* directly transfers fungal-derived sRNAs to silence host defense-associated transcripts in Arabidopsis and tomato (Weiberg et al., 2013). Small RNAs are formed from double stranded RNA and are processed by Dicer or Dicer-like proteins to yield 20-30 nucleotide fragments which are then loaded onto Argonaute (AGO) proteins that guide the sRNA to the target transcript with complementary base pairing to the sRNA (Axtell, 2013). Once at the target locations, the AGO proteins can then silence the target transcript through cleavage and degradation, DNA methylation, histone modification, or translational inhibition; thus sRNA-directed transcript regulation has high potential to

incur significant changes within the transcriptome (Axtell, 2013).

Small RNAs are often categorized based on their mode of biogenesis and length of sequence. Small RNAs include hairpin RNAs and short interfering RNAs (siRNA), which are derived from RNA forming a hairpin from a single strand or from RNA forming a double-stranded complex, respectively (Axtell, 2013). Among siRNAs, multiple types exist depending on their biogenesis. For example, phased siRNAs are secondary siRNAs that result from the RNA-dependent RNA polymerase synthesis of dsRNA from a previous sRNA-induced event. Other siRNA are called (natural antisense transcript) NAT-siRNA, which form from the hybridization of transcribed mRNAs. Among the best studied sRNAs are miRNAs which form from characteristic hairpin structures and are typically 21-24 nucleotides long in plants and usually display perfect complementarity to their targets in the 2-13 nucleotide positions (Axtell 2013; Eldem et al., 2013).

In cucumber, several studies have identified and characterized small RNAs, but none have done so in the presence of a pathogen during a resistant interaction. Among these studies, Martínez et al., identified 19 conserved families of miRNA in cucumber and identified 7 new miRNAs (Martínez et al., 2011). In a novel approach, Mao et al., combined small RNA sequencing with degradome sequencing in cucumber to identify a total of 64 miRNA and to validate 21 of their targets, which included genes involved in development, signal transduction, and transcriptional regulation (Mao et al., 2012). In a study among five different species of cucurbit, 220 miRNAs were among all the Cucurbitaceae studied and identified 41 within *C. sativus* (Hu et al., 2014). A grafting-based study identified 48 new

miRNA in pumpkin and cucumber and identified their targets (Li et al., 2013). Interestingly, they also found that an miRNA could silence tissue that was grafted onto the tissue expressing the miRNA, which could lead to potential applications in silencing cucumber genes for genetic studies. Finally, Jin and Wu have identified miRNAs in cucumber during infection with *P. cubensis* and found that 39 known miRNAs were induced in response to pathogen stress (Jin and Wu, 2015). Furthermore, they identified several novel miRNAs in cucumber, three of which were responsive to *P. cubensis* (Jin and Wu, 2015).

Cucumber genomics summary

Over the past five years, the genomic resources available for cucumber have been quickly developed. Multiple genomes have been sequenced, expressed transcripts during pathogen infection have been identified, and several studies have examined small RNAs under multiple biological conditions. These resources will allow other genomics-based studies to be conducted, which will enable us to understand how all of the genes within cucumber are regulated in response to certain environmental conditions, namely biotic or abiotic stress. More work still needs to be done in order to use these great genomic resources to improve agriculture and to use the cucumber genome sequence to conduct studies examining resistance to pathogens. In Chapter 3, the cucumber transcriptomes of a resistant and susceptible host to *P. cubensis* were compared with the intent of identifying key genes and gene pathways that are involved in regulating the resistance response. In addition, the work in Chapter 3 builds upon previous small RNA work in cucumber and identifies miRNA genes that are predicted to be involved in regulating biotic stress responses.

REFERENCES

REFERENCES

Adhikari, B. N., Savory, E. A., Vaillancourt, B., Childs, K. L., Hamilton, J. P., Day, B., and Buell, C. R. 2012. Expression profiling of *Cucumis sativus* in response to infection by *Pseudoperonospora cubensis. PLoS ONE* 7:e34954. Published online.

Ali, K., Maltese, F., Figueiredo, A., Rex, M., Fortes, A. M., Zyprian, E., Pais, M. S., Verpoorte, R., and Choi, Y. H. 2012. Alterations in grapevine leaf metabolism upon inoculation with *Plasmopara viticola* in different time-points. *Plant Sci.* 191-192:100-107.

Ando, K., Hammar, S., and Grumet, R. 2009. Age-related resistance of diverse cucurbit fruit to infection by *Phytophthora capsici*. *Jour. Am. Soc. Hort. Sci.* 134:176–182.

Arauz, L. F., Neufeld, K. N., Lloyd, A. L., and Ojiambo, P. S. 2010. Quantitative models for germination and infection of *Pseudoperonospora cubensis* in response to temperature and duration of leaf wetness. *Phytopathology* 100:959–967.

Asai, S., Rallapalli, G., Piquerez, S. J. M., Caillaud, M.-C., Furzer, O. J., Ishaque, N., Wirthmueller, L., Fabro, G., Shirasu, K., and Jones, J. D. G. 2014. Expression profiling during Arabidopsis/downy mildew interaction reveals a highly-expressed effector that attenuates responses to salicylic acid. *PLoS Pathog.* 10:e1004443.

Axtell, M. J. 2013. Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol.* 64:137–159.

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

Berkeley, M.S. and Curtis, A. 1868. *Peronospora cubensis*. *Bot. Linn. Soc.* 10:363.

Boer, den, E., Pelgrom, K. T. B., Zhang, N. W., Visser, R. G. F., Niks, R. E., and Jeuken, M. J. W. 2014. Effects of stacked quantitative resistances to downy mildew in lettuce do not simply add up. *Theor. Appl. Genet.* 127:1805–1816.

Boer, den, E., Zhang, N. W., Pelgrom, K., Visser, R. G. F., Niks, R. E., and Jeuken, M. J. W. 2013. Fine mapping quantitative resistances to downy mildew in lettuce revealed multiple sub-QTLs with plant stage dependent effects reducing or even promoting the infection. *Theor. Appl. Genet*. 126:2995–3007.

Boubakri, H., Poutaraud, A., Wahab, M. A., Clayeux, C., Baltenweck-Guyot, R., Steyer, D., Marcic, C., Mliki, A., and Soustre-Gacougnolle, I. 2013. Thiamine modulates metabolism of the phenylpropanoid pathway leading to enhancedresistance to *Plasmopara viticola* in grapevine. *BMC Plant Biol.* 13:1–1.

Brewer, H. C., Hawkins, N. D., and Hammond-Kosack, K. E. 2014. Mutations in the Arabidopsis homoserine kinase gene DMR1 confer enhanced resistance to *Fusarium culmorum* and *F. graminearum. BMC Plant Biol.* 14:317.

Buermans, H. P. J., and den Dunnen, J. T. 2014. Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta*. 1842:1932–1941.

Cabral, A., Stassen, J. H. M., Seidl, M. F., Bautor, J., Parker, J. E., and Van den Ackerveken, G. 2011. Identification of *Hyaloperonospora arabidopsidis* transcript sequences expressed during infection reveals isolate-specific effectors. *PLoS ONE* 6:e19328. Published online.

Caldwell, D., Chan, E., de Vries, J., Joobeur, T., King, J., Reina, A., and Shetty, N. V. 2014. Methods and compositions for identifying downy mildew resistance cucumber plants. US Patent 8809622. Date issued: August 19.

Call, A. D., Criswell, A. D., Wehner, T. C., Ando, K., and Grumet, R. 2012a. Resistance of cucumber cultivars to a new strain of cucurbit downy mildew. *HortScience* 47:171–17.

Call, A. D., Criswell, A. D., Wehner, T. C., Klosinska, U., and Kozik, E. U. 2012b. Screening cucumber for resistance to downy mildew caused by (Berk. and Curt.) Rostov. *Crop Science* 52:577-592.

Chen, X.-R., Xing, Y.-P., Li, Y.-P., Tong, Y.-H., and Xu, J.-Y. 2013a. RNA-Seq reveals infection-related gene expression changes in *Phytophthora capsici*. *PLoS ONE* 8:e74588. Published online.

Coates, M. E., and Beynon, J. L. 2010. *Hyaloperonospora arabidopsidis* as a pathogen model. *Annu. Rev. Phytopathol.* 48:329–345.

Colle, M., Straley, E. N., Makela, S. B., Hammar, S. A., and Grumet, R. 2014. Screening the cucumber plant introduction collection for young fruit resistance to *Phytophthora capsici*. *HortScience* 49:244–249.

Costanzo, S., Ospina-Giraldo, M. D., Deahl, K. L., Baker, C. J., and Jones, R. W. 2007. Alternate intron processing of family 5 endoglucanase transcripts from the genus *Phytophthora. Curr. Genet.* 52:115–123.

Djalali Farahani-Kofoet, R., Römer, P., and Grosch, R. 2012. Systemic spread of downy mildew in basil plants and detection of the pathogen in seed and plant samples. *Mycol. Progress* 11:961–966.

Eldem, V., Okay, S., and Unver, T. 2013. Plant microRNAs: new players in functional genomics. *Turkish Jour. Agr. Forestry.* 37:1–21.

Fabro, G., Steinbrenner, J., Coates, M., Ishaque, N., Baxter, L., Studholme, D. J., Körner, E., Allen, R. L., Piquerez, S. J. M., Rougon-Cardoso, A., Greenshields, D., Lei, R., Badel, J. L., Caillaud, M.-C., Sohn, K.-H., Van den Ackerveken, G., Parker, J. E., Beynon, J., and Jones, J. D. G. 2011. Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host plant immunity. *PLoS Pathog.* 7:e1002348.

Gao, D., Appiano, M., Huibers, R. P., Chen, X., Loonen, A. E. H. M., Visser, R. G. F., Wolters, A. M. A., and Bai, Y. 2014. Activation tagging of ATHB13 in *Arabidopsis thaliana* confers broad-spectrum disease resistance. *Plant Mol. Biol.* 86:641–653.

Gent, D. H., and Ocamb, C. M. 2009. Predicting Infection Risk of Hop by *Pseudoperonspora humuli. Phytopathology* 99:1190–1198.

Gessler, C., Pertot, I., and Perazzolli, M. 2011. *Plasmopara viticola*: a review of knowledge on downy mildew of grapevine and effective disease management. *Phytopathologia Mediterranea.* 50:3–44.

Gevens, A. J., Ando, K., Lamour, K. H., Grumet, R., and Hausbeck, M. K. 2006. A detached cucumber fruit method to screen for resistance to *Phytophthora capsici* and effect of fruit age on susceptibility to infection. *Plant Dis.* 90:1276–1282.

Govindarajulu, M., Epstein, L., Wroblewski, T., and Michelmore, R. W. 2014. Host-induced gene silencing inhibits the biotrophic pathogen causing downy mildew of lettuce. *Plant Biotech. Jour.* Published online.

Göker, M., Voglmayr, H., Riethmüller, A., and Oberwinkler, F. 2007. How do obligate parasites evolve? A multi-gene phylogenetic analysis of downy mildews. *Fun. Genet. Biol*. 44:105–122.

Granke, L. L., Quesada-Ocampo, L., Lamour, K., and Hausbeck, M. K. 2012. Advances in research on *Phytophthora capsici* on vegetable crops in the United States. *Plant Dis*. 96:1588–1600

Hok, S., Allasia, V., Andrio, E., Naessens, E., Ribes, E., Panabieres, F., Attard, A., Ris, N., Clement, M., Barlet, X., Marco, Y., Grill, E., Eichmann, R., Weis, C., Huckelhoven, R., Ammon, A., Ludwig-Muller, J., Voll, L. M., and Keller, H. 2014. The receptor kinase IMPAIRED OOMYCETE SUSCEPTIBILITY1 attenuates abscisic acid responses in Arabidopsis. *Plant Physiol.* 166:1506–1518.

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

Holmes, G. J., Wehner, T. C., and Thornton, A. 2006. An old enemy re-emerges. *Disease Manag.* 2:14–15.

Hu, T., Sun, X., Zhang, X., Nevo, E., and Fu, J. 2014. An RNA sequencing transcriptome analysis of the high-temperature stressed tall fescue reveals novel insights into plant thermotolerance. *BMC Genomics* 15:1–13.

Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., Lucas, W. J., Wang, X., Xie, B., Ni, P., Ren, Y., Zhu, H., Li, J., Lin, K., Jin, W., Fei, Z., Li, G., Staub, J., Kilian, A., van der Vossen, E. A. G., Wu, Y., Guo, J., He, J., Jia, Z., Ren, Y., Tian, G., Lu, Y., Ruan, J., Qian, W., Wang, M., Huang, Q., Li, B., Xuan, Z., Cao, J., Asan, Wu, Z., Zhang, J., Cai, Q., Bai, Y., Zhao, B., Han, Y., Li, Y., Li, X., Wang, S., Shi, Q., Liu, S., Cho, W. K., Kim, J.-Y., Xu, Y., Heller-Uszynska, K., Miao, H., Cheng, Z., Zhang, S., Wu, J., Yang, Y., Kang, H., Li, M., Liang, H., Ren, X., Shi, Z., Wen, M., Jian, M., Yang, H., Zhang, G., Yang, Z., Chen, R., Liu, S., Li, J., Ma, L., Liu, H., Zhou, Y., Zhao, J., Fang, X., Li, G., Fang, L., Li, Y., Liu, D., Zheng, H., Zhang, Y., Qin, N., Li, Z., Yang, G., Yang, S., Bolund, L., Kristiansen, K., Zheng, H., Li, S., Zhang, X., Yang, H., Wang, J., Sun, R., Zhang, B., Jiang, S., Wang, J., Du, Y., and Li, S. 2009. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41:1275-1281.

Igarashi, A., Yamagata, K., Sugai, T., Takahashi, Y., Sugawara, E., Tamura, A., Yaegashi, H., Yamagishi, N., Takahashi, T., Isogai, M., Takahashi, H., and Yoshikawa, N. 2009. Apple latent spherical virus vectors for reliable and effective virus-induced gene silencing among a broad range of plants including tobacco, tomato, *Arabidopsis thaliana*, cucurbits, and legumes. *Virology* 386:407–416.

Jiang, R. H. Y., and Tyler, B. M. 2012. Mechanisms and evolution of virulence in oomycetes. *Annu. Rev. Phytopathol.* 50:295–318.

Jiang, R. H., Tripathy, S., Govers, F., and Tyler, B. M. 2008. RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proc. Nat. Acad. Sci.* 105:4874–4879.

Jiang, W., Wu, J., Zhang, Y., Yin, L., and Lu, J. 2015. Isolation of a WRKY30 gene from *Muscadinia rotundifolia* (Michx) and validation of its function under biotic and abiotic stresses. *Protoplasma* Published online.

Jin, W. and Wu, F. 2015. Identification and characterization of cucumber microRNAs in response to *Pseudoperonospora cubensis* infection. *Gene* doi: 10.1016/j.gene.2015.05.064.

Kale, S. D., Gu, B., Capelluto, D. G. S., Dou, D., Feldman, E., Rumore, A., Arredondo, F. D., Hanlon, R., Fudal, I., Rouxel, T., Lawrence, C. B., Shan, W., and Tyler, B. M. 2010. External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell.* 142:284–295.

Kamoun, S. 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu. Rev. Phytopathol.* 44:41–60.

Kamoun, S., Furzer, O., Jones, J. D. G., Judelson, H. S., Ali, G. S., Dalio, R. J. D., Roy, S. G., Schena, L., Zambounis, A., Panabieres, F., Cahill, D., Ruocco, M., Figueiredo, A., Chen, X.-R., Hulvey, J., Stam, R., Lamour, K., Gijzen, M., Tyler, B. M., Grunwald, N. J., Mukhtar, M. S., Tomé, D. F. A., Tör, M., Van den Ackerveken, G., McDowell, J., Daayf, F., Fry, W. E., Lindqvist-Kreuze, H., Meijer, H. J. G., Petre, B., Ristaino, J., Yoshida, K., Birch, P. R. J., and Govers, F. 2014. The top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol.* 16: 413-434.

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., and Stamm, S. 2013. Function of alternative splicing. *Gene.* 514:1–30.

Koroch, A. R., Villani, T. S., Pyne, R. M., and Simon, J. E. 2013. Rapid staining method to detect and identify downy mildew (*Peronospora belbahrii*) in Basil. *App. in Plant Sci.* 1:1300032.

Krasileva, K. V., Zheng, C., Leonelli, L., Goritschnig, S., Dahlbeck, D., and Staskawicz, B. J. 2011. Global Analysis of Arabidopsis/Downy Mildew interactions reveals prevalence of incomplete resistance and rapid evolution of pathogen recognition. *PLoS ONE.* 6:e28765.

Lamour, K. H., and Hausbeck, M. K. 2001. Investigating the spatiotemporal genetic structure of *Phytophthora capsici* in Michigan. *Phytopathology.* 91:973–980.

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., Miller, N. A., Rice, B. J., Raffaele, S., Cano, L. M., Bharti, A. K., Donahoo, R. S., Finley, S., Huitema, E., Hulvey, J., Platt, D., Salamov, A., Savidor, A., Sharma, R., Stam, R., Storey, D., Thines, M., Win, J., Haas, B. J., Dinwiddie, D. L., Jenkins, J., Knight, J. R., Affourtit, J. P., Han, C. S., Chertkov, O., Lindquist, E. A., Detter, C., Grigoriev, I. V., Kamoun, S., and Kingsmore, S. F. 2012. Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici. Mol. Plant Microbe Interact.* 25:1350–1360.

Lamour, K. H., Stam, R., Jupe, J., and Huitema, E. 2011. The oomycete broad-host-range pathogen *Phytophthora capsici. Mol Plant Pathol.* 13:329–337.

Lapin, D., Meyer, R. C., Takahashi, H., Bechtold, U., and Van den Ackerveken, G. 2012. Broad-spectrum resistance of Arabidopsis C24 to downy mildew is mediated by different combinations of isolate-specific loci. *New Phytol.* 196:1171–1181.

Lebeda, A., and Cohen, Y. 2010. Cucurbit downy mildew *(Pseudoperonospora cubensis*)—biology, ecology, epidemiology, host-pathogen interaction and control. *Eur. J. Plant Pathol.* 129:157–192.

Lebeda, A., and Widrlechner, M. P. 2003. A set of Cucurbitaceae taxa for differentiation of *Pseudoperonospora cubensis* pathotypes. *Jour. Plant Dis. Protect.* 110:337.

Lebeda, A. and Urban, J. 2007. Temporal changes in pathogenicity and fungicide resistance in *Pseudoperonospora cubensis* populations. *Acta. Hortic.* 731: 327-336.

Lee, B. K., Kim, B. S., Chang, S. W., and Hwang, B. K. 2001. Aggressiveness to pumpkin cultivars of isolates of *Phytophthora capsici* from pumpkin and pepper. *Plant Dis.* 85:497–500.

Li, C., Li, Y., Bai, L., Zhang, T., He, C., Yan, Y., and Yu, X. 2013. Grafting-responsive miRNAs in cucumber and pumpkin seedlings identified by high-throughput sequencing at whole genome level. *Physiol Plantarum.* 151:406–422.

Li, J., Zhang, H., Hu, J., Liu, J., and Liu, K. 2012. A heat shock protein gene, CsHsp45.9, involved in the response to diverse stresses in cucumber. *Biochem. Genet.* 50:565–578.

Li, J.-W., Liu, J., Zhang, H., and Xie, C.-H. 2011a. Identification and transcriptional profiling of differentially expressed genes associated with resistance to P*seudoperonospora cubensis* in cucumber. *Plant Cell Rep.* 30:345–357.

Li, Z., Zhang, Z., Yan, P., Huang, S., Fei, Z., and Lin, K. 2011b. RNA-Seq improves annotation of protein-coding genes in the cucumber genome. *BMC Genomics* 12:540.

Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol.* 15:550.

Mao, W., Li, Z., Xia, X., Li, Y., and Yu, J. 2012. A combined approach of high-throughput sequencing and degradome analysis reveals tissue specific expression of microRNAs and their targets in cucumber. *PLoS ONE* 7:e33040. Published online.

Marchive, C., Léon, C., Kappel, C., Coutos-Thévenot, P., Corio-Costet, M.-F., Delrot, S., and Lauvergeat, V. 2013. Over-expression of VvWRKY1 in grapevines induces expression of jasmonic acid pathway-related genes and confers higher tolerance to the downy mildew. *PLoS ONE.* 8:e54185. Published online.

Martínez, G., Forment, J., Llave, C., Pallás, V., and Gómez, G. 2011. High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs. *PLoS ONE* 6:e19523. Published online.

McGuire, A. M., Pearson, M. D., Neafsey, D. E., and Galagan, J. E. 2008. Cross-kingdom patterns of alternative splicing and splice recognition. *Genome Biol*. 9:R50. Published online.

Merz, P. R., Moser, T., Höll, J., Kortekamp, A., Buchholz, G., Zyprian, E., and Bogs, J. 2014. The transcription factor VvWRKY33 is involved in the regulation of grapevine (*Vitis vinifera*) defense against the oomycete pathogen *Plasmopara viticola*. *Physiol Plantarum.* 153:365-380.

Michelmore, R., and Wong, J. 2008. Classical and molecular genetics of *Bremia lactucae*, cause of lettuce downy mildew. *Eur. J. Plant Pathol.* 122:19–30.

Mitchell, M. N., Ocamb, C. M., Grunwald, N. J., Mancino, L. E., and Gent, D. H. 2011. Genetic and pathogenic relatedness of *Pseudoperonospora cubensis* and *P. humuli. Phytopathol.* 101:805-818.

Mohr, T. J., Mammarella, N. D., Hoff, T., Woffenden, B. J., Jelesko, J. G., and McDowell, J. M. 2010. The Arabidopsis downy mildew resistance gene RPP8 is induced by pathogens and salicylic acid and is regulated by W box cis elements. *Mol. Plant Microbe Interact.* 23:1303–1315.

Nemri, A., Atwell, S., Tarone, A. M., Huang, Y. S., Zhao, K., Studholme, D. J., Nordborg, M., and Jones, J. D. G. 2010. Genome-wide survey of Arabidopsis natural variation in downy mildew resistance using combined association and linkage mapping. *Proc. Nat. Acad. Sci.* 107:10302–10307.

Neufeld, K. N., and Ojiambo, P. S. 2012. Interactive effects of temperature and leaf wetness duration on sporangia germination and infection of cucurbit hosts by *Pseudoperonospora cubensis. Plant Dis*. 96:345–353.

Nilsen, T. W., and Graveley, B. R. 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463:457–463

Padley, L. D., Kabelka, E. A., and Roberts, P. D. 2009. Inheritance of resistance to crown rot caused by *Phytophthora capsici* in Cucurbita. *HortScience* 44:211–213.

Palti, J. and Cohen, Y. 1980. Downy mildew of Cucurbits (*Pseudoperonospora cubensis)*: The fungus and its hosts, distribution, epidemiology, and control. *Phytoparasitica* 8:109-147.

Pontes, N. de C., Aguiar, F. M., Boiteux, L. S., Lima, Milton L. P., Oliveira, V. R., Cafe Filho, A. C., and Reis, A. 2014. Identification of sources of seedling resistance to *Phytophthora capsici* in *Cucumis melo*. *Trop. Plant Pathol.* 39:74–81.

Qiao, Y., Liu, L., Xiong, Q., Flores, C., Wong, J., Shi, J., Wang, X., Liu, X., Xiang, Q., Jiang, S., Zhang, F., Wang, Y., Judelson, H. S., Chen, X., and Ma, W. 2013. Oomycete pathogens encode RNA silencing suppressors. *Nature.* 45:330–333.

Quesada-Ocampo, L. M., and Hausbeck, M. K. 2010. Resistance in tomato and wild relatives to crown and root rot caused by *Phytophthora capsici. Phytopathology*. 100:619–627.

Quesada-Ocampo, L. M., Granke, L. L., Mercier, M. R., Olsen, J., and Hausbeck, M. K. 2011. Investigating the genetic structure of *Phytophthora capsic*i populations. *Phytopathol.* 101:1061–1073.

Quesada-Ocampo, L. M., Granke, L. L., Olsen, J., Gutting, H. C., Runge, F., Thines, M., Lebeda, A., and Hausbeck, M. K. 2012. The genetic structure of *Pseudoperonospora cubensis* populations. *Plant Dis.* 96:1459–1470.

Savory, E. A., Adhikari, B. N., Hamilton, J. P., Vaillancourt, B., Buell, C. R., and Day, B. 2012a. mRNA-Seq analysis of the *Pseudoperonospora cubensis* transcriptome during cucumber (*Cucumis sativus* L.) infection. *PLoS ONE* 7:e35796. Published online.

Savory, E. A., Granke, L. L., Quesada-Ocampo, L. M., Varbanova, M., Hausbeck, M. K., and Day, B. 2010. The cucurbit downy mildew pathogen *Pseudoperonospora cubensis. Mol. Plant Pathol.* 12:217–226.

Savory, E. A., Zou, C., Adhikari, B. N., Hamilton, J. P., Buell, C. R., Shiu, S.-H., and Day, B. 2012b. Alternative splicing of a multi-drug transporter from *Pseudoperonospora cubensis* generates an RXLR effector protein that elicits a rapid cell death. *PLoS ONE* 7:e34701. Published online.

Shen, D., Ye, W., Dong, S., Wang, Y., and Dou, D. 2011. Characterization of intronic structures and alternative splicing in *Phytophthora sojae* by comparative analysis of expressed sequence tags and genomic sequences. *Can. J. Microbiol.* 57:84–90.

Simko, I., Atallah, A. J., Ochoa, O. E., Antonise, R., Galeano, C. H., Truco, M. J., and Michelmore, R. W. 2013. Identification of QTLs conferring resistance to downy mildew in legacy cultivars of lettuce. *Sci. Rep.* 3:2875. Published online.

Stam, R., JUPE, J., Howden, A. J. M., Morris, J. A., Boevink, P. C., Hedley, P. E., and Huitema, E. 2013. Identification and characterisation CRN effectors in *Phytophthora capsici* shows modularity and functional diversity. *PLoS ONE* 8:e59517. Published online.

Stassen, J. H., and Van den Ackerveken, G. 2011. How do oomycete effectors interfere with plant life? *Curr. Opin. Plant Biol.*14:407-414.

Taler, D. , Galperin, M., Benjamin, I., Cohen, Y., and Kenigsbuch, D. 2004. Plant eR genes that encode photorespiratory enzymes confer resistance against disease. *Plant Cell*. 16:172-184.

Thabuis, A., Palloix, A., Pflieger, S., Daubèze, A.-M., Caranta, C., and Lefebvre, V. 2003. Comparative mapping of *Phytophthora* resistance loci in pepper germplasm: evidence for conserved resistance loci across *Solanaceae* and for a large genetic diversity. *Theor. Appl. Genet*. 106:1473–1485.

Thabuis, A., Palloix, A., Servin, B., Daubèze, A.-M., Signoret, P., and Lefebvre, V. 2004. Marker-assisted introgression of 4 *Phytophthora capsic*i resistance QTL alleles into a bell pepper line: validation of additive and epistatic effects. *Mol. Breed.* 14:9–20.

Tian, M., Win, J., Savory, E., Burkhardt, A., Held, M., Brandizzi, F., and Day, B. 2011. 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Mol. Plant Microbe Interac*t. 24:543–553.

Tsuchiya, T., and Eulgem, T. 2011. EMSY-like genes are required for full RPP7-mediated race-specific immunity and basal defense in Arabidopsis. *Mol. Plant Microbe Interact.* 24:1573–1581.

Tyler, B. M., Kale, S. D., Wang, Q., Tao, K., Clark, H. R., Drews, K., Antignani, V., Rumore, A., Hayes, T., Plett, J. M., Fudal, I., Gu, B., Chen, Q., Affeldt, K. J., Berthier, E., Fischer, G. J., Dou, D., Shan, W., Keller, N. P., Martin, F., Rouxel, T., and Lawrence, C. B. 2013. Microbe-independent entry of oomycete RxLR effectors and fungal RxLR-like effectors into plant and animal cells is specific and reproducible. *Mol. Plant Microbe Interact*. 26:611–616.

van Damme, M., Huibers, R. P., Elberse, J., and Van den Ackerveken, G. 2008. *Arabidopsis* DMR6 encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew. *Plant Jour.* 54:785–793.

van Damme, M., Zeilmaker, T., Elberse, J., Andel, A., de Sain-van der Velden, M., and Van den Ackerveken, G. 2009. Downy mildew resistance in *Arabidopsis* by mutation of HOMOSERINE KINASE. *Plant Cell* 21:2179-2189.

van Dijk, E. L., Auger, H., Jaszczyszyn, Y., and Thermes, C. 2014. Ten years of next-generation sequencing technology. *Trend. Genet*. 30:418–426.

Wang, Z., Gerstein, M., and Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev. Genet*. 10:57–63.

Wawra, S., Agacan, M., Boddey, J. A., Davidson, I., Gachon, C. M. M., Zanda, M., Grouffaud, S., Whisson, S. C., Birch, P. R. J., Porter, A. J., and van West, P. 2012. Avirulence protein 3a (AVR3a) from the potato pathogen *Phytophthora infestans* forms homodimers through its predicted translocation region and does not specifically bind phospholipids. *Jour. Biol. Chem.* 287:38101–38109.

Wawrzynska, A., Rodibaugh, N. L., and Innes, R. W. 2010. Synergistic activation of defense responses in *Arabidopsis* by simultaneous loss of the GSL5 callose synthase and the EDR1 protein kinase. *Mol. Plant Microbe Interact*. 23:578–584.

Weiberg, A., Wang, M., Bellinger, M., and Jin, H. 2014. Small RNAs: A new paradigm in plant-microbe interactions. *Annu. Rev. Phytopathol.* 52:495–516.

Weiberg, A., Wang, M., Lin, F. M., Zhao, H., and Zhang, Z. 2013. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* 342:118-123.

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

Win, J., Morgan, W., Bos, J., Krasileva, K. V., Cano, L. M., Chaparro-Garcia, A., Ammar, R., Staskawicz, B. J., and Kamoun, S. 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell* 19:2349–2369.

Wóycicki, R., Witkowicz, J., Gawroński, P., Dąbrowska, J., Lomsadze, A., Pawełkowicz, M., Siedlecka, E., Yagi, K., Pląder, W., Seroczyńska, A., Śmiech, M., Gutman, W., Niemirowicz-Szczytt, K., Bartoszewski, G., Tagashira, N., Hoshi, Y., Borodovsky, M., Karpiński, S., Malepszy, S., and Przybecki, Z. 2011. The genome sequence of the North-European cucumber (*Cucumis sativus* L.) unravels evolutionary adaptation mechanisms in plants. *PLoS ONE* 6:e22728. Published online.

Wroblewski, T., Piskurewicz, U., Tomczak, A., Ochoa, O., and Michelmore, R. W. 2007. Silencing of the major family of NBS-LRR-encoding genes in lettuce results in the loss of multiple resistance specificities. *Plant Jour.* 51:803–818.

Wu, T., Wang, R., Xu, X., He, X., Sun, B., Zhong, Y., Liang, Z., Luo, S., and Lin, Y. 2014. *Cucumis sativus* L-type lectin receptor kinase (CsLecRK) gene family response to *Phytophthora melonis*, *Phytophthora capsici* and water immersion in disease resistant and susceptible cucumber cultivars. *Gene.* 549:214–222.

Yaeno, T., Li, H., Chaparro-Garcia, A., Schornack, S., Koshiba, S., Watanabe, S., Kigawa, T., Kamoun, S., and Shirasu, K. 2011b. Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector AVR3a is required for its stability in host cells to modulate plant immunity. *Proc. Nat. Acad. Sci.* 108:14682–14687.

Yoshioka, Y., Sakata, Y., Sugiyama, M., and Fukino, N. 2014. Identification of quantitative trait loci for downy mildew resistance in cucumber (*Cucumis sativus* L.). *Euphytica.* 198:265–276.

Yu, Y., Zhang, Y., Yin, L., and Lu, J. 2012. The mode of host resistance to *Plasmopara viticola* infection of grapevines. *Phytopathol.* 102:1094–1101.

Zeilmaker, T., Ludwig, N. R., Elberse, J., Seidl, M. F., Berke, L., Van Doorn, A., Schuurink, R. C., Snel, B., and Van den Ackerveken, G. 2014. DOWNY MILDEW RESISTANT 6 and DMR6-LIKE OXYGENASE 1 are partially redundant but distinct suppressors of immunity in *Arabidopsis*. *Plant Jou.* 81:210–222.

Zhang, S. P., Liu, M. M., Miao, H., Zhang, S. Q., Yang, Y. H., Xie, B. Y., Wehner, T. C., and Gu, X. F. 2013. Chromosomal mapping and QTL analysis of resistance to downy mildew in *Cucumis sativus. Plant Dis.* 97:245–251.

CHAPTER 2

Alternative splicing in the obligate biotrophic oomycete pathogen, *Pseudoperonospora cubensis*

This Chapter was originally published in *Molecular Plant-Microbe Interactions.*

Burkhardt, A., Buchanan, A., Cumbie, J. S., Savory, E. A., Chang, J. H., and Day, B. 2015. Alternative splicing in the obligate biotrophic oomycete pathogen *Pseudoperonospora cubensis*. Mol. Plant-Microbe Interact. 28:298-309.

Abstract

Pseudoperonospora cubensis is an obligate pathogen and causative agent of cucurbit downy mildew. To help advance our understanding of the pathogenicity of *P. cubensis*, we used RNA-Seq to improve the quality of its reference genome sequence. We also characterized the RNA-Seq dataset to inventory transcript isoforms and infer alternative splicing during different stages of its development. Almost half of the original gene annotations was improved and nearly 4,000 previously unannotated genes were identified. We also demonstrated that approximately 24% of the expressed genome, and nearly 55% of the intron-containing genes from *P. cubensis*, had evidence for alternative splicing. Our analyses revealed that intron retention is the predominant alternative splicing type in *P. cubensis*, with alternative 5′- and alternative 3′-splice sites occurring at lower frequencies. Representatives of the newly identified genes and predicted alternatively spliced transcripts were experimentally validated. The results presented herein highlight the utility of RNA-Seq for improving draft genome annotations, and through this approach, we demonstrate that alternative splicing occurs more frequently than previously predicted. In total, the current study provides evidence that alternative splicing plays a key role in transcriptome regulation and proteome diversification in plant pathogenic oomycetes.

Introduction

Pseudoperonospora cubensis (*P. cubensis*) is the causal agent of cucurbit downy mildew, one of the most agriculturally important foliar diseases of the Cucurbitaceae (Burkhardt & Day,

2013). Like other oomycete members of the 'SAR' (Stramenopiles, Heterokonts, and Alveolates) supergroup, *P. cubensis* is highly destructive, with epidemics of downy mildew on cucumber (*Cucumis sativus* L.) causing up to 100% yield loss (Savory et al., 2010; Beakes et al., 2012; Burkhardt & Day, 2013). In recent years, particularly in the United States, the effectiveness of host resistance has decreased and is no longer sufficient for management of the disease (Holmes et al., 2006). The economic losses, coupled with rising fungicide costs associated with controlling *P. cubensis* outbreaks, highlight the importance of studying the *P. cubensis*-cucumber pathosystem.

P. cubensis is a challenging pathogen to study due to its obligatory host-associated lifestyle and its intractability to genetic manipulation. The development of genomic and transcriptomic resources are difficult because of the need to disentangle pathogen sequences from those of the host. Some of these difficulties have been overcome, and draft genome and transcriptome reference sequences have been generated for both the pathogen and its cucumber host (Tian et al., 2011; Adhikari et al., 2012; Savory et al., 2012a; Savory et al., 2012b;). Furthermore, transcriptome-wide changes have been documented for both *P. cubensis* and cucumber during infection that characterized gene expression during different life stages of *P. cubensis* including free-swimming sporangia, encystment, invasion via stomata, and hyphal expansion through a susceptible cucumber variety (Adhikari et al., 2012; Savory et al., 2012a).

The draft genome sequence of *P. cubensis* has also been mined for candidate virulence determinants, such as effector proteins that may play a role in *P. cubensis* pathogenesis

(Tian et al., 2011; Savory et al., 2012b). Oomycetes rely on both apoplastic and cytoplasmic effector proteins, which are secreted into the apoplast or the extrahaustorial matrix, respectively, to facilitate infection of their hosts (Kamoun, 2006). Candidate effectors can be predicted based on the presence of a functional signal peptide, which is a feature necessary for secretion from the pathogen. Additionally, cytoplasmic effectors that are translocated across the host membrane can be identified based on the presence of a short RXLR amino acid motif (Rehmany et al., 2005; Whisson et al., 2007), or as shown in *P. cubensis,* an alternative QXLR motif (Tian et al., 2011). Based on the original draft genome sequence, *P. cubensis* is predicted to encode 271 secreted effector proteins, of which 67 had identifiable R/QXLR motifs, that were found to be expressed at all stages of infection surveyed (Savory et al., 2012a; Savory et al., 2012b; Tian et al., 2011). The first *P. cubensis* effector identified, *PcQNE*, is a protein that contains a QXLR motif that is necessary for translocation into host cells (Tian et al., 2011). The second characterized effector protein from *P. cubensis,* PscRXLR1, is unique in that its gene has evidence for alternative splicing. The product of the constitutively spliced gene product was predicted to encode a multi-drug transporter, which in many analyses would likely be dismissed as a false positive. The alternatively spliced transcript, on the other hand is predicted to yield a candidate secreted RXLR effector due to the retention of an intron and the introduction of a premature termination codon (Savory et al., 2012b).

Alternative splicing (AS) is a fundamental and evolutionarily conserved eukaryotic process that leads to multiple transcripts from a single gene (Chen et al., 2012; McGuire et al., 2008). In higher eukaryotes, AS plays important roles in development and stress signaling

(Mastrangelo et al., 2012), circadian-dependent regulation of host physiology (James et al., 2012; Sanchez et al., 2011; Schöning et al., 2007), the modulation of gene expression in human genetic diseases (Nicholson et al., 2009; Pistoni et al., 2010), and the response of plants to pathogen infection (Rayson et al., 2012). Beyond diversifying the products of expressed genes, AS also has a role in regulating gene expression (Boothby et al., 2013; Guan et al., 2013). In microbial eukaryotes, AS is less characterized but has nevertheless been identified as an important mechanism for regulating gene expression and influencing the ability of eukaryotic microbes to infect their plant hosts (Juneau et al., 2007; Pleiss et al., 2007; Wilhelm et al., 2008; Kloppholz et al., 2011; Shen et al., 2011; Savory et al., 2012b).

Eukaryotic genes are transcribed as pre-mRNAs, and during their maturation, introns are spliced out and exons are joined via a process that is carried out by the spliceosome (Hoskins & Moore, 2012). The intron-exon boundaries of transcripts are defined via a small number of short and degenerate *cis*-regulatory sequences, the 5′ and 3′ splice sites (5′ss and 3′ss), the branch point sequence, and the polypyrimidine tract (Keren et al., 2010). Additional degenerate enhancer and silencing sequences have also been shown to recruit splicing regulatory proteins to influence both constitutive and AS (Keren et al., 2010). In the case of the latter, alternative splice sites are used to generate varying combinations of transcripts from a single gene locus, thereby regulating both the transcriptome and proteome (Kalsotra & Cooper, 2011).

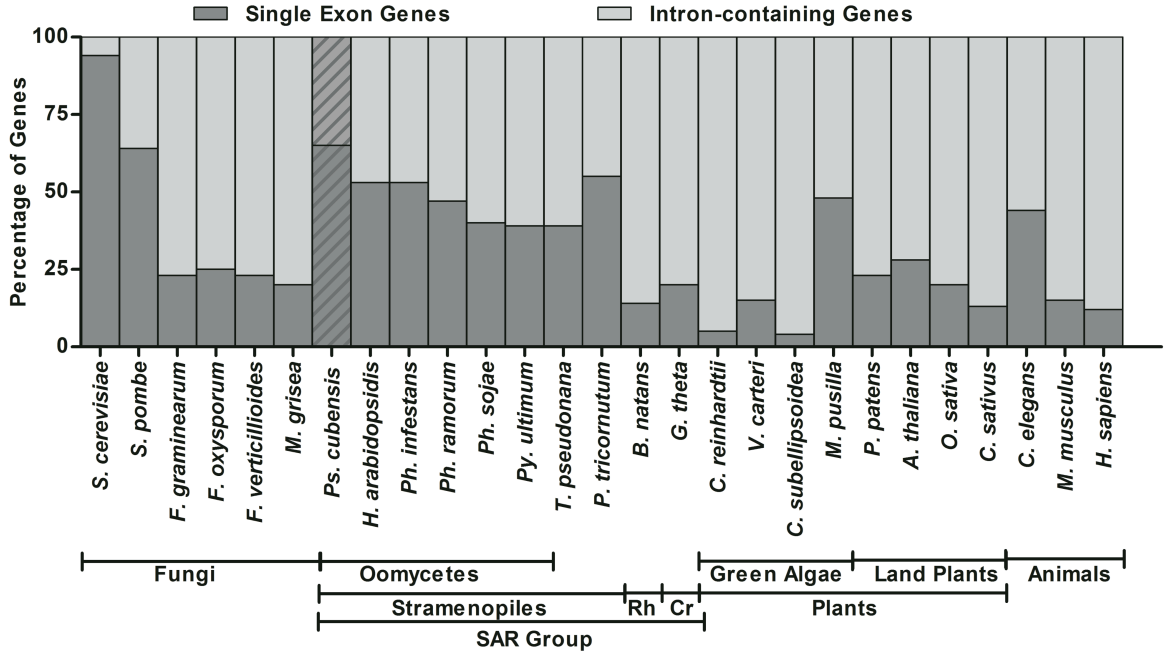
Two mechanisms of splice site selection have been proposed (Keren et al., 2010). In the first, exon definition, the spliceosome recognizes and places the initial machinery across

the exon. Generally, this mechanism is more prevalent in higher eukaryotes (Keren et al., 2010). In the second, intron definition, the spliceosome is placed across introns and recognizes splice sites at each end of the same intron. This mechanism is predicted to be more ancient and common to eukaryotic microbes (Keren et al., 2010). The differences in mechanisms impose constraints that are reflected in gene architecture and in the degree and type of AS. Higher metazoans, such as mice and humans, tend to have short exons and long introns with a high degree of AS and a preference for exon skipping (Kim et al., 2008; Pan et al., 2008; Sultan et al., 2008). Conversely, eukaryotic microbes like yeast have shorter introns and lower frequencies of AS that favor intron retention (Irimia et al., 2007; Kim et al., 2008; Meyer & Vilardell, 2008; Wilhelm et al., 2008). It is hypothesized that plants use intron definition, and in this regard, are more similar to eukaryotic microbes, with one notable exception being the high frequency of AS in plants compared to microbes (Wang & Brendel, 2006; Reddy, 2007; Filichkin et al., 2010; Marquez et al., 2012; Syed et al., 2012).

There has been limited analysis of alternative splicing in oomycetes. Previous analyses of *Phytophthora infestans* and *Phytophthora sojae* have relied on expressed sequence tag (ESTs) datasets, which could contribute to an underestimate of the true breadth of alternative splicing in this group of organisms (McGuire et al., 2008; Shen et al., 2011). For example only 405 and 66 splice variants were identified in *P. infestans* and *P. sojae*, respectively, with intron retention being predominant in both species (McGuire et al., 2008). In a more detailed study of *P sojae*, unique sequences at the 5'ss were identified and it was suggested that these sequences may function in AS (Shen et al., 2011). However, of

the 1681 predicted intron-bearing gene models, only 122 were identified as having an alternative splicing event (Shen et al., 2011). These observations of a low frequency of AS are in conflict with the low strength of the 5´ss of introns identified in *P. sojae* because low strength is typically an indicator of intron-richness and prominent AS (Irimia et al., 2007; Shen et al., 2011). Additionally, members of the SAR group have introns in 35-86% of their genes, indicating a potential for higher rates of splicing and AS than predicted from analysis of ESTs (Fig. 2).

Figure 2. *Comparison of the inventory of intron-bearing genes in Pseudoperonospora cubensis to other eukaryotes.* The percent of single exonic genes (dark grey bars) and intron-bearing genes (light grey bars) from transcriptomes of representative eukaryotes were calculated and plotted. Organisms are clustered based on their phylogenetic relationships.



Based on its original genome reference sequence, *P. cubensis* has the lowest percentage of intron-bearing genes (approximately 35%) among the oomycetes surveyed (Fig. 2). *P. cubensis* averages 0.7 introns per gene with the average intron length being 86 nucleotides. In the current study, we used RNA-Seq to identify expressed genes from *P. cubensis*, and used this information to update existing gene models, to identify new genes, and to predict AS events; the latter was prompted by a previous study from our group that identified and characterized an alternatively spliced effector-encoding gene (Savory et al., 2012b). Herein, we show that of intron-bearing genes (24% of the expressed genes), approximately 55% exhibited AS – including alternative 3′, alternative 5′, and intron retention events. Representatives of the newly identified genes and predicted alternatively spliced transcripts were validated for expression and alternative splicing. We suggest that AS in oomycete plant pathogens is substantially higher than previously predicted and is at a level closer to that of plants. Furthermore, AS was observed in genes predicted to encode signal peptide-containing proteins and putative effectors, which provides support for the hypothesis that AS could influence the virulence of this important plant pathogen.

Results

Generation and analysis of *P. cubensis* RNA-Seq datasets

We used RNA-Seq to re-sequence cDNA libraries derived from *P. cubensis* during infection of its cucumber host to quantify transcripts with evidence for alternative splicing. We used paired-end (PE) sequencing on an Illumina HiSeq and generated a total of approximately

735 million cDNA fragments (170-210 million fragments per sample) from bar-coded RNA-Seq libraries derived from two biologically replicated samples of 2, 3, 4 and 8 days post-inoculation (dpi) (Adhikari et al., 2012; Savory et al., 2012a). The PE RNA-Seq reads from the greater than 20.5 million previously sequenced fragments of *P. cubensis* sporangia (Savory et al., 2012a) were also included in this study.

More than 750 million sequenced fragments were pre-processed for quality and categorized based on origin. Consistent with the previous study, the majority of the reads mapped to the host (Savory et al., 2012a). Only approximately 40 million sequenced fragments aligned uniquely to and within the *P. cubensis* reference sequence either as an ungapped alignment with up to two mismatches or as a gapped alignment with no mismatches. Fragments from sporangia, intermediate (2-4 dpi), and late stages (8 dpi) of the infection comprised approximately 40%, 30%, and 30%, respectively, of the total aligned fragment sequences. The majority of the greater than 750 million sequenced fragments (approximately 650 million sequenced fragments), aligned with greater confidence to the *C. sativus* reference sequence. The remaining >60 million sequenced fragments did not align to either reference sequence. Approximately 75% of these sequences were likely replete with sequencing errors, as they were rare (≤10 RNA-Seq reads). The other unaligned sequenced fragments likely correspond to other microorganisms living in or on the infected cucumber leaf. The experiment was not conducted in a sterile environment and BLASTN analysis (threshold e-value = 1×10^{-5}) of 1000 randomly selected read sequences used as queries returned homologies to sequences of prokaryotic origin.

Improvements to the *P. cubensis* expressed genome annotation

To more confidently predict alternative splicing, we first used the RNA-Seq sequences to improve the annotation of the *P. cubensis* reference genome sequence (Fig. 3A). Nearly 45% of the reference genome sequence was aligned with an average depth of 218 RNA-Seq reads per sequenced nucleotide. A total of approximately 30 million sequenced fragments, aggregated across all libraries, aligned to 13,483 out of the 23,519 (57%) previously annotated genes and were classified as expressed (Savory et al., 2012a). Another 8.2 million sequenced fragments aligned to regions of the genome devoid of annotated features. Those that coalesced into distinct clusters exceeding length and coverage thresholds were used to improve the genome annotation. The majority of the clusters of reads were used to improve 11,892 genes annotated in the original genome reference sequence. Of the updated genes, ~29.7% were extended and another 13.4% had intron definitions altered; some genes were both extended and had intron definitions altered. Among the extended genes approximately 2,000 of the improvements were the addition of new coding sequences, and 10,779 were addition of 5′ and/or 3′ untranslated regions (UTRs). In addition to the alterations to loci of the original genome annotation, we also classified 4,072 features as new candidate gene loci (14.5%), thus increasing the expressed genome of *P. cubensis* to approximately 17,500 genes (approximately 64% of the total number of annotated genes; http://dx.doi.org/10.7267/N9TD9V7M). A small set of genes were selected based on the predicted level of expression in sporangia, and were tested using reverse transcriptase (RT)-PCR to confirm our prediction of new genes (Fig. 4).

Figure 3. Pseudoperonospora cubensis *genome annotation improvement using RNA-Seq*. A, Genes were first improved based on the coverage of sequencing reads to the reference sequence. B, Gene models were predicted based on the alignment of reads to intron sequences or those that aligned as gapped reads. Original: examples of gene models from the original genome annotation. Coverage: example of RNA-Seq read coverage. Predicted Introns: based on original annotation or gap alignment of RNA-Seq reads. Current: improved gene or gene model annotation. Percent of genes: ratio of the number of genes affected in each category, divided by the total number of original genes plus newly identified genes. Percent of reads: ratio of the number of RNA-Seq reads used in each category, divided by the total number of mapped reads (for A, 31.3% of reads were not used); N/A = not applicable.

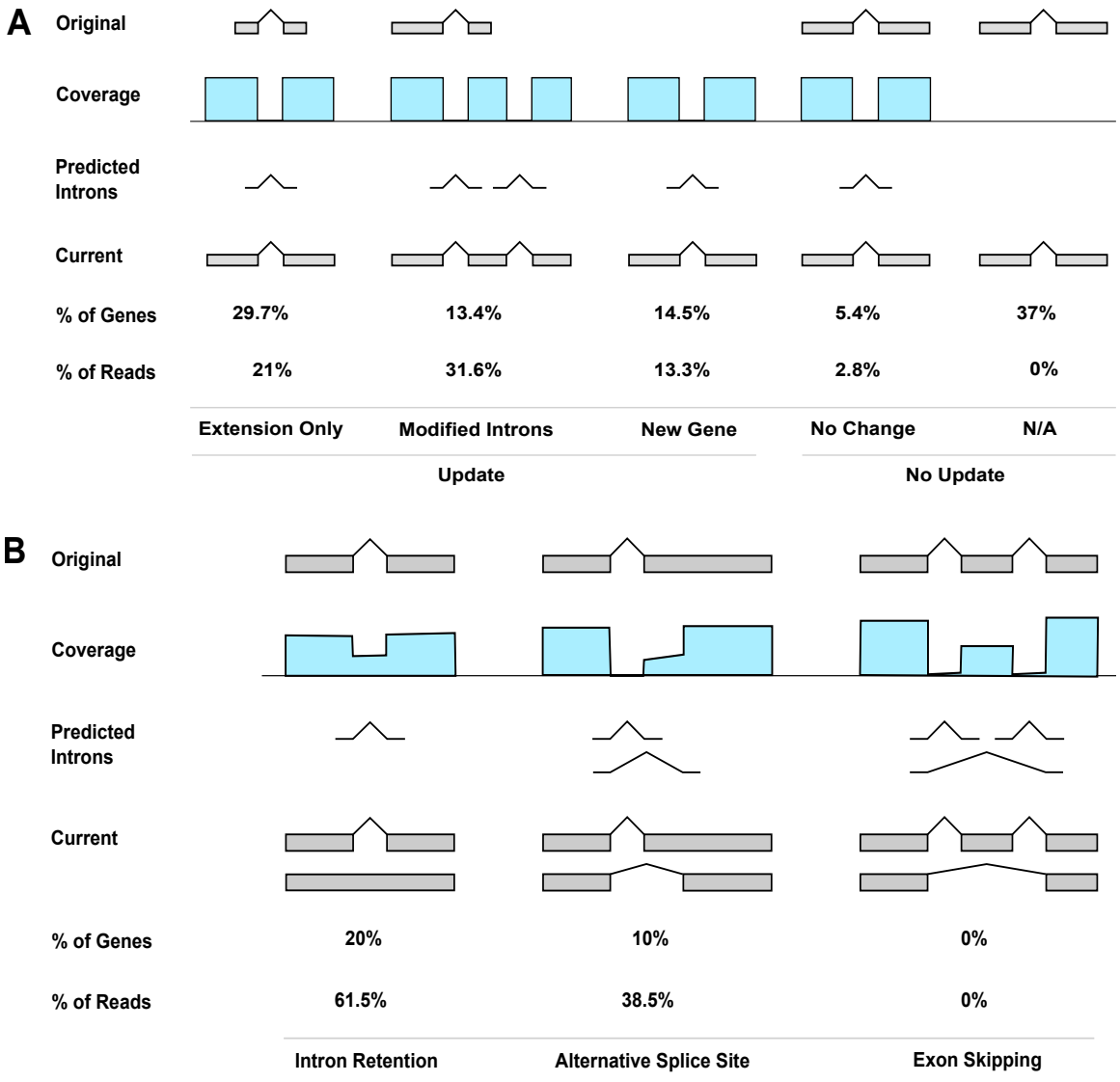
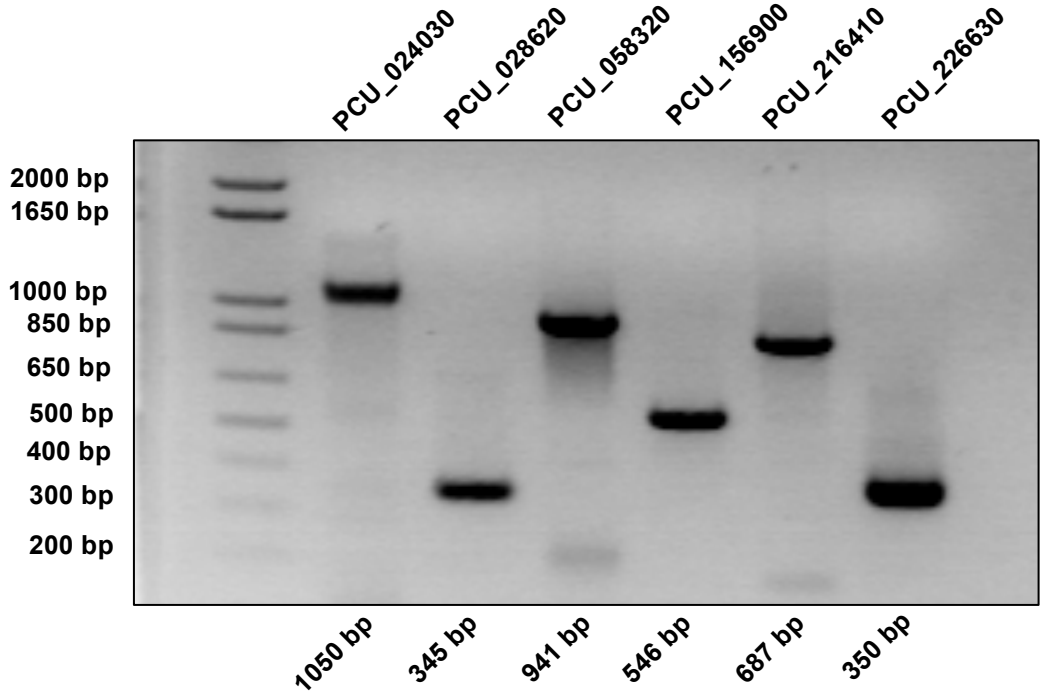


Figure 4. *Newly annotated* P. cubensis *annotated genes were amplified with reverse transcriptase (RT)-PCR using RNA extracted from sporangia and gene-specific primers.* RT-PCR products were separated on a 1.5% agarose gel with an Invitrogen 1 Kb Plus DNA ladder and stained with ethidium bromide. PCR products were sequenced, using the ABI 3730 Genetic Analyzer (Applied Biosystems, Foster City, CA, U.SA.) at the Michigan State University Research Technology Support Facility, to confirm that the sequence matched the predicted spliced sequence.



From this point on, we refer only to the expressed genome, which consists of 17,558 expressed genes identified based on analysis of the RNA-Seq datasets. Approximately 2 million sequenced fragments uniquely mapped as perfect but gapped alignments to the *P. cubensis* reference genome sequence. Using this output, we generated a list of approximately 20,000 putatively spliced sequences. Of these, 9,964 corresponded to an exon-exon junction sequence present in 6,366 expressed intron-bearing genes, which represents 65% of all possible splice junctions in the expressed transcriptome. Additionally, the majority of the predicted spliced sequences were supported by a high number of RNA-Seq reads. In 169 cases, reads corresponding to putatively spliced sequences aligned within annotated exons in which the intervening regions had sequence coverage that fell below operational thresholds for expression. Based on this, we updated the annotations of the regions to introns. The converse situation was also observed in which 700 annotated intronic sequences had high RNA-Seq coverage but no gap-aligned reads that spanned the annotated introns. These were reannotated as exons. Improvements to the genome annotation did not significantly change the average length of introns.

We also identified genes in the expressed genome that had signatures associated with genes involved in oomycete pathogenesis. We first analyzed translated sequences for predicted signal peptides. We predicted 1,736 and 2,139 secreted proteins from the translated sequences of the original annotated and updated expressed genomes, respectively. Of these, 1,320 were shared in both datasets. In the updated expressed genome, we identified an additional 819 genes predicted to encode secreted proteins that had not been annotated previously. In addition, 416 proteins predicted to be secreted

based on the original annotation were not identified in the expressed genome. Because we used the same software version to mine the original annotation and the expressed genome for genes that encode putative secreted proteins, we concluded that differences are likely a consequence of improvements via extensions at the 5′ end of gene models.

We next examined the putative secretome for proteins with RXLR and QXLR motifs (from hereafter, we will simply refer to these proteins as RXLR). We identified 59 additional protein sequences that contained the RXLR translocation motif, increasing the number of putative RXLR-effectors from 67 to 125 (http://dx.doi.org/10.7267/N9TD9V7M). Of these, 54 had been identified in the original genome annotation. Our improvements to the gene models resulted in the prediction of 59 new candidates; 13 previously annotated putative RXLR proteins were unsupported. In addition, the RNA-Seq data supports expression of 125 of these candidate RXLR effector proteins.

Genome-wide identification of alternatively spliced genes in *P. cubensis*

Using stringent criteria for the prediction of alternatively spliced transcript isoforms, we identified ~10,000 potential AS events associated with 4,205 genes (Fig. 3B; Table 2). Of these genes, 3,492, representing 20% of the expressed genome, had evidence for intron retention. A total of 1,812 genes (~10%), many of which also had evidence for intron retention, had evidence for either alternative 5′ of 3′ splice site selection (Fig. 3B). Surprisingly, despite the use of multiple *in silico* approaches, we failed to identify any incidences of exon skipping. A total of 24% of the expressed genome and 55% of the intron-bearing genes, had evidence for AS (Table 2).

Table 2. *Statistics for the* P. cubensis *transcriptome*.

Total Expressed Genes	17,558
Constitutively Spliced Genes	13,353 (76.1%)
Encodes a predicted protein (not secreted)	11,576 (65.9%)
Encodes a predicted secreted protein	1645 (9.4%)
Encodes a predicted effector	88 (0.5%)
Alternatively Spliced Genes	4205 (23.9%)
Encodes a predicted protein (not secreted)	3674 (20.9%)
Encodes a predicted secreted protein	494 (2.8%)
Encodes a predicted effector	37 (0.2%)

Next, we projected alternative splicing preferences and levels onto a phylogenetic tree with representative members of key lineages including Fungi, Stramenopiles, Plants, and Animals that had publicly available genome and/or transcriptome data (Fig. 5; Table 3). As a result of our improvements, the percent of intron-bearing genes in *P. cubensis* increased from 35% in the original annotation to 43% in the expressed genome, with an average of 0.86 per gene and 2.57 introns per intron-bearing gene. Among the members of the SAR group studied, *P. cubensis* has the fewest intron-bearing genes. Nevertheless, compared to the two members of the SAR group with available alternative splicing data available, *P. sojae* and *Phaeodactylum tricornutum*, *P. cubensis* has a higher portion of alternatively spliced genes (Bowler et al., 2008; Shen et al., 2011). The preference in AS types and the overall level of AS predicted for *P. cubensis* were more similar to those of land plants than any other lineage investigated. Plants favor intron retention but do vary in levels of alternative splicing. Specifically, 61% of multi-exonic *Arabidopsis thaliana* genes were spliced, compared to only 3% of those in the green algae *Chlamydomonas reinhardtii* (Labadorf et al., 2010; Marquez et al., 2012). In contrast, members of the Animal lineage prefer exon skipping and can have extremely high levels of AS which, in *H. sapiens*, is often tissue- or cell-specific (Pan et al., 2008; Keren et al., 2010). The fungi, with the exception of *Saccharomyces cerevisiae* and *Saccharomyces pombe*, have high (>75%) percentages of intron-bearing genes. However, they all have very low numbers of alternatively spliced genes, most of which were intron retention. The highest is *Aspergillus oryzae,* with 11.10% of intron-bearing genes alternatively spliced (Wang et al., 2010).

Figure 5. *Alternative splicing distribution across eukaryotic taxonomic branches*. Representative members of the Animal (*Homo sapiens* (95%), *Mus musculus* (28%), and *Caenorhabditis elegans* (62%)), Fungi (*Aspergillus oryzae* (11.1%), *Fusarium graminearum* (1.7%), and *Saccharomyces cerevisiae* (6%)), Plant (including land plants *Arabidopsis thaliana* (61%) and *Oryza sativa* (27.7%) and green algae (*Chlamydomonas reinhardtii* (3%)), and Chromalveolata (Stramenopiles *Phaeodactylum tricornutum* (7.32%), *Phytophthora sojae* (2.29%), and *Pseudoperonospora cubensis* (24%)) kingdoms are included. The numbers in parentheses indicated the percent of genes in each genome where AS events are found, represented by the circles in the figure (*enlarged, not to scale). Shaded bars represent the ratio of the total number of alternative splicing (AS) types found in the genomes of each species.

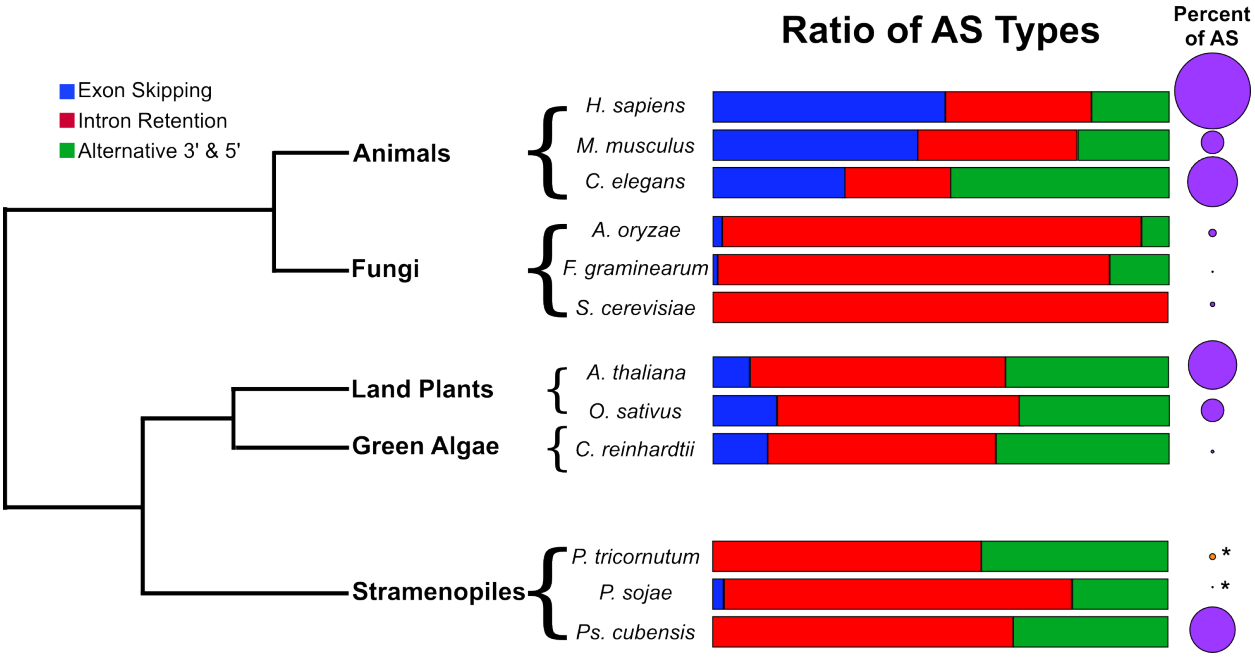


Table 3. *Genome annotation sources used to obtain information to calculate the percentage of single exonic and intron-bearing genes, ratios of alternative splicing types, and percentage of total alternative splicing in genomes*.

Organism	Database	Version	Reference
Saccharomyces cerevisiae	Ensembl[1]	EF4.70	Cherry et al., 1997
Saccharomyces pombe	Ensembl	ASM294v1.16	Wood et al., 2002
Fusarium graminearum	The Fusarium Comparative Sequencing Project[2]	PH-1_3	Ma et al., 2010, Zhao et al., 2013
Fusarium oxysporum f.sp. lycopersici	The Fusarium Comparative Sequencing Project	4287_2	Ma et al., 2010
Fusarium verticillioides	The Fusarium Comparative Sequencing Project	7600_3	Ma et al., 2010
Aspergillus oryzae	The *Aspergillus* Comparative Sequencing Project[2]	n/a	Galagan et al., 2005, Wang et al., 2010
Magnaporthe grisea	The Magnaporthe Comparative Sequencing Project[2]	70-15_6	Dean et al., 2005
Pseudoperonospora cubensis			Savory et al., 2012b
Hyaloperonospora arabidopsidis	Ensembl	HyaAraEmoy2_2.0.16	Baxter et al., 2010
Phytophthora infestans	Ensembl	ASM1429v1.16	Haas et al., 2009
Phytophthora ramorum	Ensembl	ASM14973v1.16	Tyler et al., 2006
Phytophthora sojae	Ensembl	ASM14975v1.16	Tyler et al., 2006
Pythium ultimum var. *ultimum*	Pythium Genome Database[3]	DAOM BRR144	Levesque et al., 2010
Thalassiosira pseudana	The Genome Portal of the Department of Energy Joint Genome Institute[4]	Thaps3	Armbrust et al., 2004
Phaeodactylum tricornutum	Ensembl	ASM15095v1.16	Bowler et al., 2008
Bigelowiella natans	The Genome Portal of the Department of Energy Joint Genome Institute	CCMP2755 v1.0	Curtis et al., 2012
Guillardia theta	The Genome Portal of the Department of Energy Joint Genome Institute	CCMP2712 v1.0	Curtis et al., 2012

Table 3 (cont'd)

Chlamydomonas reinhardtii	Phytozome[5]	V5.3.1	Merchant et al., 2007
Volvox carteri	Phytozome	Version 2	
Coccomyxa subellipsoidea C-169	Phytozome	V2.0	Blanc et al., 2009
Micromonas pusilla CCMP1545	Phytozome	V3.0	Worden et al., 2009
Physcomitrella patens	Phytozome	Version 1.6	Rensing et al., 2008
Arabidopsis thaliana	The Arabidopsis Information Resource[6]	TAIR10	Lamesch et al., 2011
Oryza sativa	Phytozome	MSU Release 7	Ouyang et al., 2006
Cucumus sativus	Phytozome	Phytozome 8.0	Huang et al., 2012
Caenorhabditis elegans	Ensembl	WBcel215.70	C. elegans Sequencing Consortium, 1999
Mus musculus	Ensembl	GRCm38.70	Church et al., 2009
Homo sapiens	Ensembl	GRCh37.70	Venter *et al.*, 2001

Database addresses: [1]http//:www.ensemblgenomes.org; [2]http://www.broadinstitute.org/; [3]http://www.pythium.plantbiology.msu.edu; [4]http://www. http://genome.jgi.doe.gov/; [5]http://www.phytozome.net; [6]http://www.tair.org

Experimental validation of predicted alternative splicing events

Because of the challenges associated with the small amount of biomass of an obligate biotrophic oomycete and its small intron sizes, we focused on a select number of highly expressed candidate genes that could be validated for AS using real-time PCR and/or RT-PCR. Candidate genes selected for validation were predicted to be expressed at every time point sampled and to have two isoforms. In Figure 6, the gene models predicted by RNA-Seq data (Fig. 6) were validated using real-time PCR with isoform-specific primers. PCU_123970.1 is predicted to encode a class IIb Aspartyl/Asparaginyl-tRNA synthetase (IPR002312) and lacks an identifiable signal peptide. RNA-Seq analysis suggested the gene gives rise to an alternative intron-retaining isoform (PCU_123970.2), which would introduce a premature termination codon and lead to a protein of only 65 amino acids. We could detect both the constitutively spliced and the intron-retaining isoforms during each of the developmental stages of the pathogen that were examined (Fig. 6A). However, there were clear differences in the expression patterns, consistent with inferences made from RNA-Seq analysis (Fig. 7). PCU_123970.1 showed higher relative expression than the intron-containing isoform, PCU_123970.2, at all stages examined (unpaired t-test P values for all pairwise comparisons at each time point ≤ 0.03). Furthermore, expression of PCU_123970.1 increased substantially from sporangia to its host-associated stages (unpaired t-test P values ≤ 0.02 for each pairwise comparison of sporangia to 2-8 dpi time points).

Figure 6. *Expression patterns of transcript isoforms of non-effector encoding genes exhibit different levels of relative expression over the course of infection.* Real-time PCR, using isoform-specific primers (indicated by arrows on the gene model), was used to quantify the expression of each transcript isoform relative to the *P. cubensis* internal control of the internal transcribed space region (ITS). A, Alternative splicing (AS) via the mechanism of intron retention in PCU_123970. The vertical bar in the gene model for PCU_123970.2 represents a premature stop codon in the predicted translated product. B, AS via intron retention in PCU_006210. C, AS via the use of an alternative 5′ splice site (ss) in PCU_092620. D, AS via the use of an alternative 3′ ss in PCU_065990. Real-time PCR is shown as the average with standard error of 3 biological replicates. dpi = days post-inoculation, Sp = sporangia

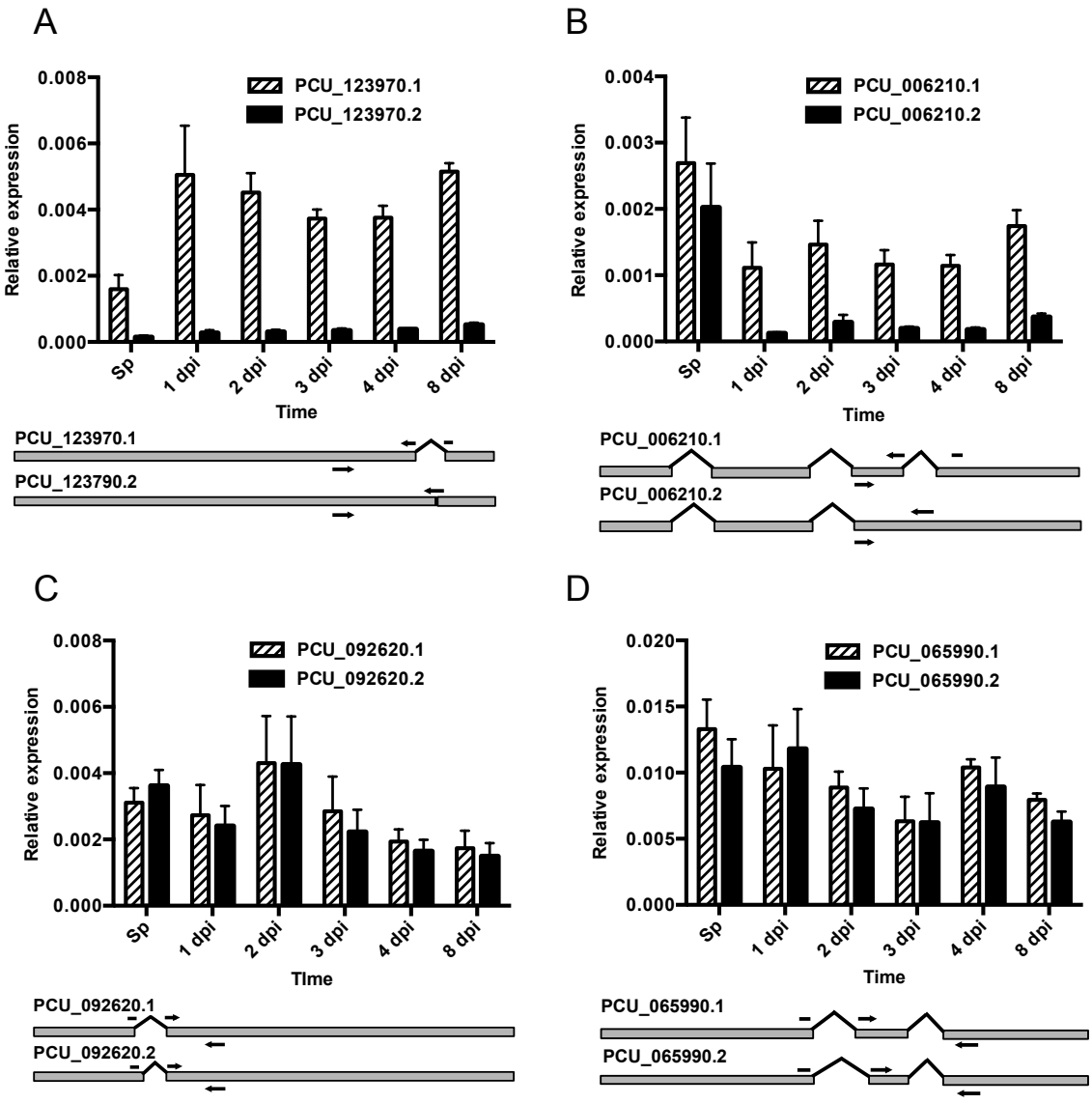
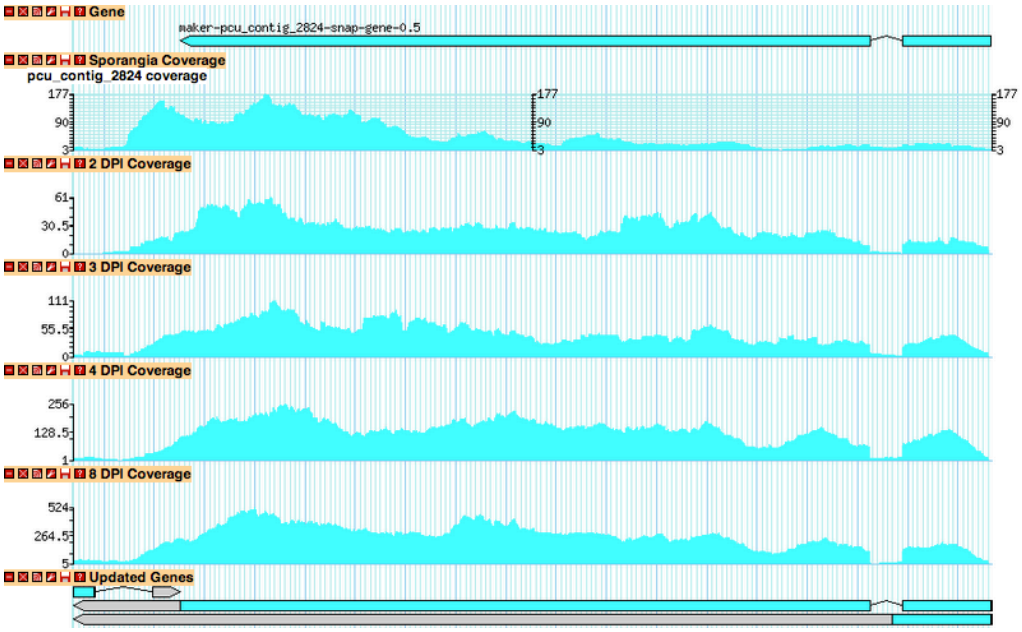


Figure 7. *RNA-Seq coverage as visualized using GBrowse for A, PCU_123970 showing intron retention, B, PCU_006210 showing intron retention, C, PCU_092620 showing use of an alternative 5′ splice site, and D, PCU_065990 showing use of an alternative 3′ splice site.*

A



B

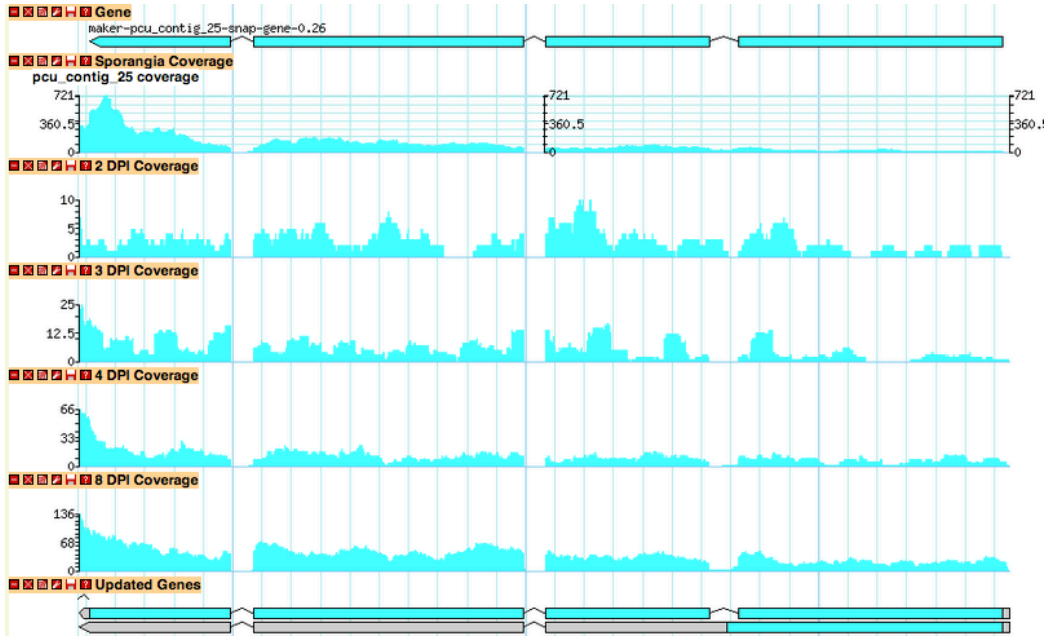
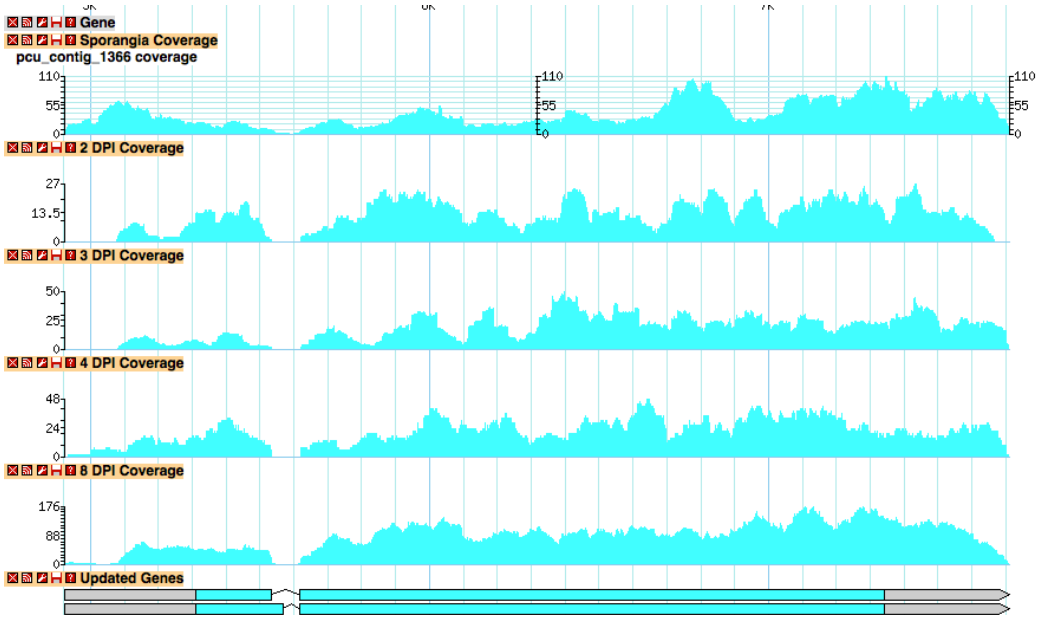
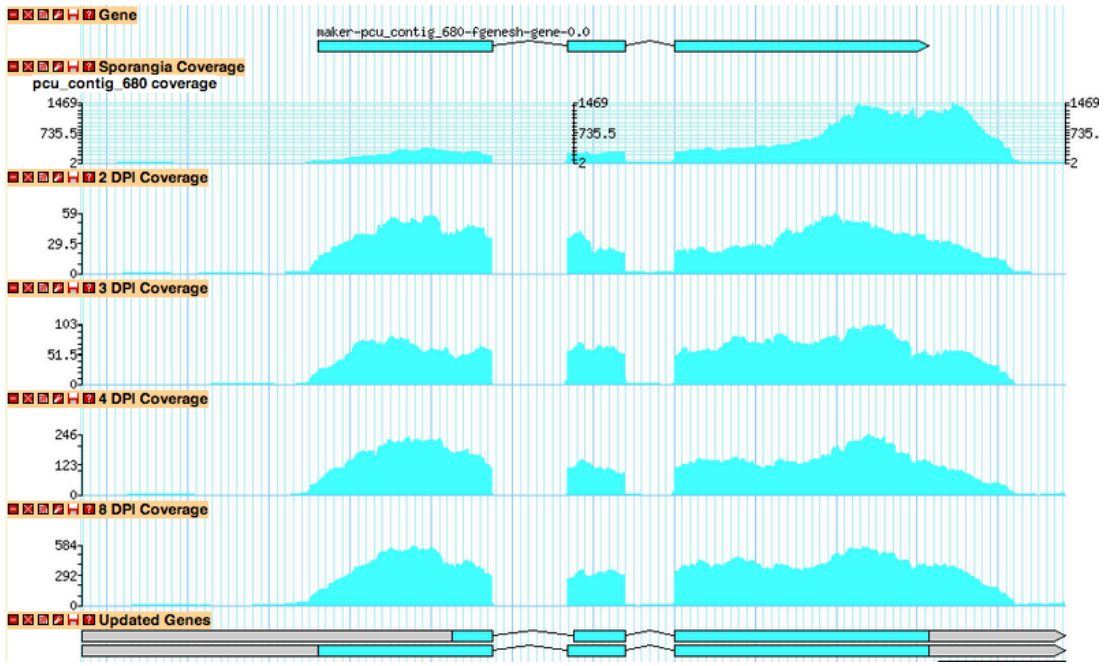


Figure 7 (cont'd)

C



D



A second gene (PCU_006210) that was tested also produces an alternative isoform that contains a retained intron (Fig. 6B). The translated product of PCU_006210.1 is a 956 amino acid protein with homology to a hypothetical protein of *Phytophthora parasitica* (GenBank: ETM43679.1). PCU_0006210.2 is predicted to encode a 313 amino acid product. The translated sequences for both isoforms are predicted to contain multiple BRCT (breast cancer carboxy-terminal; peptide- and phosphopeptide-binding modules present in proteins involved in DNA checkpoint controls and DNA repair; IPR001357) domains. Similar to the previous example, real-time PCR successfully confirmed the presence of both isoforms (Fig. 6B). In contrast, PCU_006210.2 is expressed in sporangia to relative levels comparable to that of PCU_006210.1 (unpaired t-test P value = 0.521). The expression of both isoforms decreased in *P. cubensis* during its host-associated stages, with PCU_006210.2 expression dropping to barely detectable levels.

We also used real-time PCR to confirm predicted alternative 5′ and 3′ ss usage. PCU_092620 has a predicted N-terminal lipin domain (IPR007651) and is similar to a predicted lipin-like protein from *Phytophthora infestans* T30-4 (XP_002903839.1). Both the fully spliced PCU_92620.1 and its 5' alternatively spliced isoform PCU_092620.2 appear to encode functional proteins of 649 and 660 amino acids, respectively. The two isoforms of PCU_092620 were relatively highly expressed in all stages we examined, with both showing highest levels of expression at 2 dpi (Fig. 6C). PCU_065990.1 is predicted to encode a 200 amino acid long member of the Rab5 family GTPase superfamily (IPR003579; GO:0005525), which function in endocytosis and early endosome fusion (Schwartz et al., 2008). The second isoform is alternatively spliced via a 3′ alternative splice site and

because of a frame shift, gives rise to a putative 91 amino acid protein. Both isoforms of PCU_065990 also showed relatively high expression throughout all stages tested but steadily decreased in expression over the course of sampling (Fig. 6D).

Twelve percent of the genes with evidence for alternative splicing are predicted to encode a secreted protein (Table 2). PCU_071910.1 is predicted to encode a secreted glutathione peroxidase (IPR000889; GO:0004602). The gene has two introns and based on RNA-Seq coverage, we predict that PCU_071910.2, which retains the second intron will result in a truncated, 141 amino acid long protein. The relative expression levels of PCU_071910.1 and PCU_071910.2 were similar in sporangia (Fig. 8A). During host infection, the PCU_071910.1 isoform was significantly more expressed at 8 dpi relative sporangia (t-test P value = 0.035) and PCU_071910.1 was significantly more expressed (t-test P value < 0.05) relative to PCU_071910.2 levels from 2-8 dpi (Fig. 8A). These data are consistent with the RNA-Seq results which showed a similar decrease in the coverage of reads in the intron in samples collected during host infection relative to the sample collected from sporangia (Fig. 8A). Using Sanger sequencing, we confirmed the sequences for both transcript isoforms. To test whether AS affected localization of PCU_071910 both isoforms were fused to cyan fluorescent protein (CFP) and expressed transiently in *Nicotiana benthamiana.* We did not observe any change in the localization of the protein isoforms (Fig. 9). PCU_115490 has a single intron and the translated sequence of the spliced form is predicted to encode a member of the endonuclease/exonucleases/peroxidase family (IPR005135) (Fig. 8B). Use of RT-PCR and real-time PCR confirmed the presence of an alternative intron-retaining form, which is predicted to lack all of the functional domains, but retain its signal peptide

sequence. Over the course of infection, PCU_115490.1 shows a rapid increase in expression, peaking at 2 dpi, and gradually declining over the course of infection (Fig. 8B). The intron retaining form, PCU_115490.2, shows a consistent low level of expression. Thus, there is a dramatic change in the ratio of the predicted functional to putative non-functional isoforms for this gene, as PCU_115940.1 consists of about 35% of the total PCU_115940 gene expression in the sporangia and over 95% of the total expression by 1 dpi (Fig. 8B). We also fused CFP to both isoforms of PCU_115490 and expressed transiently the genes in *N. benthamiana.* Similar to previous observations, changes in the localization of their protein isoforms were not observed (Fig. 9).

Figure 8. *Genes encoding secreted gene products displaying alternatively splicing via intron retention.* RNA-Seq coverage of each gene is shown as visualized through GBrowse. Coverage is shown for sporangia, 2dpi, 3 dpi, 4 dpi, and 8 dpi. Gene models are provided for each isoform with a vertical bar indicating the location of a stop codon. Real-time PCR was used to determine the relative expression of alternatively spliced isoforms as normalized to the *P. cubensis* internal transcribed spacer region, ITS. Data is reported as the average relative expression and standard error from 3 biological replicates. The isoform products amplified from reverse-transcriptase (RT)-PCR were validated by Sanger sequencing and separated on a 1.5% agarose gel with a 1 kb plus DNA ladder. Genomic *P. cubensis* DNA is shown as a positive PCR control. A, PCU_071910; B, PCU_115490. dpi = days post-inoculation, Sp = sporangia, DNA = PCR of sporangia genomic DNA.

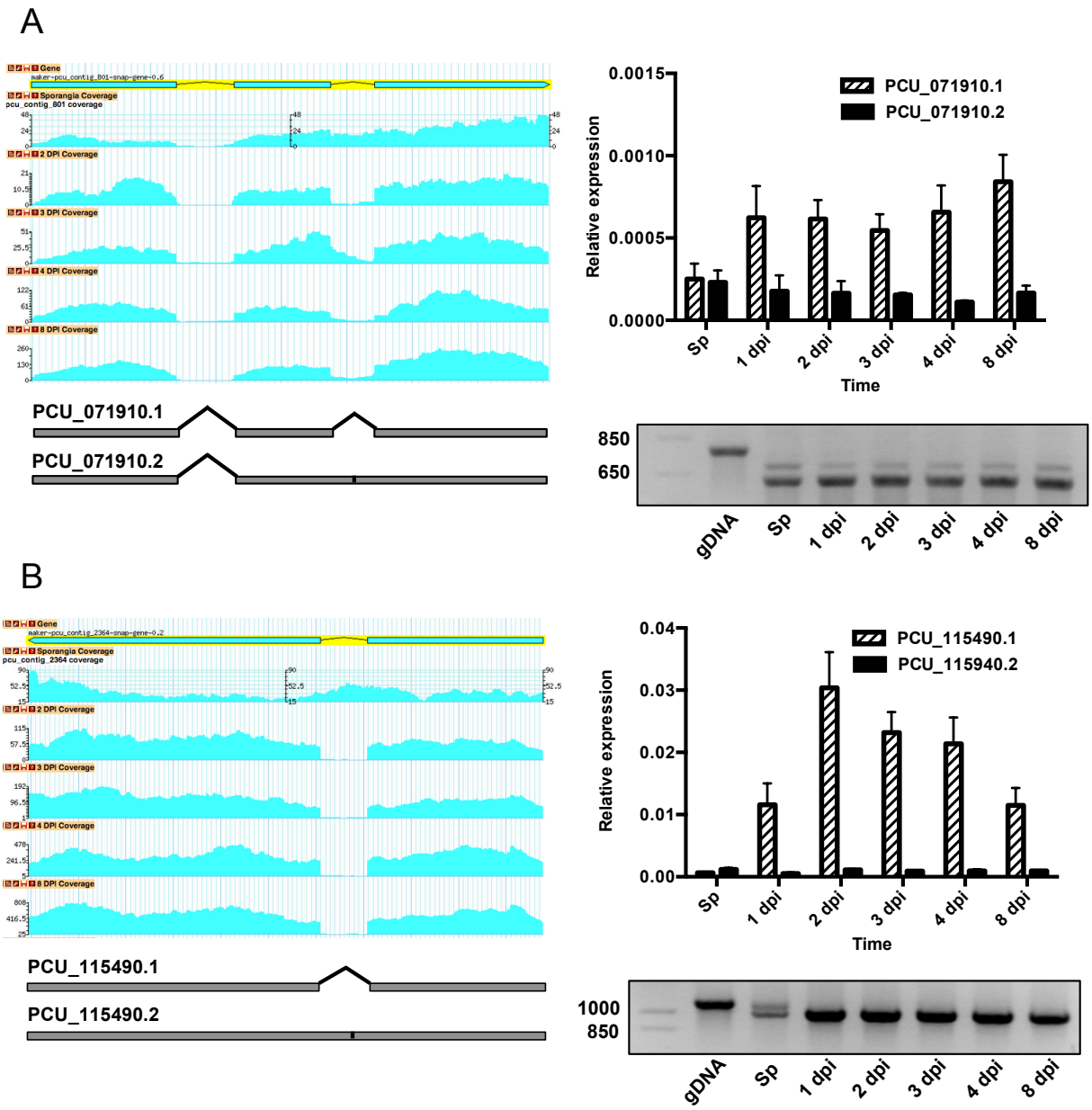
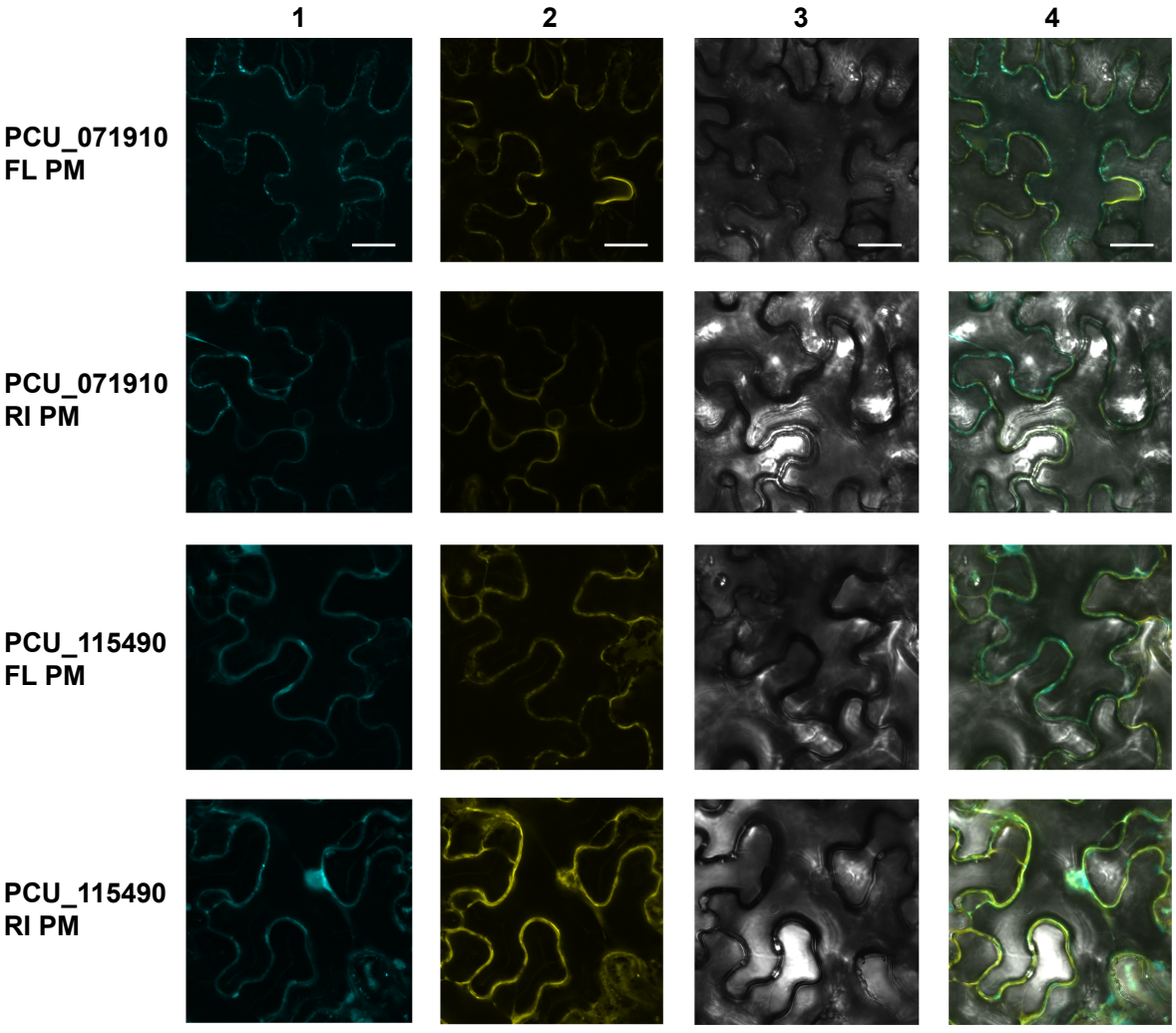
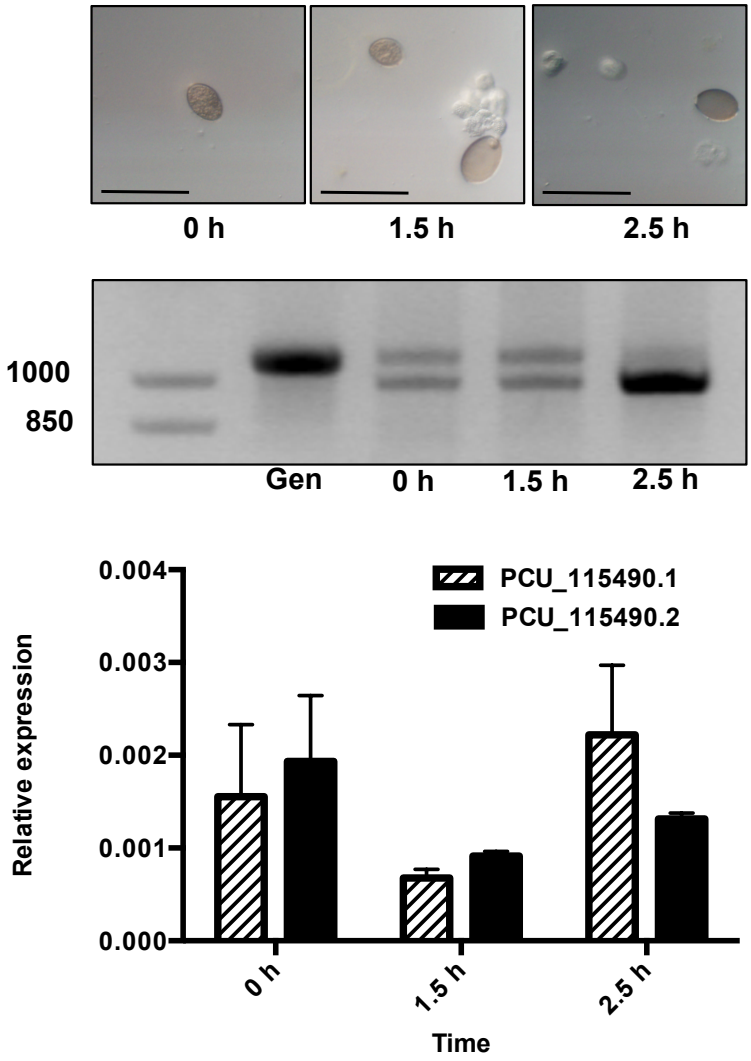


Figure 9. *Proteins encoded by full-length (FL) and retained intron (RI) transcripts localize to the plasma membrane (PM).* Full-length and retained intron cDNAs of PCU_071910 and PCU_115490 were cloned into pVKH18En6gw-cCFp. *Agrobacterium tumefaciens* carrying the CFP-fusion constructs were infiltrated into *Nicotiana benthamiana* (column 1) and co-infiltrated with *A. tumefaciens* carrying a plasma membrane marker *At*PIP2A-YFP (column 2). An Olympus Fluoview 1000 laser scanning confocal microscope with a 40x oil objective was used for brightfield imaging (column 3). Image overlay of the CFP and YFP channels is shown (column 4). Scale bars represent 20 μm.



The transition from sporangia to zoospore may be a key development change that correlates with changes in alternative splicing patterns. This hypothesis was developed based on a recent publication showing that transcripts containing introns were spliced after certain environmental cues, which allowed for rapid development to occur post-transcriptionally (Boothby et al., 2013). We therefore examined the changes for PCU_115490 at three different points during the sporangia to zoospore transition period that occurs under exposure to water which triggers germination (Fig. 10). In sporangia (0 h), and during the early stages of zoospore formation (1.5 h), both isoforms were expressed at similar levels. At the latest stage of zoospore formation, PCU_115490.1 accumulated to higher levels than PCU_115490.2.

Figure 10. *PCU_115490 shows differential alternative splicing during pathogen developmental transitions from sporangia to zoospore***.** Representative developmental stages are shown in images collected over the time course. Images were taken using an Olympus IX71 inverted light microscope. Scale bars represent 200 μm. Alternative isoforms were amplified using reverse transcriptase (RT)-PCR with *Pseudoperonospora cubensis* RNA collected at 0, 1.5, and 2.5 hours (h) after exposure of sporangia to water. gDNA, PCR of genomic DNA from sporangia; 0 h, sporangia; 1.5 h, beginning of zoospore germination; 2.5 h, fully germinated zoospores. Alternative isoform expression was quantified using real-time PCR with transcript normalization to the *P. cubensis* internal transcribed spacer region, ITS. Real-time PCR is shown as the average with standard error of 3 biological replicates. dpi = days post-inoculation, Sp = sporangia



Previously, we reported on the role of AS in generating a candidate effector protein (Savory et al., 2012b). Our approach described in this study successfully identified *Psc*RXLR1 (PCU_071030) as a candidate RXLR with evidence for AS (http://dx.doi.org/10.7267/N9TD9V7M). We further suggest that nearly 1% of the alternatively spliced genes are predicted to encode effectors. Within this class of genes, only intron retention was identified as an alternative splicing mechanism. PCU_240880 encodes a predicted RXLR effector. The retained intron inferred from RNA-Seq occurs in its 3′ UTR and is not predicted to affect function, but it may regulate expression (Fig. 11A). Both of the gene isoforms show similar patterns of expression, with peak levels observed at 2 dpi (Fig. 11A). More importantly, the ratios of PCU_240880.1 to PCU_240880.2 changed, as more PCU_240880.1 could be detected at higher levels in *P. cubensis* during its host-associated stages (> 60%), compared to in sporangia (about 40%). PCU_143920 is also predicted to encode an RXLR effector. This gene has two introns and data suggests that only the second intron is retained (Fig. 11B). The relative expression of PCU_143920.1 significantly decreased (t-test P-value < 0.05) in levels from sporangia to the host-associated stages (Fig. 11B). Likewise, the relative expression of PCU_143920.2 also decreased but its overall relative expression levels were consistently low. Lastly, PCU_031660 is predicted to encode a QXLR effector. It has two introns, and inferences from RNA-Seq data suggest the first intron is retained, with strongest evidence being from sporangia tissue (Fig. 11C). The relative expression of PCU_031660.1 significantly increased (t-test P value < 0.05) from sporangia to the host-associated stages of *Ps cubensis*. In contrast, the expression of PCU_031660.2 remained relatively low and constant.

Figure 11. *Predicted alternatively spliced genes encoding secreted gene products predicted to contain a signal peptide and an RXLR/QXLR motif.* RNA-Seq coverage of each gene is shown as visualized through GBrowse. Coverage is shown for sporangia, 2 days post inoculation (dpi), 3 dpi, 4 dpi, and 8 dpi. Gene models are provided for each isoform, with a vertical bar indicating the location of a stop codon. Real-time PCR was used to determine the relative expression of alternatively spliced isoforms as normalized to the *P. cubensis* internal transcribed spacer region, ITS. Arrows on the models indicate the binding position of isoform-specific primers. Data is reported as the average relative expression and standard error from 3 biological replicates. A, PCU_240880 encodes a predicted RXLR effector and is alternative spliced via the mechanism of intron retention in the 3′ untranslated region (UTR). B, PCU_143920 encodes a predicted RXLR effector and is alternative spliced in the coding sequence (CDS) of the gene. C, PCU_031660 encodes a predicted QXLR effector whose CDS is alternatively spliced.

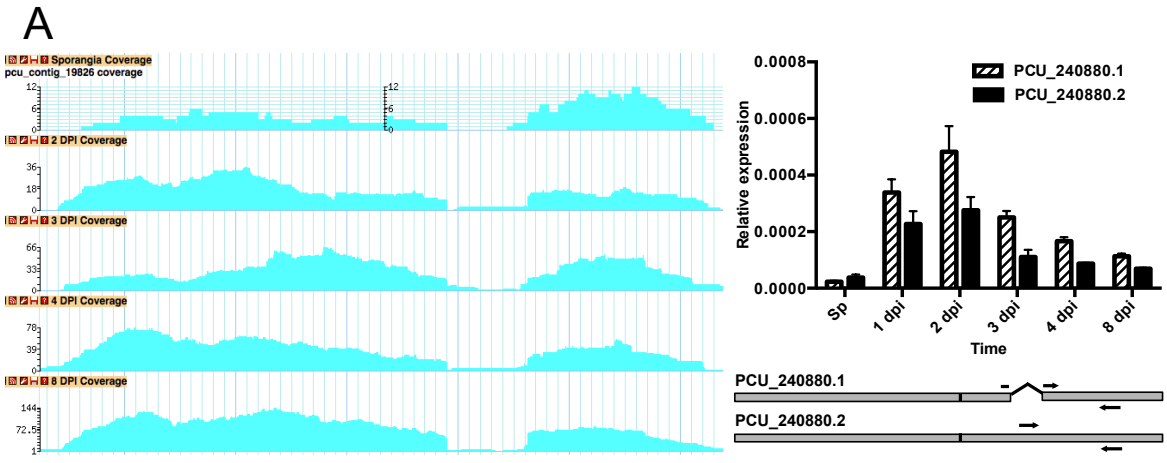
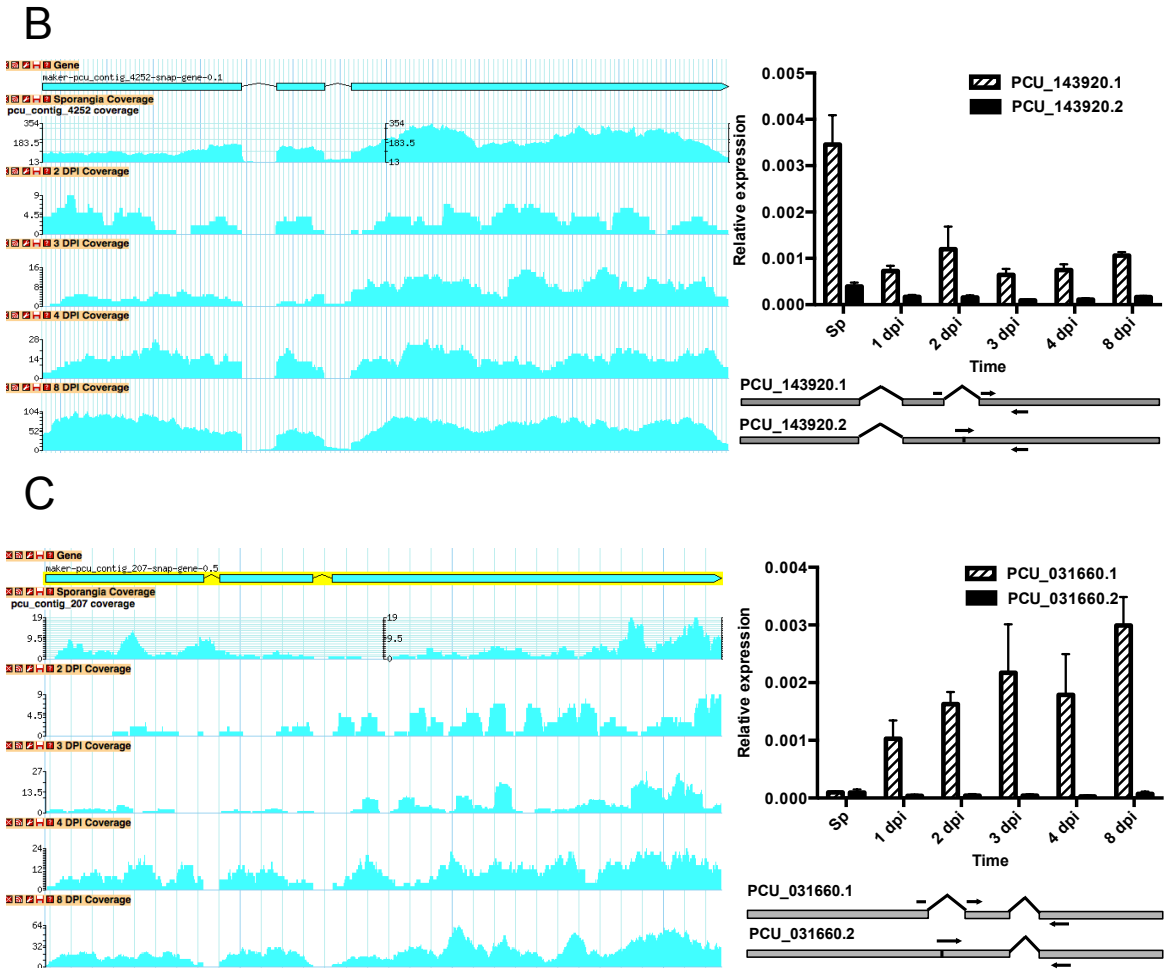


Figure 11 (cont'd)



Discussion

In this study, we used RNA-Seq to characterize the transcriptome of *P. cubensis*. We sampled multiple time points during the infection cycle of *P. cubensis* on cucumber. This enabled us to capture gene expression in multiple pathogen structures, including sporangia, encysted zoospores, appressoria, hyphae, and haustoria (Savory et al., 2012a). By using the Illumina HiSeq we were able to obtain approximately 40 million sequenced fragments that uniquely aligned to *P. cubensis*, more than doubling the number of usable sequences that were previously generated using the Illumina GAII (Savory et al., 2012a). Inferences based on unique alignments revealed approximately 17,500 expressed genes including approximately 4,000 new genes. This enabled us to both improve the existing genome annotation and provided sufficient coverage for the first transcriptome-wide view into alternative splicing of an obligate oomycete plant pathogen.

The original draft sequence of *P. cubensis* consists of a large number of contigs, 23,519 annotated genes, and few predicted gene models. With the RNA-Seq reads, we could support 57% of those genes and also identified over 4,000 previously unannotated genes. Moreover, more than half of the previously annotated genes were modified to include additional coding sequences, UTRs, and/or modifications to their gene models.

The current draft genome sequence, improved using RNA-Seq datasets, has 27,591 annotated genes of which 64% are predicted to be expressed. While this may seem low, it is likely due to the fact that not all stages of the *P. cubensis* lifestyle, such as sporulation and sexual reproduction, were represented in the sampled time points. An additional

confounding factor is the relatively low depth of transcriptome sequencing caused by the high percentage of host transcripts in each sample, which again highlights one of the challenges of studying this pathosystem. Additionally, or alternatively, the low percent of expressed genes could also suggest that the original genome sequence included contaminating sequences and that many of the annotated genes were erroneously predicted. Nonetheless, our analysis indicates that the original draft genome of *P. cubensis* did not likely suffer from missing gene sequences. We were able to align nearly 10 million sequenced fragments to the reference sequence that were previously unannotated. Most of the unaligned read sequences were determined to be the result of sequencing errors and not predicted to be from *P. cubensis*. Thus, the use of RNA-Seq led to dramatic changes to the reference sequence and contributed to a higher quality resource for advancing our understanding of this critically important pathogen.

The extent and contribution of alternative splicing to transcriptome and proteome diversity in Stramenopiles is less characterized than in fungi, plants, and mammals. Results from this study suggested that AS occurs with a much higher frequency than had been originally predicted. In *P. cubensis*, we identified 24% of the expressed genes as having evidence for alternative splicing (Table 2). These data are not consistent with previous work that predicted lower rates of AS in plant-pathogenic oomycetes (McGuire et al., 2008; Shen et al., 2011). However, a critical difference between the current work and previously published studies is the technology used in transcriptome sampling. The current work, which relied on RNA-Seq provided a significantly deeper and more comprehensive data set from which the *P. cubensis* transcriptome can be modeled for AS, in spite of the

overwhelming number of reads that were derived from the cucumber host. Indeed, RNA-Seq-based analysis of *Bigelowiella natans,* a chlorarachniophyte algae, revealed that AS occurs at levels remarkably higher than previously observed for other unicellular eukaryotes (Curtis et al., 2012). This observation is significant because oomycetes are more closely related to *B. natans* and plants than to fungi (Simpson & Roger, 2004; Hackett et al., 2007; Seidel et al., 2011; Burki et al., 2012). Thus, it is distinctly possible that AS in oomycetes occurs at rates more in line to that of plants.

The higher frequency of intron retention observed in *P. cubensis* is consistent with previous studies of AS in oomycetes (McGuire et al., 2008; Shen et al., 2011), and is consistent with high intron retention rates identified in plants (Filichkin et al., 2010). Intron retention has been proposed as a mechanism of auto-regulation for controlling both the timing and expression of the genome. The splicing of introns can be regulated in a stage- and stress-dependent manner, as has been demonstrated for influencing the expression of meiotic genes and ribosomal protein-encoding genes in response to amino acid starvation (Juneau et al., 2007; Pleiss et al., 2007; Wilhelm et al., 2008). In the fern *Marsilea vestita*, intron retention, has been shown to be important for regulating gamete development, whereby transcripts with retained introns prevent the early translation of a protein but upon developmental cues, are rapidly spliced to trigger developmental changes (Boothby et al., 2013). In P. cubensis, notable changes in splicing patterns occurred during the transition from sporangia to host-associated life stages. This is consistent with a model suggesting that alternative splicing can suppress translation and maintain the oomycete in a quiescent but poised state that, upon splicing, contributes to a rapid transition into an infective stage

(Figs. 6, 8, and 10).

A proportion of the alternative spliced genes could lead to different protein isoforms. In *Phytophthora sojae*, alternatively spliced Crinkler effector transcripts have been reported (Shen et al., 2011). Similarly, an effector gene from *Glomus intraradices* is alternatively spliced to generate multiple isoforms that vary in the number of C-terminal repeats (Kloppholz et al., 2011). In *Colletotrichum lindemuthianum* and *Botrytis cinerea*, exon skipping results in the exclusion of a zinc finger domain from the STE12 transcription factor (Hoi et al., 2007; Schamber et al., 2010). As a function of the host-pathogen interaction in the case of *C. lindemuthianum*, expression of the alternatively spliced truncated form of STE12 has been shown to negatively affect pathogenicity through a down-regulation in expression coincident with appressorium formation (Hoi et al., 2007). In future studies, it will be important to identify and characterize genes that give rise to different protein isoforms that contribute to the virulence of *P. cubensis*.

The relevance of the extraordinarily high numbers of estimated alternatively spliced transcripts has been repeatedly challenged. At the core of the debate is the fate of the alternatively spliced transcript. Many isoforms are predicted to encode nonfunctional proteins and regardless of functionality, most lack support from expressed protein variants in databases (Hegyi et al., 2011). As a consequence, sloppiness in splicing (i.e., stochastic noise) is often invoked to explain AS and the generation of large numbers of seemingly nonproductive transcripts (Melamud & Moult, 2009; Pickrell et al., 2010). The contributions of stochastic noise and technical biases in estimates of AS in the RNA-Seq

datasets described herein are difficult to quantify. In humans, for example, splicing errors are conservatively estimated to occur in approximately 2% of the transcripts from the average genes (Pickrell et al., 2010). However, this estimate was made using unannotated spliced sequences that were poorly supported, represented by only 1.7% of the gap-aligned RNA-Seq read sequences. In this study, a number of filters were implemented to limit poorly supported AS events and reduce the misclassification of noise as relevant AS events. As a result of our stringent analysis, the alternative 5′ and 3′ splice site type events were supported by 6.3% of the gap-aligned RNA-Seq read sequences. The prediction of intron retention was more challenging and could be a consequence of sequencing pre-mRNAs. However, in PCU_071910 and PCU_143920, retention was only observed for one of their two introns, arguing against artifacts. Conservatively, because AS is often developmentally regulated or stress-induced, we hypothesize that the number of splice events in the *P. cubensis* transcriptome and in the transcriptomes of other oomycetes is actually higher than predicted. To fully define the role of AS in the regulation of pathogen development, in response to environment, and in mediation of host defense signaling, future studies are needed to examine the splicing variation under differing conditions; these include abiotic stresses (e.g., changes in temperature) and biotic stresses such as variation of host range and differing levels of host resistance. As presented herein, the current work provides a foundation for such studies, which will help reveal the true extent of AS and better understand the role of AS in oomycete development and virulence of plant hosts.

Materials and Methods

Plant growth and *Pseudoperonospora cubensis* inoculation

Uninfected cucumbers cv. "Vlaspik" were grown at 22°C in 16h light/8h dark cycles. *P. cubensis* isolate MSU-1 was propagated on cucumber as previously described (Tian et al., 2011). Time course experiments were performed as described by Savory et al. (2012a), in which 10 μL of a 1 x 10^5 suspension of *P. cubensis* sporangia/ml in distilled water were drop-inoculated onto the abaxial leaf surface of four- to five-week-old cucumber plants. Infected plants were kept in the dark at 100% humidity for 24 hours and were subsequently maintained at near 100% humidity at 22°C in 12 light/12h dark cycles (Savory et al., 2012a). Infected leaf samples were collected at the same time each day of the time course using a #3 cork borer, were flash-frozen in liquid nitrogen, and stored at -80 °C until sample processing.

Next-generation sequencing

Previously prepared samples (Savory et al., 2012a) were re-sequenced more deeply for this study. The libraries from biological replicates of cucumber leaves infected with *P. cubensis* for 2, 3, 4, and 8 days post inoculation (dpi) were re-sequenced using 100-mer paired-end (PE) sequencing on four channels of an Illumina HiSeq at the Michigan State University Research Technology Support Facility. Previously sequenced *P. cubensis* sporangia sample reads (Savory et al. 2012a) were used for bioinformatic analyses. The short reads are

linked to Bioproject PRJNA247812 and also available from SRA (SRP042019).

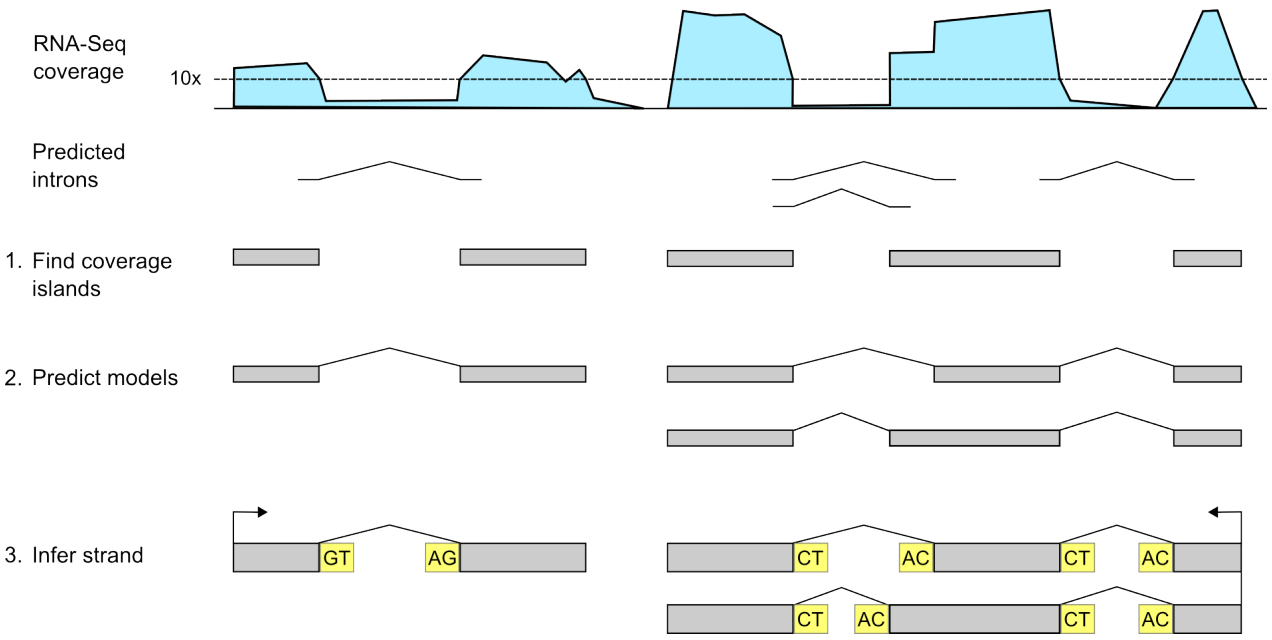
Alignment and processing of RNA-Seq read sequences

Bowtie version 0.12.7 was used to align RNA-Seq read sequences with up to two mismatches to the *C. sativus L.* (genome and contig accessions ACHR00000000-ACHR01059995) and *P. cubensis* (genome and contig accessions AHJF00000000-AHJF01035539) reference sequences (Huang et al., 2009; Savory et al., 2012a). Paired-end (PE) RNA-Seq reads were aligned independently. The quality of the reads was determined by assessing mismatches as a factor of nucleotide position. Based on the bias of mismatches towards the ends of the sequencing reads, between 5 and 31 nucleotides were trimmed from the ends. Using a custom C++ pipeline, high quality pairs were first identified based on the criteria that one read must align per strand and the alignments must be between -60 to 200 nucleotides (nt) apart. A second step using an iterative process that adjusted the distance based on sizes of intervening annotated introns was employed to identify pairs that exceeded the distance criterion because of an annotated intron. SuperSplat modified to output SAM formatted data was used for gapped alignments (Bryant et al., 2010). Alignment parameters included a perfect match on either side of the gap with a minimum of at least 15 nt on one side of the gap and a gap of 20 to 4,000 nt in length. All alignments were stored in BAM format using SAMtools (Li et al., 2009). Both PE and a subset of single, unpaired RNA-Seq reads were used, which are collectively referred to as “sequenced fragments”.

Gene improvement and discovery

Custom Perl and Python scripts were developed to update and identify genes based on coverage and predicted splice junction sequences from the aggregated sequenced fragments that uniquely aligned to the *P. cubensis* reference genome sequence. Genes in the original annotation were extended if adjacent positions had ≥ 10X coverage or had a gap-aligned fragment sequence that was consistent with it being on the sense strand of the gene (Fig. 12).

Figure 12. *Schematic of the method for using RNA-Seq reads to identify new expressed gene features.* The RNA-Seq reads were aligned to the original genome reference of *P. cubensis* (AHJF00000000-AHJF01035539). Reads that mapped to regions lacking annotated features were used to identify new genes. RNA-Seq coverage: example of RNA-Seq alignments (with the 10X read-depth threshold indicated) and having no gaps > 5 nt that were not supported by gap alignments. Predicted introns: gap-aligned fragment sequences used to infer introns. 1. Find coverage islands: coalescence of aligned sequences in regions exceeding 300 nt in length. 2. Predict models: association of expressed features based on gap-aligned fragment sequences. 3. Infer strand: the dinucleotide sequence of the gap aligned fragment sequences was used to predict the strand of the gene.



Comparative analysis of alternative splicing among eukaryotes

Custom Perl scripts were used to calculate the percentage of single exonic and intron-bearing genes, ratios of AS types, and percentage of total AS from representative eukaryotic genomes. Genome annotation information was obtained from the appropriate genome databases.

Prediction of candidate secreted and effector proteins

SignalP (version 4.0; using the noTM network with a D-cutoff of 0.34) was used to predict signal peptide cleavage sites in translated sequences to identify candidate secreted proteins (Petersen et al., 2011). A string search for "RXLR" within a distance of 90 amino acids of a predicted signal peptide cleavage site was used to query the candidate secreted proteins for putative effector proteins.

Prediction of alternatively spliced isoforms

To identify potential intron retention events, only introns in which the coverage of the flanking 30 nt ≥ 10X were considered as a binary measure of expression. The "local intron coverage" ratios were calculated by normalizing the coverage of introns to the coverage of the flanking 30 nt. Retained introns were expected to exceed a threshold of 13%, which was set based on it being nearly two standard deviations to the right of the mean local intron coverage of all expressed introns. We also required intron retention events to be

supported by reads that spanned the spliced intron as well as those that aligned to the exon-intron border sequence. To distinguish an alternatively spliced sequence from an erroneous annotation, we required additional correspondence of putatively spliced sequences to an intronic region that had support for at least one other spliced variant affecting the same region (e.g., the annotated splice junction and an alternatively spliced sequence). Finally, the local intron coverage ratio could not exceed 80% to minimize the categorization of mis-annotated exonic sequences as intron retention events. All of the predicted intron retention events were confirmed using an independent approach. Intron retention events were inferred based on the unambiguous mapping of reads to exon-intron junctions flanking an intron and the gapped mapping of reads to exon-exon junctions that spanned the intron.

RNA and DNA isolation

The RNeasy Plant Mini Kit (Qiagen, Germantown, MD, U.S.A.) was used to extract total RNA from sporangia and flash-frozen leaf tissue. An on-column DNase treatment (Qiagen) was used to remove contaminating genomic DNA from the RNA samples. The Qiagen DNeasy Plant Mini kit was used to extract genomic DNA from sporangia.

Reverse transcriptase (RT)-PCR analysis

First-strand cDNA was synthesized from RNA isolated from sporangia and infected plant tissue, using oligo-dT primers and the first strand cDNA synthesis kit (USB Affymetrix,

Santa Clara, CA, U.S.A.). RT-PCR was performed using GoTaq (Promega, Madison, WI, U.S.A.). DNA primer sequences used for amplification are listed in Table 4. The following cycling parameters were used: 1 hold at 95°C (1 min); 50 cycles of 95°C (30 sec), 55°C (1 min), 72°C (3 min); 1 hold at 72°C (10 min). RT-PCR products were separated on a 1.5% agarose gel with an Invitrogen 1 kb Plus DNA ladder and stained with ethidium bromide. Genomic *P. cubensis* DNA was used as a positive PCR control. PCR products were sequenced, using the ABI 3730 Genetic Analyzer (Applied Biosystems, Foster City, CA, U.S.A.) at the Michigan State University Research Technology Support Facility, to confirm that the sequence matched the predicted spliced sequence.

Table 4. *List of primer sequences used for reverse-transcriptase PCR (RT-PCR) of alternatively spliced transcripts from* Pseudoperonospora cubensis*.*

Primer Name	Sequence (5´→3´)
PCU_024030 F	CAAAGACCGCAGTCCAAGGATATTG
PCU_024030 R	CTGGTGTGGCGGTACGAACGAAG
PCU_028620 F	GACCTACTGAAGAAGCTCTATCGACATG
PCU_028620 R	GTAGAACAGATTGACGGTCGATTTGC
PCU_058320 F	CAGGCGACAAGAAGCGAAAGAAAGC
PCU_058320 R	GTTGCCGTGTTGGCGTAACTTGGA
PCU_156900 F	CGAAGTCGACGGGTTGGATTGAC
PCU_156900 R	CCTCAACTCTCTCTTCTCGTGAC
PCU_216410 F	GATAACATCAGCGACTGACTTGTGACC
PCU_216410 R	CTTCTGGATTCGCGCTCGGTCTG
PCU_226630 F	GAATTCTCATTGTGTCGATATCGGC
PCU_226630 R	CGAAGTAGCGCAGTCCTCTCG
PCU_071910 CACC F	CACCATGCGGTTGCTACTCCTGAT
PCU_071910 NS R	CAGCTCATCGTGAGCGTCAC
PCU_115490 CACC F	CACCATGGCCTTTGTGAAAGGACT
PCU_115490 NS R	CAGCCTTTCTTCCTTAATAACGCG

Quantitative real-time PCR

First-strand cDNA was synthesized using the USB first-strand cDNA synthesis kit with random hexamer primers (USB Affymetrix), as described above, using 1 μg of total RNA. Samples were prepared using the HotStart SYBR Green qPCR Master Mix (2x; USB) and the DNA primers listed in Table 5. Quantitative real-time PCR was performed using a Mastercycler ep Realplex real-time PCR machine (Eppendorf AG, Hamburg, Germany). *P. cubensis*-specific internal transcribed spacer (ITS) was also amplified (Tian et al., 2011) and used to normalize for the level of *P. cubensis* in infected tissue samples. The following cycling parameters were used: 1 hold of 95°C for 2 min, 40 cycles of 95°C (15 sec), 56°C (15 sec), and 72°C (30 sec). Relative expression for each splice variant was calculated where relative expression = $2^{(-\Delta Ct)}$ and where $\Delta Ct = Ct_{\text{gene of interest}} - Ct_{\text{ITS,}}$ as previously described (Porter et al., 2012). A melt curve analysis was run for each primer set to ensure that a unique product was amplified. Error bars represent the standard error from three biological replicates. Data were analyzed and processed using Prism (GraphPad Software, Inc., San Diego, CA, U.S.A.). Unpaired two-tailed t-test analyses were performed using the mean and standard error of the mean for pairwise sample comparisons with a significance P value cutoff of 0.05.

Table 5. List of primer sequences used for real-time PCR of alternatively spliced transcripts from *Pseudoperonospora cubensis.*

Primer Name	Sequence (5´→3´)
PCU_123970.1 F	CGTCTACAAGTCGATAAAAGACCT
PCU_123970.2 F	CTTGAAGTTACGATGTGCAATGAATATG
PCU_123970 Rev	TCTGCTCACGCAGCGTGATAAAG
PCU_006210.1 F	GTTGACCGATGACAGATTGGAGGC
PCU_006210.2 F	GCTCTCATTGATATTGACACCAATTTG
PCU_006210.1 Rev	CACCTTGGCCTCGACGGTCT
PCU_092620.1 F	TCCACATGAAGTTAGGTGCTGCT
PCU_092620.2 F	GAAGAAGAGATGTTAGGTGCTGCTG
PCU_092620 Rev	CGAGACATTGTTCTCTAAAGGAC
PCU_065990.1 F	GAAAAGTACCACAGCCTGGCAC
PCU_065990.2 F	GGCCAAGAAAAGCCTGGCACCC
PCU_065990 Rev	GCTCTTGTTACCTGCAATGGC
PCU_071910.1 F	GCTCTTTTCTAAGGTGGATGTGAATG
PCU_071910.2 F	GGATGAGATGCAAAGAGGGGAGG
PCU_071910 Rev	CAATCTCCAGTGGCGACGTGG
PCU_115490.1 F	CGTTTGCGCTGAGCGATACAC
PCU_115490.2 F	GCTTGCATCGTGCGGAAGTCG
PCU_115490 Rev	CAGCATGCACACTACTCGATAAG
PCU_143920.1 F	TGCTTTCAAGTTTTCTCGGCTG
PCU_143920.2 F	GTGATATGGCAATTCTTACCTTG
PCU_143920 Rev	GCCATGCGTATTTGGGATCAC
PCU_240880.1 F	CCACAAGTCGAAGAATCAAACA
PCU_240880.2 F	GCTTACTTCTGATGCTGACTC
PCU_240880 Rev	CCGAATTCTTGGTCTCGTCG
PCU_031160.1 F	GAAATGCAGTGTGTCGCACG
PCU_031160.2 F	GTGATTCTGATACTTGCTATTGC
PCU_031160 Rev	GCTGGTGTTACGCCACGC

Transient gene expression and protein localization in *Nicotiana benthamiana*

PCR products of the constitutively and alternatively spliced transcripts (to the first premature stop codons) of PCU_071910 and PCU_115490 were amplified in RT-PCRs and ligated to the pENTR/D-TOPO Gateway entry vector (Invitrogen, Carlsbad, CA, U.S.A.). The cDNAs were recombined into pVKH18En6gw-cCFp (Tian et al., 2011), using LR Clonase (Invitrogen). DNA sequences were confirmed using the ABI 3730 Genetic Analyzer (Applied Biosystems).

Plasmid constructs were electroporated into *Agrobacterium tumefaciens* strain C58C1. For *in planta* localization, 4-6 week old *N. benthamiana* were co-infiltrated with strains of *Agrobacterium* carrying pVKH18En6gw-cCFp variants paired with a strain carrying *At*PIP2A-YFP (Nelson et al., 2007; Savory et al., 2012b). The latter encodes a translational fusion between a plasma membrane marker and Yellow Fluorescent Protein (YFP). Images were collected 48 hours post-infiltration using the Olympus FV1000 laser scanning confocal microscope (Olympus America, Center Valley, PA, U.S.A.). Images for negative controls were also collected to ensure that there was no fluorescence crossover; cells transiently expressing *At*PIP2A-YFP were not detectable when visualized using the CFP channel settings used for image collection.

Sporangia time course sampling

Ungerminated sporangia emerging from sporangiophores on leaf surfaces did not contact

water, and thus did not germinate prior to flash-freezing; these samples were considered 0 hour (0 h) sporangia. Flash-frozen leaf tissues harboring 0 h sporangia were shaken in flasks of distilled water at 120 rpm at room temperature to dislodge sporangia. Samples were filtered through a 70-micron filter to remove debris, and centrifuged at 3220 *g* to pellet the sporangia. Sporangia samples from multiple tubes were pooled and centrifuged again at 4000 rpm in a microcentrifuge to further pellet the sporangia for RNA extraction. Isolated samples were visually inspected using light microscopy to ensure that flash-frozen sporangia did not germinate during the sample preparation process. To collect sporangia for zoospore germination, fresh sporulating leaves were processed as described above except infected leaf tissues were not flash-frozen. Biological samples were collected after 1.5 hours and 2.5 hours in sterile distilled water. Germination of sporangia was observed using an Olympus IX71 inverted light microscope. After 1.5 hours in water, 30-50% of the sporangia were germinating into zoospores and after 2.5 hours post-water exposure nearly 100% of the sporangia had germinated. Three biological replicates of this time course were collected.

Acknowledgements

The authors would like to thank Siobhan Cusack for assistance with sample preparation. EAS is supported by USDA NIFA post-doctoral fellowship #2013-67012-21139. Work in the laboratory of JHC is supported by funding from the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM104977. Research in the laboratory of BD is supported by funding from the Michigan State

University Rackham Foundation, Michigan State University Project GREEEN (GR13-007), and a grant from the United States Department of Agriculture Specialty Crops Research Initiative (2011-51181-30661).

Author contributions

Performed experiments: Aly. B, Alex B., J.S.C., E.A.S.; Analyzed data: Aly. B., Alex B., J.S.C., E. A. S., J.H.C., B.D.; Wrote the manuscript: Aly. B., E.A.S., J.H.C, B.D.

Author-recommended internet resources

Updated *Pseudoperonospora cubensis* genome annotation is archived at Oregon State University Libraries (http://dx.doi.org/10.7267/N9TD9V7M). *P. cubensis* gene list, including gene updates, annotation, and alternative splicing predictions (http://dx.doi.org/10.7267/N9TD9V7M).

REFERENCES

REFERENCES

Adhikari, B. N., Savory, E. A., Vaillancourt, B., Childs, K. L., Hamilton, J. P., Day, B., and Buell, C. R. 2012. Expression profiling of *Cucumis sativus* in response to infection by *Pseudoperonospora cubensis PLoS ONE* 7: e34954. Published online.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., Zhou, S., Allen, A. E., Apt, K. E., Bechner, M., Brzezinski, M. A., Chall, B. K. , Chiovitti, A. , Davis, A. K., Demarest, M. S., Detter, J. C., Glavina, T., Goodstein, D., Hadi, M. Z., Hellsten, U., Hildebrand, M., Jenkins, B. D., Jurka, J. Kapitonov, V. V., Kröger, N., Lau, W. W. Y., Lane, T. W., Larimer, F. W., Lippmeier, C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M. S., Palenik, B., Pazour, G. J., Richardson, P. M., Rynearson, T. A., Saito, M. A., Schwartz, D. C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F. P., and Rokhsar, D. S. 2004. The genome of the diatom *Thalassiosira pseudonana:* ecology, evolution, and metabolism. *Science* 306: 79-86.

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

Beakes, G.W., Glockling, S.L., and Sekimoto, S. 2012. The evolutionary phylogeny of the oomycete "fungi". *Protoplasma* 249: 3-19.

Blanc, G., Agarkova, I., Grimwood, J., Kuo, A., Brueggeman, A., Dunigan, D., Gumon, J., Ladunga, I., Lindquist, E., and Lucas, S. 2012. The genome of the polar eukaryotic microalga *Coccomyxa subellipsodiea* reveals traits of cold adaption. *Gen. Biol.* 13: R39. Published online.

Boothby, T. C., Zipper, R. S., van der Weele, C. M., and Wolniak, S. M. 2013. Removal of retained introns regulates translation in the rapidly developing gametophyte of *Marsilea vestita*. *Dev. Cell* 24: 517-529.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., and Otillar, R. P. 2008. The *Phaedodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456:239-244.

Bryant, D. W., Shen, R., Priest, H. D., Wong, W. K., and Mockler, T. C. 2010. Supersplat--spliced RNA-seq alignment. *Bioinformatics* 26: 1500-1505.

Burkhardt, A. and Day, B. 2013. A genomics perspective on cucurbit-oomycete interactions. *Plant Biotech.* 30: 265-271.

Burki, F., Okamoto, N., Pombert, J-F., and Keeling, P. J. 2012. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *P. Roy. Soc. B-Biol. Sci.* 279: 2246-2254.

C. elegans Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 11:2012-118.

Chen, L., Tovar-Corona, J. M., and Urrutia, A.O. 2012. Alternative splicing: A potential source of functional innovation in the eukaryotic genome. *Int. J. Evol. Biol.* 2012: 1110.

Cherry, J.M., Ball, C., Weng, S., Juvik, G., Schmidt, R., Adler, C., Dunn, B., Dwight, S., Riles, L., and Mortimer, R.K. 1997. Genetic and physical maps of *Saccharomyces cerevisiae. Nature* 6632: 67-73.

Church, D.M., Goodstadt, L., Hillier, L.W., Zody, M.C., Goldstein, S., She, X., Bult, C.J., Agarwala, R., Cherry, J. L., DiCuccio, M., Hlavina, W., Kapustin, Y., Meric, P., Maglott, D., Birtie, Z., Marques, A. C., Graves, T., Zhou, S., Teague, B., Potamousis, K., Churas, C., Place, M., Herschleb, J., Runnheim, R., Forrest, D., Amos-Landgraf, J., Schwartz, D. C., Cheng, Z., Lindblad-Toh, K., Eichler, E. E., Ponting, C. P., and Mouse Genome Sequencing Consortium. 2009. Lineage-specific biology revealed a finished genome assembly of the mouse. *PLoS Biol.* 7: e1000112. Published online.

Curtis, B.A,. Tanifuji, G., Burki, F., Gruber, A., Irimia, M., Maruyama, S., Arias, M. C., Ball, S. G., Gile, G. H., Hirakawa, Y., Hopkins, J. F., Rensing, S. A., Schmutz, J., Symeonidi, A., Elias, M., Eveleigh, R. J., Herman, E. K., Klute, M. J., Nakayama, T., Oborník, M., Reyes-Prieto, A., Armbrust, E. V., Aves, S. J., Beiko, R. G., Coutinho, P., Dacks, J. B., Durnford, D. G., Fast, N. M., Green, B. R., Grisdale, C. J., Hempel, F., Henrissat, B., Höppner, M. P., Ishida, K., Kim, E., Kořený, L., Kroth, P. G., Liu, Y., Malik, S. B., Maier, U. G., McRose, D., Mock, T., Neilson, J. A., Onodera, N. T., Poole, A. M., Pritham, E. J., Richards, T. A,. Rocap, G., Roy, S. W., Sarai, C., Schaack, S., Shirato, S., Slamovits, C. H., Spencer, D. F., Suzuki, S., Worden, A. Z., Zauner, S., Barry, K., Bell, C., Bharti, A. K., Crow, J. A., Grimwood, J., Kramer, R,. Lindquist E., Lucas, S., Salamov, A., McFadden, G. I., Lane, C. E., Keeling, P. J., Gray, M. W., Grigoriev, and I. V., Archibald J.M. 2012. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* 492: 59-65.

Dean, R. A., Talbot, N. J., Ebbole, D. J., Farman, M. L., Mitchell, T. K., Orbach, M. J., Thon, M., Kulkarni, R., Xu, J., Pan, H., Read, N. D., Lee, L. H., Carbone, I., Brown, D., Oh, Y. Y., Donofrio, N., Jeong, J. S., Soanes, D. M., Djonovic, S., Kolomiets, E., Rehmeyer, C., Li, W., Harding, M., Kim, S., Lebrun, M. H., Bohnert, H., Coughlan, S., Butler, J., Calvo, S., Ma, L. J., Nichol, R., Purcell, S., Nusbaum, C., Galagan, J. E., and Birren, B. W. 2005. The genome sequence of the rice blast fungus *Magnaporthe grisea. Nature* 434: 980-986.

Filichkin, S.A., Priest, H. D., Givan, S. A., Shen, R., Bryant, D. W., Fox, S. E., Wong, W. K., and Mockler, T. C. 2010. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Gen. Res.* 20: 45-58.

Galagan, J. E., Calvo, S. E., Cuomo, C., Ma, L.J., Wortman, J. R., Batzoglou, S., Lee, S.I., Basturkmen, M., Spevak, C. C., Clutterbuck, J., Kapitonov, J. Jurka, J., Scazzocchio, C., Farman, M., Butler, J., Purcell, S., Harris, S., Braus, G. H., Draht, I., Busch, S., D'Enfert, C., Bouchier, C., Goldman, G. H., Bell-Pedersen, D., Griffiths-Jones, S., Doonan, J. H., Yu, J., Vienken, K., Pain, A., Freitag, M., Selfker, E. U., Archer, D. B., Peñalva, M. A., Oakley, B. R., Momany, M., Tanaka, T., Kumagi, T., Asai, K., Machida, M., Nierman, W. C., Denning, D. W., Caddick, M., Hynes, M., Paoletti, M, Fischer, R., Miller, B., Dyer, P., Sachs, M. S., Osmani, S. A., and Birren, B. W. 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438: 1105-1115.

Guan, Q., Wu, J., Zhang, Y., Jiang, C., Liu, R., Chai, C., and Zhu, J. 2013. A DEAD box RNA helicase is critical for pre-mRNA splicing, cold-responsive gene regulation, and cold tolerance in Arabidopsis. *Plant Cell* 25: 342-356.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozhurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M, Conti, L. , Constanzo, S., Ewan, R., Fahlgran, N, Fischbach, M. A., Fugelstad, J., Gilroy, E. M., Gnerre, S., Green, P. J., Grenville-Briggs, L. J., Griffith, J., Grünwald, N. J., Horn, K., Horner, N. R., Hu, C. H., Huitema, E., Jeong, D. H., Jones, A. M., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro C. A., O'Neill, K., Ospina-Giraldo, M., Pinzón, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., savidor, A., Schornack, S., Schwartz, D. C., Schumann, U. D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D. J., Sykers, S., Thines, M., van de Vondervoort, P. J., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C. , van West, P., Ristaino, J, Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans. Nature* 46: 393-398.

Hackett, J.D., Yoon, H.S., Li, S., Reyes-Prieto, A., Rümmele, S.E., and Bhattacharya, D. 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol. Biol. Evol.* 24: 1702-1713.

Hegyi, H., Kalmar, L., Horvath, T., and Tompa, P. 2011. Verification of alternative splicing variants based on domain integrity, truncation length and intrinsic protein disorder. *Nucl. Acids Res.* 39: 1208–1219.

Hoi, J. W.S., Herbert, C., Bacha, N., O'Connell, R., Lafitte, C., Borderies, G., Rossignol, M., Rougé, P., and Dumas, B. 2007. Regulation and role of a STE12-like transcription factor from the plant pathogen *Colletotrichum lindemuthianum*. *Mol. Microbiol.* 64: 68-82.

Holmes, G., Wehner, T., and Thornton, A. 2006. An old enemy re-emerges. *Am. Vegetable Grow.* Feb, 14-15.

Hoskins, A. A. and Moore, M. J. 2012. The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* 37: 179-188.

Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., Lucas, W.J., Wang, X., Xie, B., Ni, P. Ren, Y., Zhu, H., Li, J., Lin, K, Jin, W., Fei, Z., Li, G., Staub, J., Kilian, A., van der Vossen, E., Wu, y., Gui, J., He, J., Jia, Z., Ren, Y., Tian, G., Lu, Y., Ruan, J., Qian, E., Wang, M., Huang, Q., Li, B., Xua, Z., Cao, J., Asan, Wu, Z., Cai, Q., Bai, Y., Zhao, B., Han, Y., Li, Y., Li, X., Wang, S., Shi, Q., Liu, S., Cho, W. K., Kim, J. Y., Xu, Y., Heller-Uszyska, K., Miao, H., Cheng, Z., Zhang, S., Wu, J., Yang, Y., Kang, H., Li, M., Liang, H., Ren, X., Shi, Z., Wen, M., Jian, M., Yang, H., Zhang, G., Yang, Z., Chen, R., Liu, S., Li, J., Ma, L., Liu, S., Li, J., Zhou, Y., Zhao, J., Fang, X., Li, G., Fang, L., Li, Y., Liu , D., Zheng, H, Zhang, Y., Qin, N., Li, Z., Yang, G., Yang, S., Bolund, L., Kristiansen, K., Zheng, H., Li, S., Zhang, X., Yang, H., Want, J., Sun, R., Zhang, B., Jiang, S., Wang, J., Du, Y., and Li, S. 2009. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41: 1275-1281.

Irimia, M., Penny, D., and Roy, S. W. 2007. Coevolution of genomic intron number and splice sites. *Trends Genet.* 23: 321-325.

James, A. B., Syed, N. H., Bordage, S., Marshall, J., Nimmo, G. A., Jenkins, G. I., Herzyk, P., Brown, J. W. S, and Nimmo, H. G. 2012. Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. *Plant Cell* 24: 961-981.

Juneau, K., Palm, C., Miranda, M., and Davis, R. W. 2007. High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. *Proc. Natl. Acad. Sci. U.S.A.* 104: 1522-1527.

Kalsotra, A. and Cooper, T. A. 2011. Functional consequences of developmentally regulated alternative splicing. *Nat. Rev. Genet.* 12: 715-729.

Kamoun, S. 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu. Rev. Phytopathol.* 44: 41-60.

Keren, H., Lev-Maor, G., and Ast, G. 2010. Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* 11: 345-355.

Kim, E., Goren, A., and Ast, G. 2008. Alternative splicing: current perspectives. *BioEssays* 30: 38-47.

Kloppholz, S., Kuhn, H., and Requena, N. 2011. A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy. *Curr. Biol.* 21: 1204-1209.

Labadorf, A., Link, A., Thomas, J., Reddy, A. S., and Ben-Hur, A. 2010. Genome-wide analysis of alternative splicing in *Chlamydomonas reinhardtii*. *BMC Genomics* 11:114.

Lamesch, P., Berardini, T. Z., Donghui, L., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., and Huala, E. 2011. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucl. Acids Res.* 40:D1202-D1210.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R., 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078-2079.

Ma, L. J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M. J., Di Pietro, A., Dufresne, M., Freitag M., Grabherr, M., Henrissat, B., Houterman, P. M., Kang, S., Shim, W. B., Woloshuk, C., Xie, X., Xu, J. R., Antoniw, J., Baker, S. E., Bluhm, B. H., Breakspear, A., Brown, D. W., Butchko, R. A., Chapman, S., Coulson, R., Coutinho, P. M., Danchin, E. G., Diener, A., Gale L. R., Gardiner, D. M., Goff, S., Hammond-Kosack, K. E., Hiburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C. D., Koehrsen, M., Kumar, L., Lee, Y .H., Li, L., Manners, J. M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S. Y., Proctor, R. H., Regev, A., Ruiz-Roldan, M. C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D. C., Turgeon, B. G., Wapinkski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C. A., Kistler, H. C., and Rep, M. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature* 464: 367-373.

Marquez, Y., Brown, J. W. S., Simpson, C., Barta, A., and Kalyna, M. 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.* 22: 1184-1195.

Mastrangelo, A. M., Marone, D., Laidò, G., De Leonardis, A. M., and De Vita, P. 2012. Alternative splicing: Enhancing ability to cope with stress via transcriptome plasticity. *Plant Sci.* 185-186: 40-49.

Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W. C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science* 12:245-250.

McGuire, A. M., Pearson, M. D., Neafsey, D. E., and Galagan, J. E. 2008. Cross-kingdom patterns of alternative splicing and splice recognition. *Gen. Biol.* 9: R50.

Melamud, E. and Moult, J. 2009. Stochastic noise in splicing machinery. *Nucl. Acids Res.* 37: 4873–4886.

Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Marechal-Drouard, L., Marshall, W. F., Qu, L. H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C. L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M, Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J. P., Riaño-Pachón, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova, K., Chen, C. J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S. Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Homick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martinez, D., Ngua, W.C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. 2006. The *Chlamydomonas* genome reveals the evolution of key animal and plant function. *Science* 318: 245-250.

Meyer, M. and Vilardell, J. 2008. The quest for a message: budding yeast, a model organism to study the control of pre-mRNA splicing. *Brief. Func. Gen. Prot.* 8: 60–67.

Nelson, B. K., Cai, X., and Nebenführ, A. 2007. A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J.* 51: 1126-1136.

Nicholson, P., Yepiskoposyan, H., Metze, S., Orozco R. Z., Kleinschmidt, N., and Mühlemann, O. 2009. Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cell Mol. Life Sci.* 67: 677-700.

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thbaud-Nissen, F., Malek, R.L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, J., and Buell, C. R. 2007. The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucl. Acids Res.* 35:D883-887.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genet.* 40: 1413-1415.

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8: 785-786.

Pickrell, J. K., Pai, A. A., Gilad, Y., and Pritchard, J. K. 2010. Noisy splicing drives mRNA isoform diversity in human cells. *PLoS Genet.* 6: e1001236.

Pistoni, M., Ghigna, C., and Gabellini, D. 2010. Alternative splicing and muscular dystrophy. *RNA Biol.*7: 441-452.

Pleiss, J. A., Whitworth, G. B., Bergkessel, M., and Guthrie, C. 2007. Rapid, transcript-specific changes in splicing in response to environmental stress. *Mol. Cell* 27: 928-937.

Porter, K., Shimono, M., Tian, M., and Day, B. 2012. Arabidopsis actin-depolymerizing factor-4 links pathogen perception, defense activation and transcription to cytoskeletal dynamics. *PLoS Pathog.* 8: e1003006. Published online.

Rayson, S., Arciga-Reyes, L., Wootton, L., De Torres-Zabala, M., Truman W, Graham, N., Grant, M., and Davies, B. 2012. A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in *Arabidopsis thaliana* NMD mutants. *PLoS ONE* 7: e31917. Published online.

Reddy, A. S. N. 2007. Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu. Rev. Plant Biol.* 58: 267-294.

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

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Linquist, E. A., Kamisugi, Y. Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin-I, T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W. B., Barker, E., Bennetzen, J. L., Blankenship, R., Cho, S. H., Dutcher, S. K., Estelle, M., Fawcett, J. A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K. A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D. R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P. J., Sanderfoot, A., Schween, G., Shiu, S. H., Stueber, K., Theodoulou, F. L., Tu, H., Van de Peer, Y., Verrier, P. J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A. C., Hasebe, M., Lucas, S., Mishler, B. D., Reski, R., Grigoriev, I. V., Quatrono, R. S., and Boore, J. L. 2008. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science* 4: 64-69.

Sanchez, S. E., Petrillo, E., Kornblihtt, A. R., and Yanovsky, M. J. 2011. Alternative splicing at the right time. *RNA Biol.* 8: 954-959.

Savory, E. A., Adhikari, B. N., Hamilton, J. P., Vaillancourt, B., Buell, C. R., and Day, B. 2012a. mRNA-Seq analysis of the *Pseudoperonospora cubensis* transcriptome during cucumber (*Cucumis sativus* L.) infection. *PLoS ONE* 7: e35796. Published online.

Savory, E. A., Granke, L. L., Quesada-Ocampo, L. M., Varbanova, M., Hausbeck, M. K., and Day, B. 2010. The cucurbit downy mildew pathogen *Pseudoperonospora cubensis*. *Mol. Plant Pathol.* 12: 217-226.

Savory, E. A., Zou, C., Adhikari, B. N., Hamilton, J. P., Buell, C. R., Shiu, S-H., and Day, B. 2012b. Alternative splicing of a multi-drug transporter from *Pseudoperonospora cubensis* generates an RXLR effector protein that elicits a rapid cell death. *PLoS ONE* 7: e34701. Published online.

Schamber, A., Leroch, M., Diwo, J., Mendgen, K., and Hahn, M. 2010. The role of mitogen-activated protein (MAP) kinase signalling components and the Ste12 transcription factor in germination and pathogenicity of *Botrytis cinerea*. *Mol. Plant Pathol.* 11: 105-119.

Schöning, J. C., Streitner, C., Page, D. R., Hennig, S., Uchida, K., Wolf, E., Furuya, M., and Staiger, D. 2007. Auto-regulation of the circadian slave oscillator component AtGRP7 and regulation of its targets is impaired by a single RNA recognition motif point mutation. *Plant J.* 52: 1119-1130.

Schwartz, S. L., Cao, C. Pylypenko, O., Rak, A., and Wandinger-Ness, A. 2008. Rab GTPases at a glance. *J. Cell Sci.* 120: 3905-3910.

Seidl, M.F., Van den Ackerveken, G., Govers, F., and Snel, B. 2011. A domain-centric analysis of oomycete plant pathogen genomes reveals unique protein organization. *Plant Physiol.* 155: 628-644.

Shen, D., Ye, W., Dong, S., Wang, Y., and Dou, D. 2011. Characterization of intronic structures and alternative splicing in *Phytophthora sojae* by comparative analysis of expressed sequence tags and genomic sequences. *Can. J. Microbiol.* 57: 84-90.

Simpson, A. G. B. and Roger, A. J. 2004. The real 'kingdoms' of eukaryotes. *Curr. Biol.* 14: R693-R696.

Sultan, M., Schulz, M. H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A., Parkhomchuk, D., Schmidt, D., O'Keeffe, S., Haas, S., Vingron, M., Lehrach, H., and Yaspo, M. L. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321: 956-960.

Syed, N. H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J. W. S. 2012. Alternative splicing in plants – coming of age. *Trends Plant Sci.*17: 616-623.

Tian, M., Win, J., Savory, E., Burkhardt, A., Held, M., Brandizzi, F., and Day, B. 2011. 454 genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Mol. Plant Microbe Interact.* 24: 543-553.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H. Y., Aerts, A., Arredondo, F. D., Baxter, L., Bensasson, D., Beynon, J. L., Chapman, J., Damasceno, C. M., Dorrance, A. E., Dou, D., Dickerman, A. W., Dubchak, I. L., Garbelotto, M., Gijzen, M., Gordon, S. G., Govers, F., Grunwald N. J., Huang, W., Ivors, K. L., Jones, R. W., Kamoun, S., Krampis, K., Lamour, K. H., Lee, M. K., McDonald, W. H., Medine, M., Meijer, H. J., Nordberg, E. K., Maclean, D. J., Ospina-Giraldo, M. D., Morris, P. F., Phuntumart, V., Putnam, N. H., Rash, S., Rose, J. K., Sakihama, Y., Salamov, A. A., Savidor, A., Scheuring, C. F., Smith, B. M., Sobral, B. W., Terry, A., Torto-Alalibo, T.A., Win, J., Xu, J., Zhang, H., Grigoriev, I. V., Rokhsar, D. S., and Boore, J. L. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313: 1261-1265.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., and Holt, R. A. 2001. The sequence of the human genome. *Science* 16:1304-1351.

Wang, B. B. and Brendel, V. 2006. Genomewide comparative analysis of alternative splicing in plants. *Proc. Natl. Acad. Sci. U.S.A.* 103: 7175-7180.

Wang, B., Guo, G., Wang, C., Lin, Y., Wang, X., Zhao, M., Guo, Y., He, M., Zhang, Y., and Pan, L. 2010. Survey of the transcriptome of *Aspergilus oryzae* via massively parallel mRNA sequencing. *Nucl. Acids Res.* 38: 5075-5087.

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M., Armstrong, M. R., Grouffaud, S., van West, P., Chapman, S., Hein, I., Toth, I. K., Pritchard, L., and Birch, P. R. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* 450: 115-119.

Wilhelm, B. T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C. J., Rogers, J., and Bähler, J. 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453: 1239-1243.

Wood, V., Gwilliam, R., Rajandream, M. A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., Basham, D., Bowman, S., Brooks, K., Brown, D., Borwn, S., Chillingworth, T., Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P., Feltwell, T., Fraser, A., Gentles, S., Goble, A., Hamlon, N., Harris, D., Hidalgo, J., Hodgson, G., Holroyd, S., Hornsby, T., Howarth, S., Huckle, E. J., Junts, S., Jagels, K., James, K., Jones, L., Jones, M., Leather, S., McDonald, S., McLean, J., Mooney, P., Moule, S., Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neill, S., Pearson, D., Quail, M. A., Rabbinowitsch, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M., Squares, S., Stevens, K., Taylor, K., Talyor, R. G., Tivery, A., Walks, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Grymonprez, B., Weltjens, I., Vanstreels, E., Rieger, M., Schäfer, M., Müller-Auer, S., Gabel, C., Fuchs, M., Düsterhöft, A., Fritzc, C., Holzer, E., Moestl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T. M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E. Dréano, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S. J., Xiang, Z., Hunt, C., Moore, K., Hurst, S. M., Lucas, M., Rochet, M., Gaillardin, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R. R., Cruzado, L., Jimenez, J., Sánchez, M., del Rey, F., Benito, J., Domínguez, A., Revuelta, J. L., Moreno, S., Armstrong, J., Forsburg, S. L., Cerutti, L., Lowe, T., McCombie, W. R., Paulsen, I., Potashkin, J., Shpakovski, G. V., Ussery, D., Barrell, B. G., and Nurse, P. 2002. The genome sequence of *Schizosaccharomyces pombe. Nature* 415: 871-880.

Worden, A. Z., Lee, J. H., Mock, T., Rouze, P., Simmons, M. P., Aerts, A. L., Allen, A. E., Cuvelier, M. L., Derelle, E., Everett, M. V., Foulon, E., Grimwood, J., Gundlach, H., Henrissat, B., Napoli, C., McDonald, S. M., Parker, M. S., Rombauts, A., Von Dassow, P., Badger, J. H., Coutinho, P. M., Demir, E., Dubchak, I., Gentemann, C., Eikrem, W., Gready, J.E., John, U., Lanier, W., Lindquist, E. A., Lucas, S., Mayer, K. F., Moreau, H., Not, F., Otillar, R., Panaud, O., Pangilanan, J., Paulsen, I., Piegu, B., Poliakov, A., Robbens, S., Schmutz, J., Toulza, E., Wyss, T., Zelensky, A., Zhou, K., Armbrust, E. V., Bhattacharya, D., Goodenough, U. W., Van de Peer, Y., and Grigoriev, I. V. 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes Micromonas. *Science* 324: 268-272.

Zdobnov, E. M. and Apweiler, R. 2001. InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17: 847-848.

Zhao, C., Waalwijk, C., de Wit, P. J. G. M., Tang, D., and van der lee, T. 2013 RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. *BMC Genomics* 14: 1-16.

CHAPTER 3

Transcriptome and small RNAome dynamics during a resistant and susceptible interaction between cucumber and downy mildew

This Chapter is currently under review:

Burkhardt and Day. Transcriptome and small RNAome dynamics during a resistant and susceptible interaction between cucumber and downy mildew. Plant Genome. *Submitted.*

Abstract

Cucumber downy mildew, which is caused by the obligate oomycete pathogen *Pseudoperonospora cubensis*, is the primary factor limiting cucumber production. Although sources of resistance have been identified, such as plant introduction line PI 197088, the genes and processes involved in mediating resistance are still unknown. In the current study, we conducted a comprehensive transcriptome and small RNAome analysis of a resistant (PI 197088) and susceptible (Vlaspik) cucumber during a time course of *P. cubensis* infection using Illumina sequencing. We identified significantly differentially expressed genes within and between the resistant and susceptible cucumber leaves over a time course of infection. Weighted gene correlation network analyses created coexpression modules containing genes with unique expression patterns between Vlaspik and PI 197088. Recurring data trends indicated that resistance to cucumber downy mildew is associated with an earlier response to the pathogen, hormone signaling, and the regulation of nutrients. In total, candidate resistance genes were identified from multiple transcriptome analyses and literature support. In addition, parallel sequencing of small RNAs from cucumber and *P. cubensis* during the infection time course was used to identify and quantify novel and existing miRNA in both species. Predicted miRNA targets of cucumber transcripts suggest a complex interconnectedness of gene expression regulation in this plant-pathogen system. In total, this work bioinformatically uncovered gene expression patterns that are involved in the mediation of and or response to *P. cubensis* resistance. Herein, we provide the foundation for future work to validate candidate resistance genes and miRNA-based regulation proposed in this study.

Introduction

Pseudoperonospora cubensis, the casual agent of cucurbit downy mildew, is an obligate oomycete pathogen capable of infecting more than 20 genera within the Cucurbitaceae, including *Cucumis sativus* (cucumber), *Cucumis melo* (melon), and *Citrullus lanatus* (watermelon) (Savory et al. 2010). Within the first 24 hours of the host-pathogen association, pathogen infection occurs, which includes sporangia germination, zoospore encystment, and the initiation of hyphae formation. In parallel, the onset of host symptoms include the development of angular chlorotic lesions on the upper leaf surface that expand and coalesce, with pathogen sporulation occurring on the lower surface. In the following 2-6 days, pathogen hyphal development continues, leading to the formation of specialized structures called haustoria, through which metabolites, protein, and nucleic acids are transferred between the pathogen and host, many of which likely function in enhancing pathogen virulence. In the final stage of its life cycle, *P. cubensis* forms sporangiophores, which release sporangia into the air, leading to the establishment of new infection cycles.

The molecular and genetic bases of host resistance and pathogen virulence in downy mildew species has been best characterized through the study of *Arabidopsis thaliana* and *Hyaloperonospora arabidopsidis*, an obligate pathosystem through which the genetic diversity of host and pathogen has revealed a complex interaction of several host resistance (*R*) genes and race-specific effectors from in *H. arabidopsidis* (Nemri et al. 2010; Krasileva et al. 2011; Coates and Beynon 2010; Mohr et al. 2010). In addition, other studies have characterized the contribution of broad spectrum resistance responses, collectively

demonstrating the genetic interactions between quantitative and qualitative trait loci within the host (Lapin et al. 2012, Gao et al. 2014). In some cases, specific transcription factors have been shown to contribute to resistance, which is the case for transcription factors *Vv*WRKY33 and *Vv*WKRY1 which mediate resistance to grapevine downy mildew (*Plasmopara viticola*) (Marchive et al. 2013; Merz et al. 2014). Transcription-based resistance has also been identified in cucurbits, the best example of which is through resistance to *P. cubensis* in melon, which is conferred through the constitutive overexpression of glyoxylate aminotransferases (*At1* and *At2*) (Taler 2004). In addition to host genes that contribute to resistance signaling, numerous other genes have been associated with susceptibility, including those in Arabidopsis in response to *H. arabidopsidis* infection, such as 2-oxoglutarate (2OG)-Fe(II) oxygenase DMR6 (van Damme et al. 2008), the malectin-like receptor kinase IOS1 (Hok et al. 2011; 2014), and the negative regulator of plant defense response GSL5/PMR4 (Wawrzynska et al. 2010). In total, these genes were initially identified based on transcriptional differences and/or were found to control transcriptional activities in the plant during infection. Collectively, these studies have demonstrated the importance of transcriptional regulation of defense and susceptibility during host-oomycete interactions.

In addition to the identification of host genes associated with resistance signaling during pathogen infection, a growing body of literature has demonstrated that small RNAs (sRNAs) play essential roles in plant-pathogen interactions through the regulation of transcriptional, or post-transcriptional, processes. For example, host-derived sRNAs have been demonstrated to promote defense signaling in response to pathogen infection, or

reciprocally, by the pathogen to promote infection and the down-regulation of host immune responses (Qiao et al. 2013; Weiberg et al. 2013; 2014). Among the best characterized of these sRNAs are microRNAs (miRNAs), a class of single stranded, 20-22 nucleotide RNA molecules that post-transcriptionally regulate gene expression. As a group, miRNAs are hypothesized to play a key role as regulators of plant defense signaling including regulating processes associated with PAMP-triggered and effector-triggered defense responses (Weiberg et al. 2014; Zhao et al. 2015). Specifically, recent work has shown that sRNAs from *Botrytis cinerea* can move from the pathogen to the host, where they function to silence resistance-associated genes (Weiberg et al. 2013). In cucumber, miRNAs have also been identified (Mao et al. 2012; Martínez et al. 2011; Hu et al. 2014; Li et al. 2013; Jin and Wu 2015). However, none of these studies investigated the expression of sRNAs during a resistant response to biotic stress.

Historically, cucumbers have been bred for resistance to *P. cubensis*, yet in recent years, downy mildew epidemics in the United States suggest that the widely incorporated cucumber recessive resistance locus, *dm-1*, is no longer sufficient in providing durable resistance (Call et al. 2012a; 2012b). To define alternate sources of resistance, screening of numerous cucumber cultigens has resulted in the identification of a plant introduction line, PI 197088, which displays robust resistance to a field isolate of *P. cubensis* (Call et al. 2012b). More recently, work by Caldwell et al. (2014), used quantitative trait loci (QTL) mapping to identify three loci located on Chromosome 2, 4, and 5 as having a contribution to resistance to *P. cubensis,* with the QTL on Chromosome 5 being the most robust (~24%) (Caldwell et al. 2011). However, a separate study by Yoshioka et al. (2014) identified a

different constellation of QTLs for *P. cubensis* resistance from PI 197088-derived cucumber lines. Taken together, the sum of these two these studies support the hypothesis that downy mildew resistance in cucumber is likely due to the combinatorial effects of several genes.

In the current study, we advance previous genome- and transcriptomics-centered research in the *P. cubensis*-cucumber pathosystem (Tian et al. 2011; Burkhardt et al. 2015; Adhikari et al. 2012; Savory et al. 2012b; 2012a) through employing a comprehensive parallel transcriptome and small RNAome sequencing analysis to identify changes in gene and small RNA expression that are associated with a resistant or susceptible interaction. By analyzing temporal changes in gene expression over a time course of infection in susceptible (Vlaspik) and resistant (PI 197088) cucumber varieties, we identified more than 10,000 host genes that were differentially expressed. Using this approach, we compared the expression of these transcripts, including the pathogen-induced responses during resistance and susceptibility, to provide compelling evidence in support of the hypothesis that multiple genes – specifically at the earliest stages of the host-pathogen interaction – are responsible for the initiation of resistance to *P. cubensis*. Finally, we characterize the co-expression of Dicer-derived sRNAs from both host and pathogen and identify novel miRNAs in both *C. sativus* and *P. cubensis* that are predicted to regulate host gene expression during pathogen infection. Taken together, the work presented herein provides the first comprehensive parallel analysis of the transcriptome and small RNAome changes during the *P. cubensis* – *C. sativus* interaction, shedding light on processes that control resistance and susceptibility during obligate parasitism.

Materials and methods

Plant growth and *Pseudoperonospora cubensis* inoculation

Cucumber cv. Vlaspik (Osborne seeds, Mt. Vernon, WA sourced from Seminis hybrid Vlaspik M F1 88% Vlaspik, 12% Sire) and plant introduction line PI 197088 (seeds provided by Todd Wehner, NC State) were grown at 22°C in 16h light/8h dark cycles in Redi-earth potting mix (Sun Gro Horticulture, Agawam, MA) in a BioChambers Bigfoot Series Model AC-60 growth chamber (Winnipeg, Manitoba, Canada). *P. cubensis* isolate MSU-1 was propagated on the susceptible cucumber 'Vlaspik', as previously described (Tian et al. 2011). Infected plants were maintained at near 100% humidity at 22°C under a 12 light/12h dark cycle. Time course of infection experiments were performed by pipetting multiple 10 μL droplets of a 1 x 10^5 sporangia/ml solution of *P. cubensis* onto the abaxial side of fully expanded 4-week-old leaves. For mock-inoculated samples, water was pipetted onto the leaf surface. Inoculated plants were stored in the dark at nearly 100% humidity before being moved to the growth chamber (Adhikari et al. 2012). Inoculated leaf samples were collected using a #3 cork borer.

Disease phenotype analysis

For microscopic analyses of infected samples, samples were cleared in 100% ethanol and stained with trypan blue in a 1:1:1 solution of glycerol, lactic acid, and water. Visualization of destained tissues was performed using an Olympus IX71 inverted light microscope as

previously described (Savory et al. 2012a).

RNA isolation

RNA for real-time PCR was extracted from sporangia and from flash-frozen leaf tissue using the RNeasy Plant Mini Kit (Qiagen, Germantown, MD). Contaminating genomic DNA was removed from the RNA sample by performing the on-column DNase treatment (Qiagen). Total RNA was extracted from sporangia and flash-frozen leaf tissue for mRNA and sRNA sequencing using the AllPrep DNA/RNA/miRNA Universal kit (Qiagen) and contaminating DNA was removed using on-column DNase treatment. RNA quality was assessed using the 2100 Bioanalyzer (Agilent) with the Agilent RNA 6000 Pico kit.

Pathogen quantification qPCR

cDNA was prepared from RNA using the USB first-strand cDNA synthesis kit with random hexamer primers (USB Affymetrix, Santa Clara, CA). Samples were prepared using the HotStart SYBR Green qPCR Master Mix (2x; USB). A Mastercycler ep Realplex real-time PCR (Eppendorf AG, Hamburg, Germany) was used to perform quantitative real-time PCR (qPCR). *P. cubensis* internal transcribed space (ITS) primers were used to quantify the level of pathogen in the infected leaf material, as previously described (Tian et al. 2011). Values were normalized based on the relative expression of the pathogen ITS region as compared to cucumber-specific actin primers. For qPCR analysis, the following cycling parameters were used: 1 hold of 95°C for 2 min, 40-50 cycles of 95°C (15 sec) 56°C (15 sec), and 72°C

(30 sec). Relative expression was calculated where relative expression = $2^{(-\Delta Ct)}$ and where $\Delta Ct = Ct_{ITS} - Ct_{actin}$ (Porter et al. 2012). Error bars represent the standard error from four biological replicates. Data were analyzed and processed using Prism (GraphPad Software, Inc., San Diego, CA). A two-way ANOVA was used to compare time and plant, and the Sidak's multiple comparisons test was used to correct for multiple comparisons. Alpha ≤ 0.05 was used to determine significance.

Library preparation and next generation sequencing

Libraries from two biological replicates, each, of Vlaspik and PI 197088 cucumber leaves that were mock-inoculated or collected 1, 2, 3, 4, and 6 days post inoculation (dpi) with *P. cubensis* were prepared using the Illumina TruSeq Stranded mRNA Library Preparation Kit LT for mRNA or the Illumina TruSeq Library Preparation Kit for sRNA by the Michigan State Research Technology Support Facility (RTSF). For sRNA samples, products were gel purified using a 6% polyacrylamide gel. Samples were pooled and run on the Illumina HiSeq 2500 Rapid Run flow cell (v1) using Rapid SBS reagents. Each mRNA sample was sequenced with at least 20 million 50 nucleotide (nt) single end (SE) reads and each sRNA sample was sequenced with at least 10 million 50 nt SE reads.

Processing and alignment of mRNA-Seq reads

Illumina Real Time Analysis (RTA) software v1.17.21.3 was used to call bases. The output of RTA was de-multiplexed and converted to FastQ files with Illumina Bcl2Fasta v1.8.4.

Reads were deposited in the Sequence Read Archive (SRA) as a BioProject at the National Center for Biotechnology Information under the accession number PRJNA285071. The reads were evaluated using FastQC (Babraham Bioinformatics), and the sequencing adapters were removed using Cutadapt v. 1.4.1. The first 14 bases were trimmed using fastx_trimmer from the FASTX v. 0.0.13 program. The processed reads were uniquely mapped to the annotated *C. sativus* reference genome of Chinese long version 2 (ftp://www.icugi.org/pub/genome/cucumber/Chinese_long/v2/) using Bowtie v.1.0.0 (Langmead et al. 2009) and TopHat v1.4 (Trapnell et al. 2009). The minimum and maximum intron sizes allowed were 50 bp and 50000 bp, respectively (Adhikari et al. 2012). Accepted BAM hits from TopHat were converted to SAM files using SAMTools v0.1.18 (Li et al. 2009). The number of reads that aligned to each annotated gene of the cucumber genome were counted using HTSeq count v. 0.60 using the –s reverse and –t gene options, which is part of the HTSeq Python package (Anders et al. 2014). The percent of uniquely mapped reads was calculated by dividing the total number of reads that mapped to genes using HTSeq count by the total number of reads that were obtained from sequencing. Normalized read counts were calculated by dividing the raw counts from HTSeq count by the library size calculated for each sample by DESeq2 (Love et al. 2014). The data were then log2 transformed and negative values were converted to zero. The correlation of biological replicates was determined and the values for the replicates were averaged for each time point for downstream weighted gene correlation network analysis (WGCNA) analyses. The spliced version of each gene was used to report the annotation and gene ontology (GO) terms of each gene.

Analysis of differential gene expression

The raw HTSeq count data were used as the input to measure differential gene expression using DESeq v. 1.18.0 implemented in R (http://cran.r-project.org/) (Seyednasrollah et al. 2013; Anders et al. 2013). DESeq analyses were performed using the standard procedures for pairwise comparisons with the p-adjusted value < 0.05 as the cutoff for determining significance. DESeq was used to find significantly differentially expressed genes between Vlaspik and PI 197088 at each time point and between all time points, including the mock. DESeq2 v. 1.6.3 was used to identify genes that were differentially expressed across the time course (Love et al. 2014). Count data from HTSeq was used as input, and the data set design included the variables of plant line, time and the interaction between the plant line and time. The likelihood ratio test with a reduced design allowed for the contributions of each variable to be measured by comparing the full model to the reduced model as a method to explain the variation in the calculated values of gene expression; the p values reported the likelihood that the source of variation within the model could be attributed the removed variable. In this way, the significance of the contributions of the plant line, time, and the interaction of the plant line and the time to the pattern and changes in gene expression observed in all plant lines at all time points was determined.

Weighted gene correlation network analysis

Genes with correlated expression patterns were identified using WGCNA implemented in R (http://cran.r-project.org/) (Langfelder and Horvath 2008). Genes included in this analysis

were such that a gene was required to have a DESeq2 p-adjusted value of less than 1×10^{-5} to be included in the analysis. Such a stringent cutoff was used to reduce the number of genes classified as significant. The input expression values for each gene were the processed HTSeq counts as described above. Modules were generated from these 12,532 genes using data for Vlaspik, PI 197088, or both. When genes did not exhibit sufficient variance or had zero value, they were excluded from the modules. All WGCNA parameters were used in their default settings with the following modifications: the soft threshold (β) parameter of 16 was used to create the Vlaspik modules and a β of 18 was used to create the PI 197088 and Vlaspik/PI 197088 combined modules. For all modules, the treecut parameter was 0.9, and the resulting modules were merged using a distance threshold of 0.2. Modules were visualized by plotting the Z-scores for all genes within a given module (Childs et al. 2011).

Gene ontology enrichment analysis

Gene Ontology (GO) terms for each predicted cucumber gene were obtained by running the predicted cucumber protein sequences through InterProScan (Zdobnov and Apweiler 2001). A custom script was written to parse the file to create a list of all of the unique GO terms for each protein translated from the spliced version of each gene. The GO terms for each gene in an enrichment list (i.e., a list of genes within a specific module) were selected for GO term enrichment analysis, which was performed using the singular enrichment analysis (SEA) online tool on the AgriGo website (http://bioinfo.cau.edu.cn/agriGO/index.php). This analysis was run using the standard

parameters, which were the Fisher statistical method with the Yekutieli multi-test adjustment method at a significance level of 0.05.

Processing and alignment of small RNA reads

Illumina Real Time Analysis (RTA) software (v1.17.21.3) was used for base calling. The output of RTA was de-multiplexed and converted to FastQ files with Illumina Bcl2Fasta v1.8.4. Reads were evaluated using FastQC, and sequencing adapters were removed using Cutadapt (v1.4.1). Sickle (v. 1.33) was used to trim reads with a quality score below 30 and reads longer than 15 nucleotides were used for further analysis. The high quality, trimmed reads were aligned to both the *P. cubensis* genome and the *C. sativus* genome using Bowtie v. 1.0.0 using the options –v 1, --best, -strata, -k 50 as was used to identify sRNA in *Arabidopsis thaliana* (Axtell 2013b). Reads mapping to only one of the genomes were used for sRNA analyses in either *P. cubensis* or *C. sativus.* Small RNAs were identified and mature miRNAs were quantified using the default settings of the plant-optimized sRNA bioinformatics program ShortStack v. 2.1.0, which utilized other sRNA programs including RNAfold and RNAplot to identify hairpins and miRNA (Axtell 2013b; Shahid and Axtell 2013). The number of reads mapping to each identified sRNA cluster were counted with HTSeq using the alignment files generated by ShortStack. The miRNA read counts were divided by the DESeq2 library size, and the resulting data was log2 transformed and negative values were converted to zero. Biological replicates were combined.

miRNA family identification and target prediction

The families of each predicted miRNA locus were identified from the 2-3 most abundant mature miRNA sequences predicted by ShortStack. These mature miRNA were then compared to the miRNAs present in miRBase through a BLASTN search using an E-value cutoff of 1.0 to be considered as part of the known miRNA family. The miRBase database was also used to identify conserved miRNA in all species present in the database. Cucumber targets of the mature miRNA sequences were identified using psRNATarget with a maximum expectation value of 3.0 and with default program setting (Dai and Zhao, 2011). TAPIR, the target prediction for plant microRNAs (Bonnet et al. 2010), was also used to predict targets from among the *C. sativus* v. 2 transcripts using the score cutoff of 4 and the free energy ratio cutoff of 0.7. A gene was considered to be a target of a predicted sRNA if it was identified using both prediction methods.

DESeq2 for miRNA

Differentially expressed miRNA were identified using DESeq2 (Love et al. 2014) following the procedures described for mRNA. The cutoff for determining significance was a p-adjusted value of less than 0.05. The expression patterns of significantly expressed miRNA were correlated with the expression patterns of their predicted targets by calculating and plotting the z-scores for the normalized count values.

qPCR validation of RNA-Seq data

RNA was extracted with the Qiagen All Prep kit and was synthesized into cDNA as previously described. Samples were prepared for qPCR as previously described but using the HotStart SYBR Green qPCR Master Mix with ROX. A 7500 Fast real-time PCR machine (Applied Biosystems) was used to perform quantitative real-time PCR using the following parameters: 1 hold of 95°C for 10 min and 40 cycles of 95°C (15 sec) and 72°C (1:00 sec). A melt curve analysis was run to ensure that only amplicon was generated per reaction. Primers used to amplify the genes of interest are in Table 6. Relative expression was calculated whereby relative expression = $2^{(-\Delta Ct)}$ and where $\Delta Ct = Ct_{\text{gene of interest}} - Ct_{\text{actin}}$ (Porter et al. 2012). Error bars represent the standard error from three biological replicates. A Grubbs' test was performed to remove outliers (alpha < 0.05).

Table 6. *List of primer sequences used for qPCR to validate expression patterns of mRNA*

Primer Name	Sequence (5´->3´)
Actin F	TGCTGGATTCTGGTGATGGTGTGA
Actin R	AGGTCCAAACGGAGAATGGCATGA
Csa1G051760 F	GGAAGAAGCTCAATGATAGGCTG
Csa1G051760 R	GGAGGTCGTTGATTCTTTGCAG
Csa4G641000 F	TGCCGCCCTTGACATTACAG
Csa4G641000 R	AGTTCAGTTACCTGAATGGATAGCA
Csa3G017320 F	CGTCGGTTGGGTACCATCGC
Csa3G017320 R	GCCGCAGCCACCTCCGT
Csa3G135080 F	GGGCTTTCATTGCTGCTGTG
Csa3G135080 R	CTGTGGGACAGTAGAGGCG
Csa7G432520 F	GTGACACCGCTGGGTGGG
Csa7G432520 R	CTCTGACTACATCATCGTCGGG
Csa2G028490 F	GGAGTTAACCCTATTCTCATTTCTCGTC
Csa2G028490 R	CACTGAGTCCTTCCAAGCCATCC
Csa5G638300 F	CTGCTGGAATGTTGGATACTGTG
Csa5G638300 R	AGCTCTTGCCCTCAGATCTC
Csa1G008580 F	CAACCACCACCACCACCTGC
Csa1G008580 R	CTCCATAAGCACGAATCAGTGG
Csa5G604920 F	GACTACGATCTTCTCAAGAG
Csa5G604920 R	TGATTCTGTTGGCTCTCTAC
Csa3G078800 F	CGTATTGCCATCATCGAAC
Csa3G078800 R	CATAGTCGTCTTGCTCGTC
Csa1G528590 F	CACTGCTCAATTTGCAGAG
Csa1G528590 R	TCATTACACCACTCAAGCTG
Csa6G433300 F	GAGACAGCAACTGGTGTGAC
Csa6G433300 R	GCCACATCTACTCCCCAATC

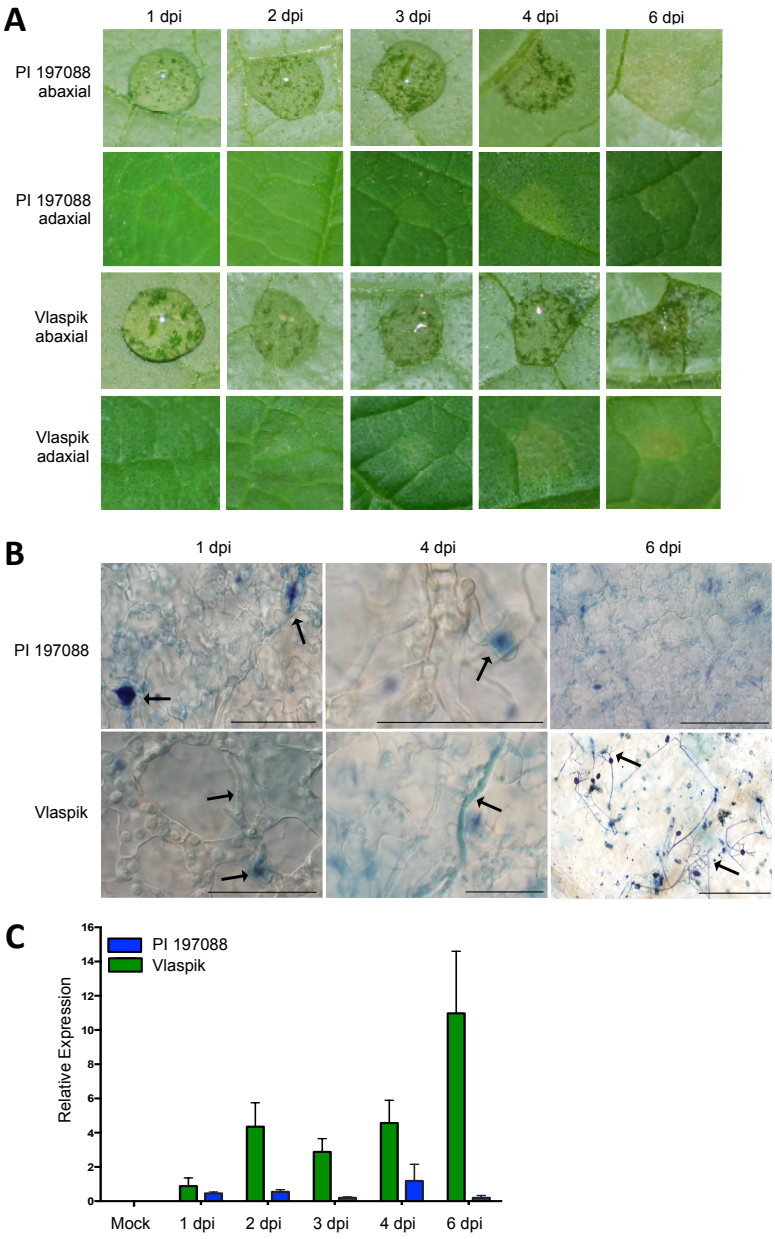
Results and discussion

P. cubensis growth and development is abrogated in the resistant cucumber accession PI 197088

A time course of *P. cubensis* infection of resistant (PI 197088) and susceptible (Vlaspik) cucumber varieties was conducted to investigate changes in phenotypes and gene expression between the two plant lines during the infection cycle. As shown in Figure 13, pathogen infection of PI 197088 and Vlaspik showed increasingly divergent phenotypes, as demonstrated through microscopy and leaf lesions, as well as supported by quantitative real-time PCR to measure pathogen proliferation. In the susceptible cultivar Vlaspik, the area of the leaf inoculated with *P. cubensis* showed a water-soaking phenotype and developing chlorotic lesions (Fig. 13A). In contrast, the resistant variety PI 197088 showed smaller, less chlorotic lesions. Microscopic examination of infected lesions supported the observed whole leaf phenotypes, revealing that sporangiophore development occurred as early as 6 days post-inoculation (dpi) in Vlaspik but was not observed in infected PI 197088 (Fig. 13B). The lack of sporulation on PI 197088 does not appear to be the result of a block in pathogen entry, as indicated by encystment and hyphal growth (Fig. 13B; arrows at 1 and 4 dpi). However, it is unclear from microscopy observations if the pathogen is living. Furthermore, we did not observe pathogen elicitation of the hypersensitive response, which is consistent with breeding data (Caldwell et al. 2014; Yoshioka et al. 2014).

To quantitatively assess pathogen growth as a means to correlate pathogen load with precise stages of host infection, qPCR was used (Fig. 13C). We observed a steady increase in the abundance of *P. cubensis* in Vlaspik over time, with an approximate 12-fold increase at 6 dpi, as compared to samples analyzed from 1 dpi. In contrast, the level of *P. cubensis* in PI 197088 remained relatively low, and the level of pathogen growth never exceeded 3-fold of that observed at 1 dpi (Fig. 13C). The greatest contrast in pathogen growth on either Vlaspik or PI 197088 was observed at 6 dpi. A two-way ANOVA indicated that the differences attributed to time, plant line, and the interaction were all statistically significant. As noted above, a key feature of resistance observed in PI 197088 is that the pathogen was not able to sporulate; however, phenotypic and qPCR data indicate the early stages of infection are occurring in the resistant line. We posit that once the initial entry of the pathogen occurs, the pathogen is capable of growing, if only for a brief period of time, in the resistant plant line. Given that *P. cubensis* is an obligate biotroph and requires host resources for all of its nutrition, we hypothesize that one possible contributing mechanism that functions in the abrogation of pathogen growth is host restriction of *P. cubensis* nutrient acquisition. Alternatively, or in addition, a second hypothesis is that changes in plant metabolism, such as changes in hormones or nutrients, may play a role in resistance.

Figure 13. *Phenotypic and quantitative analysis of* Pseudoperonospora cube*nsis infection of Vlaspik (susceptible) and PI 197088 (resistant) cucumber.* (A) Leaf phenotypes of the lower (abaxial) and upper (adaxial) leaf surface inoculated with 10µL of 1 x 10^{-5} *P. cubensis* sporangia per ml during a time course. (B) Light microscopy of inoculated Vlaspik and PI 197088 leaf punches were cleared and stained with trypan blue. Scale bars = 50 µm for Vlaspik at 2-4 dpi, 100 µm for Vlaspik at 1 dpi and PI 197088 at 1-4 dpi, 200 µm for PI 197088 at 6 dpi, and 500 µm for Vlaspik at 6 dpi. Arrows indicate pathogen structures. (C) Pathogen quantification using quantitative real time PCR analysis. DNA primers specific for *P. cubensis* ITS normalized to *Cucumis sativus* actin. Statistical analysis was performed using a two-way ANOVA using the Sidak's multiple comparisons test. The differences attributed to time, plant line, and the interaction were all statistically significant ($p < 0.05$). Error bars represent the standard error of four biological replicates. dpi = days post inoculation.



Transcriptome profiling of the *P. cubensis*-cucumber interaction

RNA-Seq was used to undertake comprehensive transcriptome profiling for five infection time points – 1, 2, 3, 4, and 6 days post inoculation (dpi) – and mock-inoculated leaf samples from both Vlaspik and PI 197088. Two independent and parallel time course inoculations were conducted on both Vlaspik and PI 197088 to yield two biological replicates. On average, following trimming and removing lower quality reads, 74% of the sequenced reads from each sample uniquely mapped to the cucumber genome (Table 7). The correlation of biological replicates for the normalized read counts of all the genes between each sample was calculated to have an R^2 value of at least 0.95 (Fig. 14). The normalized read counts for each gene at each time point can be found in the supplemental files in the author-recommended online resources. In total, 19,581 genes were expressed during at least one time point in at least one cultigen.

Table 7. *The percent of trimmed, high quality reads uniquely mapping to the Chinese long* Cucumis sativus *genome v. 2 using Bowtie 1.0.0 for each biological replicate of Vlaspik and PI 197088.* *Low percentage mapping is due to rRNA contamination. This was not found to affect the downstream analyses and over 17 million reads still uniquely mapped.

Sample	Percent Uniquely Mapped Reads
Vlaspik 1dpi BioRep1	84%
Vlaspik 2dpi BioRep1	72%
Vlaspik 3dpi BioRep1	75%
Vlaspik 4dpi BioRep1	73%
Vlaspik 6dpi BioRep1	72%
Vlaspik Mock BioRep1	75%
Vlaspik 1dpi BioRep2	73%
Vlaspik 2dpi BioRep2	74%
Vlaspik 3dpi BioRep2	75%
Vlaspik 4dpi BioRep2	81%
Vlaspik 6dpi BioRep2	82%
Vlaspik Mock BioRep2	83%
PI 197088 1dpi BioRep1	73%
PI 197088 2dpi BioRep1	31% *
PI 197088 3dpi BioRep1	65%
PI 197088 4dpi BioRep1	76%
PI 197088 6dpi BioRep1	77%
PI 197088 Mock BioRep1	75%
PI 197088 1dpi BioRep2	73%
PI 197088 2dpi BioRep2	74%
PI 197088 3dpi BioRep2	75%
PI 197088 4dpi BioRep2	76%
PI 197088 6dpi BioRep2	77%
PI 197088 Mock BioRep2	75%

Figure 14. *Correlation of expression values in two biological replicates of PI 197088 and Vlaspik inoculated with* Pseudoperonospora cubensis *or mock-inoculated with water*. Reads were mapped to the *Cucumis sativus* genome v. 2 using Bowtie 1.0.0 and TopHat v. 1.4 The number of reads mapping to each gene were counted using HTSeq v. 0.6 and were normalized divided by the library size for each sample calculated from DESeq2. The data was then log2 transformed and negative values were converted to zero. R^2 is the correlation coefficient. dpi = days post inoculation.

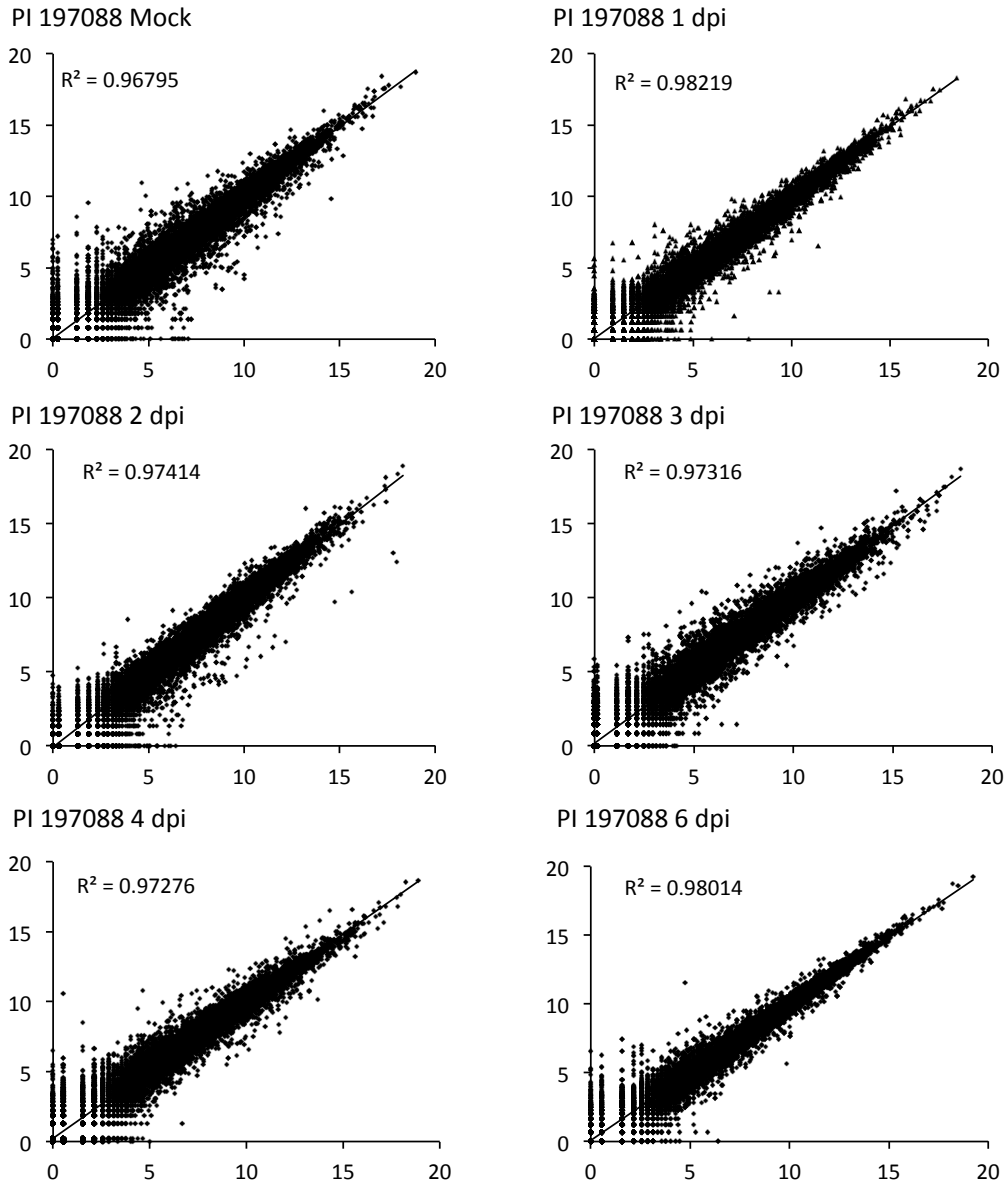
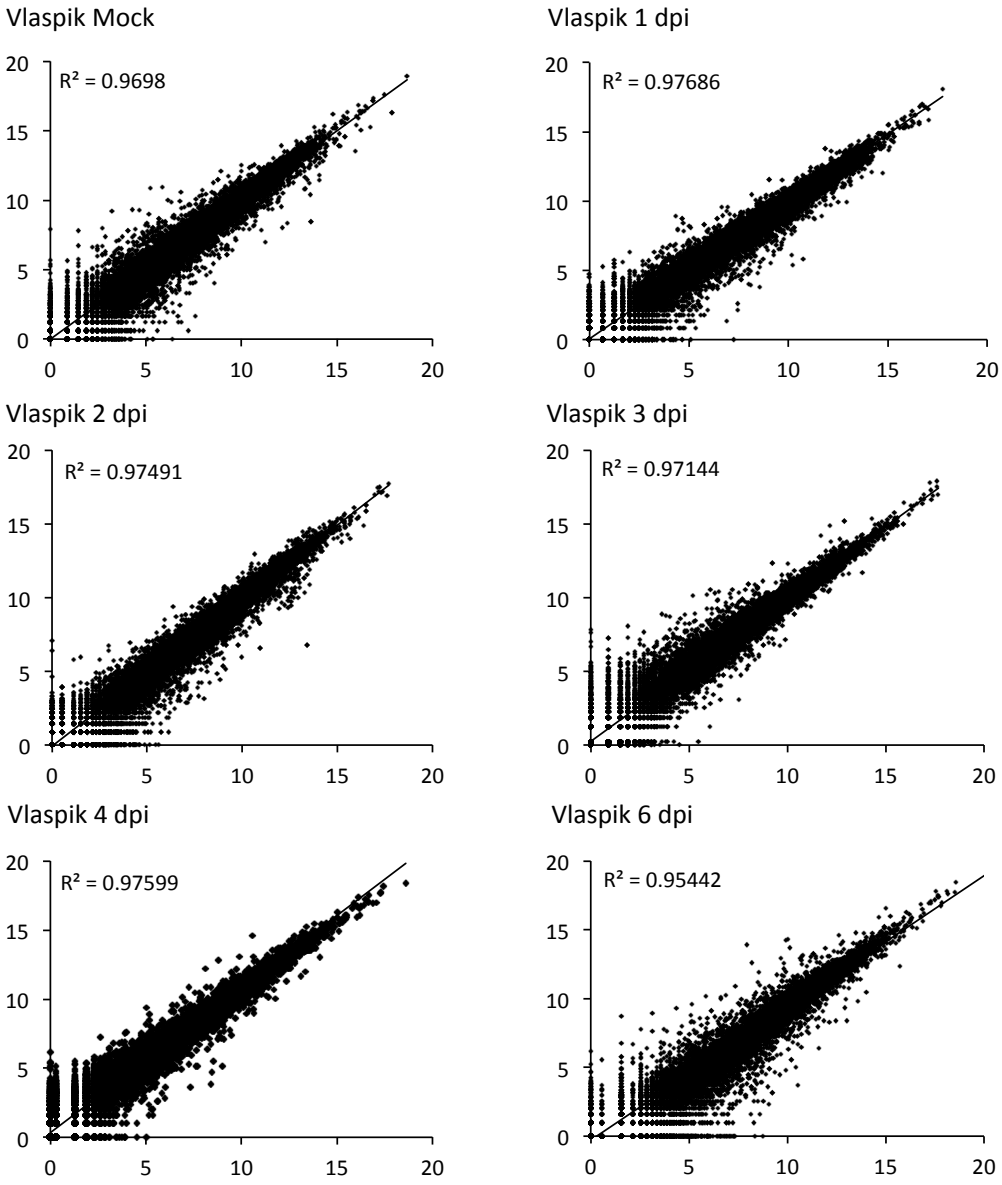


Figure 14 (cont'd)



Differentially expressed genes in PI 197088 and Vlaspik reveal early time points are important in *P. cubensis* infection response

To identify differentially expressed host genes during a resistant and susceptible interaction, we next used DESeq analyses to perform a pairwise comparison between Vlaspik and PI 197088. Using this approach, we found that the greatest number of differentially expressed genes existed between the mock-inoculated plant and each inoculation time point (Table 8). The number of DE genes compared to mock was lowest at 1 dpi for both plants, yet compared to Vlaspik, PI 197088 still had several thousand more DE genes compared to mock-inoculated samples harvested at 1 dpi. For 2-6 dpi, the number of DE genes was similar compared to mock with the greatest number of DE genes occurring at 4 dpi in Vlaspik and 2 dpi in PI 197088. Together, these data suggest that PI 197088 responds earlier to the pathogen than Vlaspik. Furthermore, the number of genes that were DE when comparing the later time points in either cultigen were very low, with the number of DE genes in the hundreds when comparing 3 dpi or 4 dpi to 6 dpi and below 100 when comparing the nearest time points. This could suggest that more gradual changes in gene expression exist once the pathogen has established itself within the host (in the case of Vlaspik) or has been resisted by the plant (in the case of PI 197088).

Table 8. *Differentially expressed genes as determined by DESeq from Vlaspik (susceptible) and PI 197088 (resistant) inoculated with* P. cubensis*.* The percentages in parentheses indicate the percentage of significantly (padj <0.05) down-regulated genes. dpi = days post inoculation.

Vlaspik	1 dpi	2 dpi	3 dpi	4 dpi	6 dpi
Mock	1969 (18%)	6526 (43%)	5610 (39%)	6652 (43%)	5733 (40%)
1 dpi		3699 (5%)	2782 (36%)	3347 (38%)	2689 (34%)
2 dpi			23 (30%)	242 (38%)	938 (50%)
3 dpi				28 (79%)	85 (29%)
4 dpi					7 (14%)

PI 197088	1 dpi	2 dpi	3 dpi	4 dpi	6 dpi
Mock	4864 (58%)	7639 (42%)	6671 (46%)	6901 (45%)	7059 (45%)
1 dpi		5990 (39%)	4987 (34%)	5625 (37%)	7237 (43%)
2 dpi			51 (55%)	15 (27%)	302 (74%)
3 dpi				2 (0%)	717 (46%)
4 dpi					56 (73%)

In addition, trends were observed in the number of significantly DE genes that were up- or down-regulated in either Vlaspik or PI 197088 at some time points. For example in Vlaspik, roughly 18% of the genes were down-regulated at 1 dpi compared to the mock, while in PI 197088 over 58% of the genes were down-regulated at 1 dpi compared to the mock (Table 8). Similarly, in Vlaspik only about 5% of the genes were down-regulated at 2 dpi compared to 1 dpi while in PI 197088, the same comparison resulted in about 39% of the genes being down-regulated (Table 8). Among the down-regulated genes of PI 197088 at 1 dpi, several significant gene ontology (GO) terms were associated with photosynthesis processes, transcription factor activity, or signaling cascades. The GO term of response to hormone stimuli was attributed to several auxin response factor genes; this is significant because auxin has been shown to promote pathogen growth and is often negatively regulated in initial responses to pathogens (Huot et al. 2014).

Metabolism and hormone signaling may play a role in resistance signaling

Pairwise comparisons between each time point and the mock samples were further analyzed by comparing lists of genes, which were commonly or oppositely up- or down-regulated, at each time point (Fig. 15). Among the expressed genes identified as uniquely up-regulated in PI 197088, common GO terms included photosynthesis and transcription factor activity at 2-6 dpi. In contrast, glycolysis related GO terms were common among the gene groups only up-regulated in Vlaspik during infection. This might indicate that while PI 197088 is able to continue with photosynthesis during infection, Vlaspik must break down more of its carbohydrates. Genes associated with exocytosis were also enriched in up-

regulated genes in only Vlaspik at 4 and 6 dpi. Among the genes that were uniquely down-regulated in PI 197088, were intracellular transport associated GO terms including those related to the Golgi and ER. In addition, zinc ion binding genes and genes associated with translation and signal transduction, including protein phosphatases, were also included in PI 197088 only down-regulated genes. Beyond the uniquely induced or down-regulated genes from each plant line, the genes that were oppositely regulated between Vlaspik and PI 197088 were of particular interest, and generally increased in number across the time course (Fig. 15). For example, two separate ethylene responsive transcription factors, Csa3G017320 and Csa2G191300, were down-regulated in PI 197088 between 2-6 dpi and up-regulated in Vlaspik; this pattern in differential regulation supports the role of hormone signaling in defense regulation (Fig. 16).

Figure 15. *Venn diagrams of significantly differentially expressed genes from DESeq at each time point compared to mock between Vlaspik and PI 197088*. (A) 1 dpi (B) 2 dpi (C) 3 dpi (D) 4 dpi (E) 6 dpi. dpi = days post inoculation.

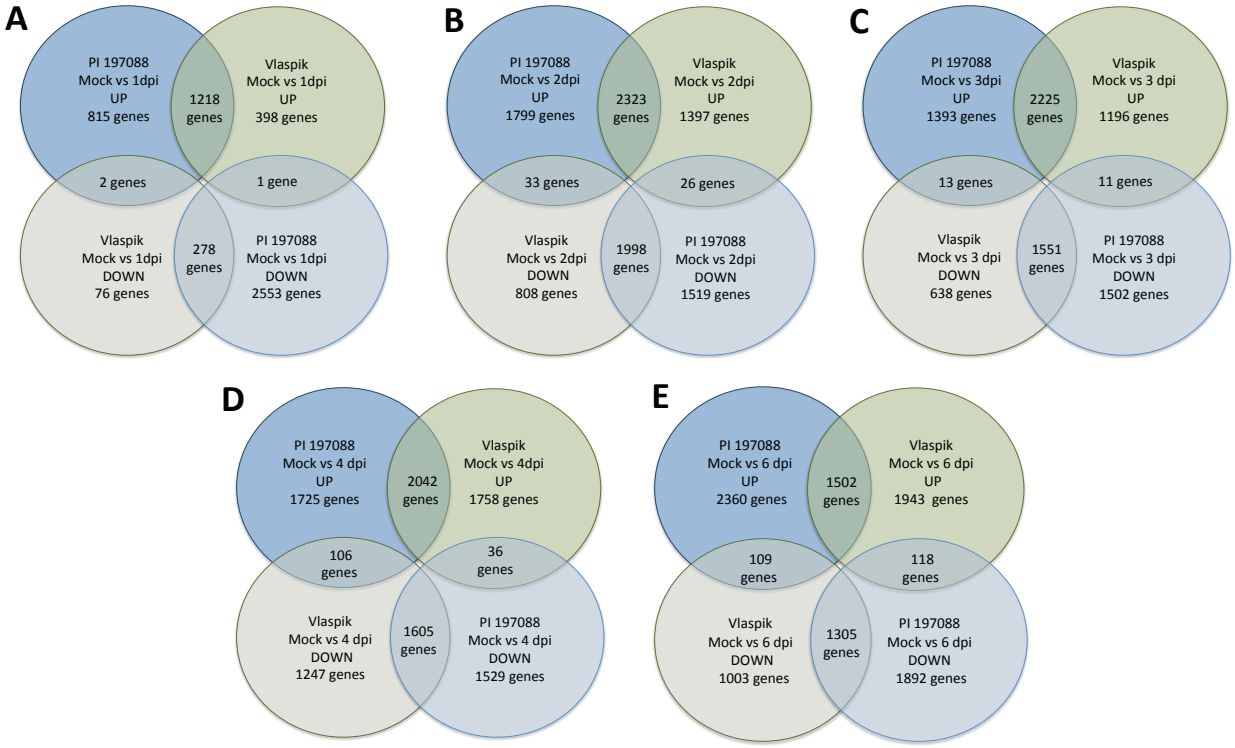


Figure 16. *Validation of RNA-Seq data from select genes with qPCR.* RNA-Seq plots are log2 transformed normalized RNA-Seq counts with error bars representing the standard error from 2 biological replicates. qPCR plots represent the relative expression of each gene of interest normalized to actin with the error bars representing the standard error of 3 biological replicates.

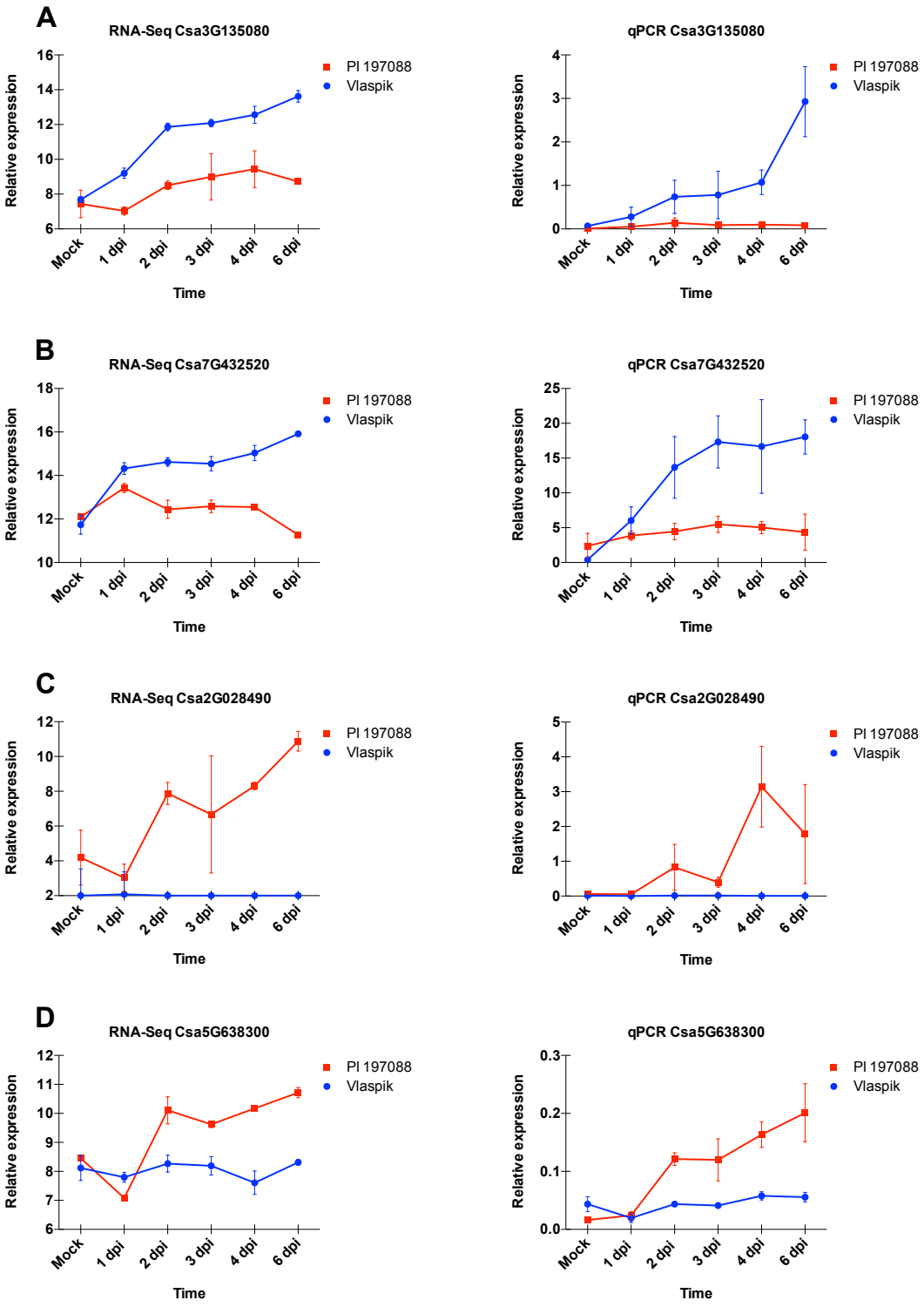


Figure 16 (cont'd)

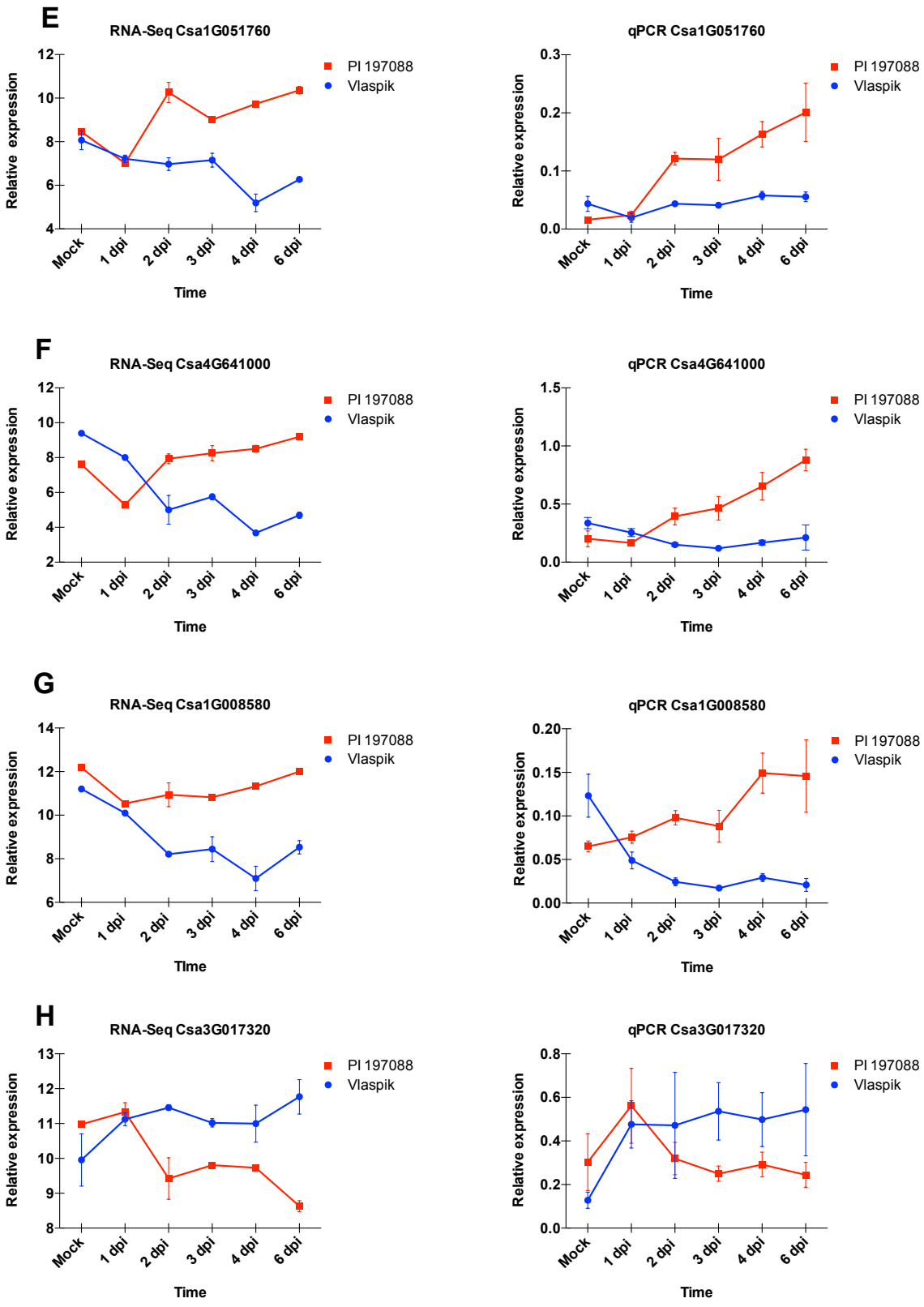
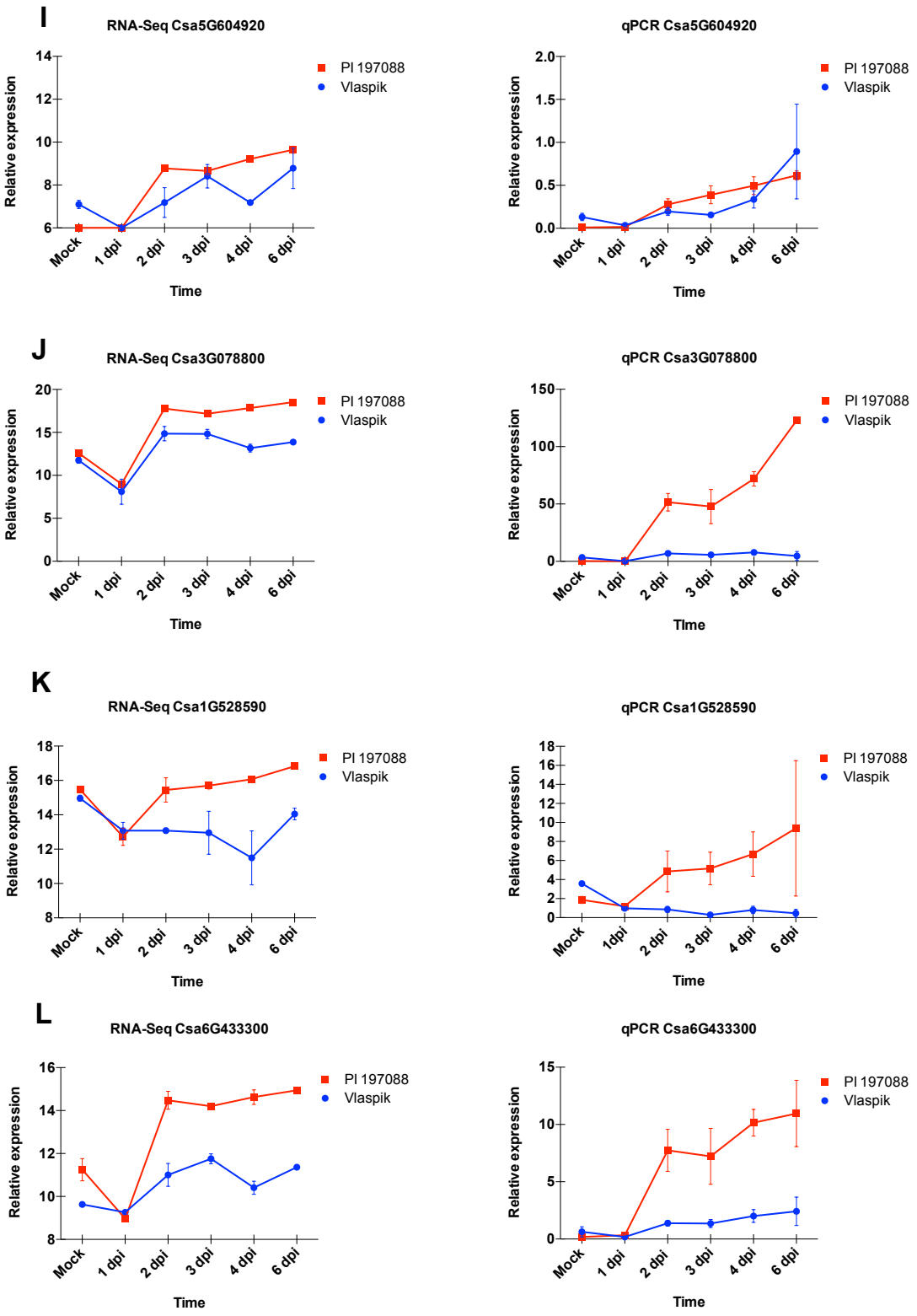


Figure 16 (cont'd)



Next, DESeq was used to compare genes expressed at each time point between plant lines. From this analysis, we identified 655 DE genes between the Vlaspik and PI 197088 mock-inoculated plants. Among the genes identified, the only enriched GO terms were those related to photosynthesis. As a result, additional analyses, below, included the mock in DESeq2 analysis without using it as a baseline for comparison. Still, in the pairwise comparisons, it was evident that an important change in gene transcription patterns was occurring between 1 and 2 dpi, as the number of DE genes between Vlaspik and PI 197088 at 1 dpi was relatively low (ca. 472), while the number of DE genes at 2 dpi was considerably high (ca. 3538). Interestingly, a GO term that was significantly enriched at 1 dpi among the down-regulated genes in PI 197088 was related to secondary active transmembrane transporter activity, and included genes Csa2G175170 and Csa3G135080, which encode phosphate transporters, and Csa3G094510 and Csa6G409380, which encode general solute carriers (Fig. 16). From this, we posit that if any of these transporters are functioning in the delivery of nutrients to *P. cubensis,* it supports the hypothesis that a host-driven restriction in nutritional compounds could be a significant contributor to host resistance (Stassen and Van den Ackerveken 2011). As noted above, in contrast to the low number of DE genes at 1 dpi, there were 3538 genes DE between Vlaspik and PI 197088 at 2 dpi, with 1624 being down-regulated in PI 197088 (66 enriched GO terms) and 1914 being up-regulated (45 enriched GO terms). Among the GO terms that were significantly enriched among the down-regulated genes at 2 dpi were GO terms associated with molecular processes such as protein phosphorylation and ubiquitination processes. Both of these processes have been extensively characterized as key processes associated with

immune and defense signaling pathways in plants (Antolín-Llovera et al. 2014; Coca and San Segundo 2010; Furniss and Spoel 2015; Belkhadir et al. 2014).

Compared to the early time points, the number of DE genes between Vlaspik and PI 197088 was highest at 4 and 6 dpi, where the number of DE genes was 5297 and 6867, respectively. This pattern of expression suggests a strong divergence in the biology of the two genotypes; perhaps because one had successfully resisted infection and the other was succumbing to the pathogen. A unique GO term in the down-regulated genes at 4 dpi included the molecular process of extracellular-glutamate gated ion channel activity, for which several genes contributing to this GO term belong to the plant defense-related clade 3 of this protein family (Forde and Roberts 2014). Common GO terms associated with cellular processes in down-regulated genes in PI 197088 compared to Vlaspik at 4 dpi and 6 dpi included those associated with the proteasome complex, the endoplasmic reticulum, and the exocyst. Some of these genes could be involved in the formation of the extrahaustorial membrane surrounding haustoria and therefore might be allowing less nutrient flow to the pathogen when they are down-regulated (Lu et al. 2012). Within the groups for genes up-regulated at 4 dpi and 6 dpi in PI 197088, the heat shock protein GO term, which has prior evidence for being involved in downy mildew resistance in cucumber (Li et al. 2012).

Significantly differentially expressed genes group into distinct modules of co-expression

In order to reduce the number of biologically relevant genes for WGCNA, DESeq2 was used to filter the number of genes that were significantly DE not only in pairwise comparisons, but also in multiple comparisons across both time and between plant lines. Using this method, we found that the gene with the most significance for the interaction between plant line and time of infection was Csa1G051760, which encodes an inducer of *CBF* expression and was highly induced only in PI 197088 (Fig. 16). As a function of plant defense signaling, previous work showed that a homolog of this gene was induced by jasmonic acid (Hu et al. 2013), and was also regulated by ABA, a key hormone associated with downy mildew defense signaling in Arabidopsis (Knight 2004; Hok et al. 2014; 2011). The second most significantly differentially regulated gene for the interaction factor of plant and time was Csa4G641000, an N-hydroxycinnamoyl/benzoyltransferase, a homolog of which catalyzes a committed step in phytoalexin synthesis (Yang et al. 1997), which was induced and highly expressed only in PI 197088 (Fig. 16). This is of particular significance, as phytoalexin is an important secondary metabolite in plant defense signaling (Hammerschmidt 1999; Zhao et al. 2005).

Following DESeq2 filtering, approximately 12,000 genes were further segregated based on patterns of expression. Using this approach, when either Vlaspik or PI 197088 expression values were analyzed, 12 separate modules were constructed, while 9 modules were constructed when Vlaspik and PI 197088 expression values were analyzed together. For example, as shown in Figures 17 and 18, modules were found to illustrate distinct patterns

of gene expression, whereby the plotted values (i.e., Z-scores) describe how many standard of deviations each data point is from the mean of the normalized expression values for each gene across all time points. Genes with the same or exact opposite expression patterns were grouped in the same module. For example, in PI 197088, Module K contained the fewest number of genes (i.e., 34), while Module G contained the highest number of genes (i.e., 4236 genes). The average module size was approximately 1000 genes (Fig. 17A, Fig. 19). For the grouping of genes from Vlaspik, Module H contained the fewest number of genes (i.e., 48), while Module E contained the greatest number of genes (i.e., 5156) with the average module size of approximately 1000 genes (Fig. 17B, Fig. 19). Overall, modules constructed for Vlaspik contained genes that were regulated in the same directional pattern, with few oppositely regulated genes, while in PI 197088, several of the modules included genes with opposite regulation patterns. PI 197088 also had more genes in modules that changed in expression pattern at an early time point (i.e., 1 dpi) compared to Vlaspik, whose modules showed delayed transcriptional changes. For the combined Vlaspik and PI 197088 data, the average module size was approximately 1200 genes, with the lowest number of genes (84) in Module B and the higher number of genes in Module D have (ca. 2853) (Fig. 18, Fig. 19).

Figure 17. *Trend plots of normalized gene expression values from PI 197088 and Vlaspik grouped through weighted gene correlation network analysis.* Z-score indicates the number of standard of deviations away from the mean. PI 197088 Module E genes display a dichotomous pattern in the mock and relatively high or low expression throughout the infection time course. PI 197088 Module C genes have either high expression at 1 dpi and low expression approaching 6 dpi or low expression at 1 dpi and high expression approaching 6 dpi. PI 197088 Module G genes have a strong high or low expression at 1 dpi. Vlaspik Module G genes have increased expression at 1 and 6 dpi. Vlaspik Module D genes have increased expression at 2 dpi and reduced expression at 4 dpi. Vlaspik Module K genes have variable expression, with lower expression at 1 and 4 dpi and higher expression at 3 and 6 dpi. dpi = days post inoculation.

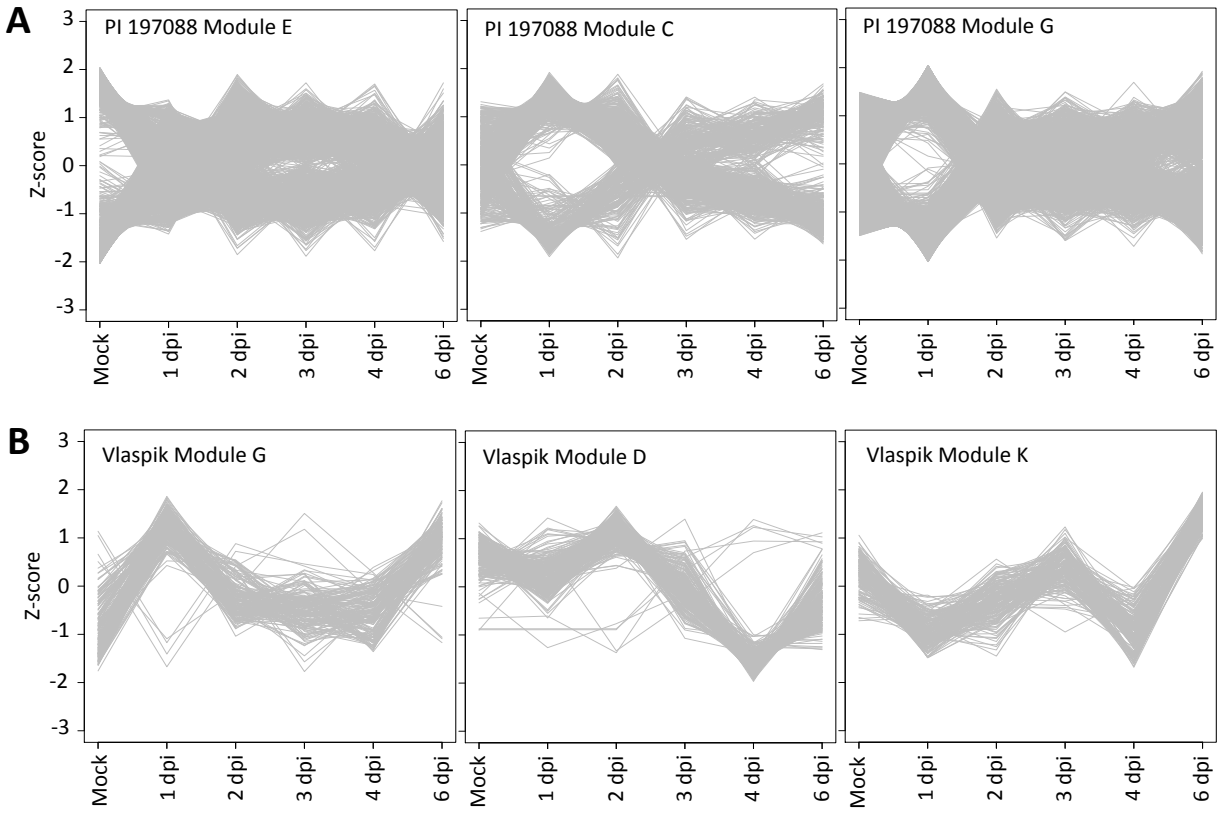


Figure 18. *Trend plots of normalized gene expression values from Vlaspik and PI 197088 grouped through weighted gene correlation network analysis.* Z-score indicates the number of standard of deviations away from the mean. Module A genes generally increase in expression in PI 197088 and decrease in expression in Vlaspik. Module B genes are either constitutively expression in Vlaspik and not PI 197088 or vice versa. Module D genes generally decrease in expression level in PI 197088 and increase in expression in Vlaspik over the time course. Module F genes are highly expressed in the mock of each plant line and more lowly expressed during the time course, but this decrease in expression occurs more rapidly in PI 197088 and increases at Vlaspik 6 dpi but not in PI 197088 6 dpi. dpi = days post inoculation.

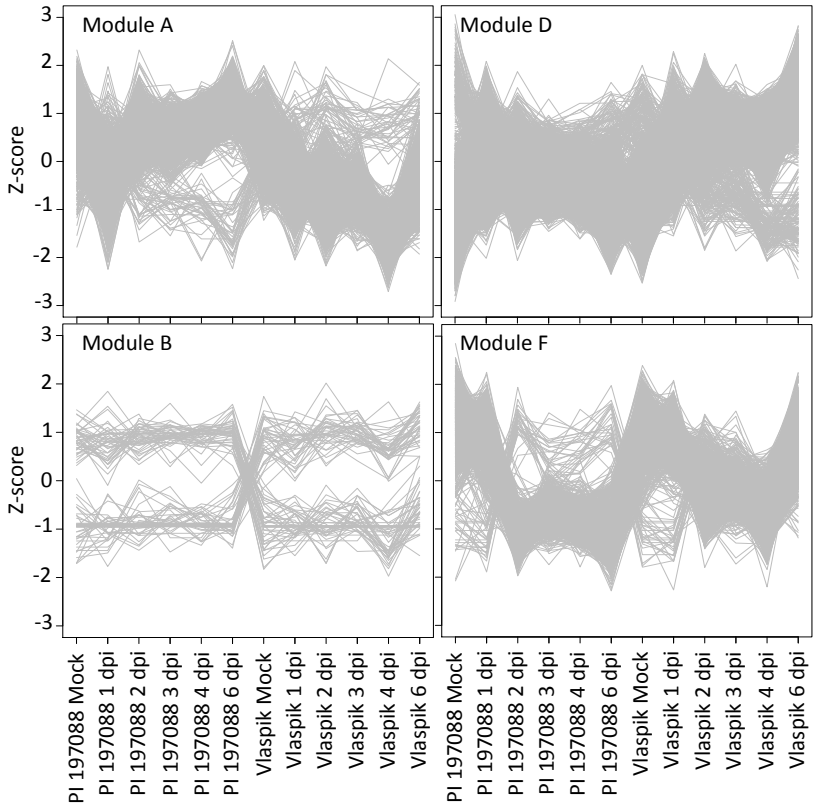


Figure 19. *Trend plots from the weighted gene correlation network analysis of normalized gene expression values from Vlaspik, PI 197088, and combined Vlaspik and PI 197088*. dpi = days post inoculation.

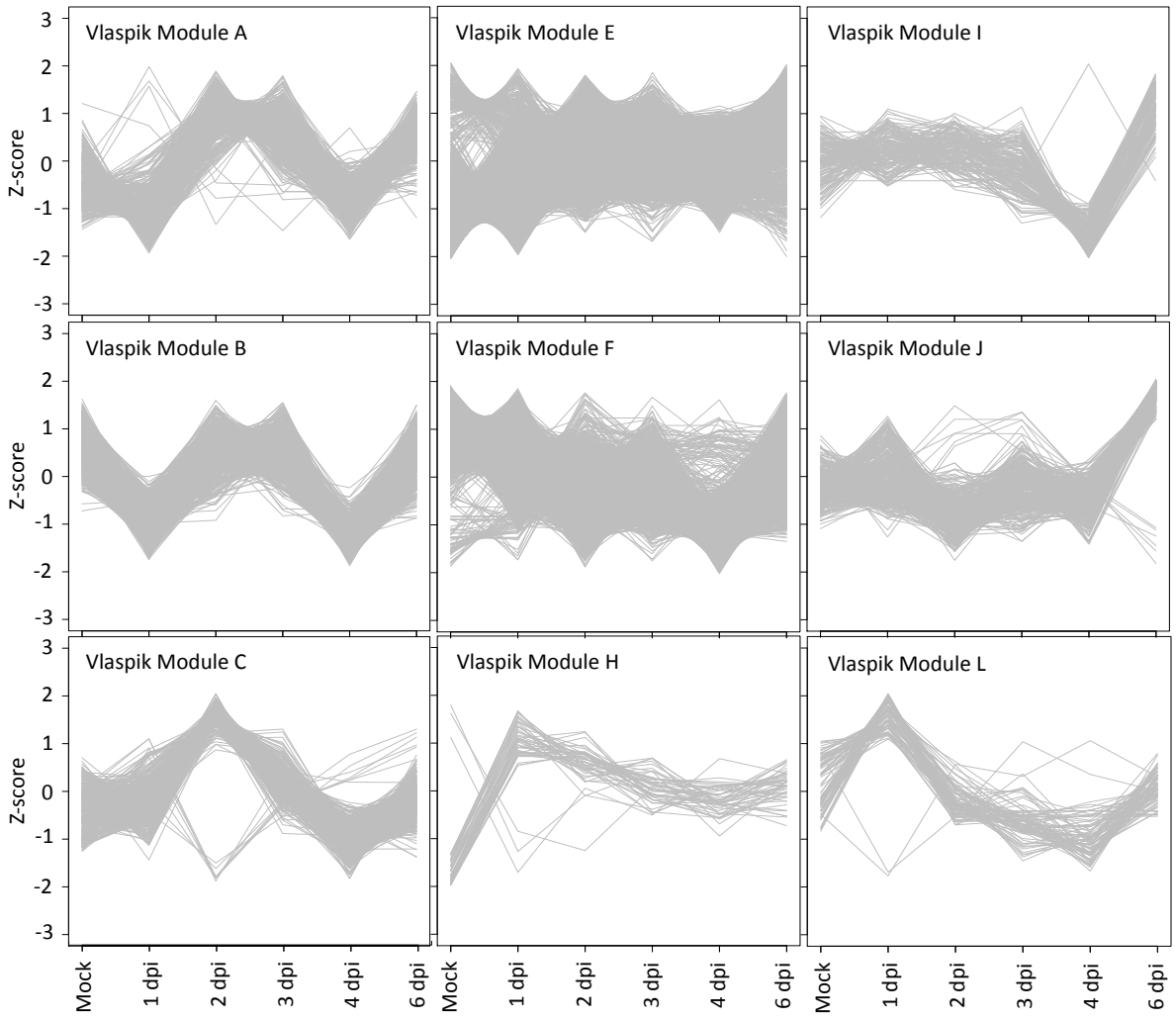


Figure 19 (cont'd)

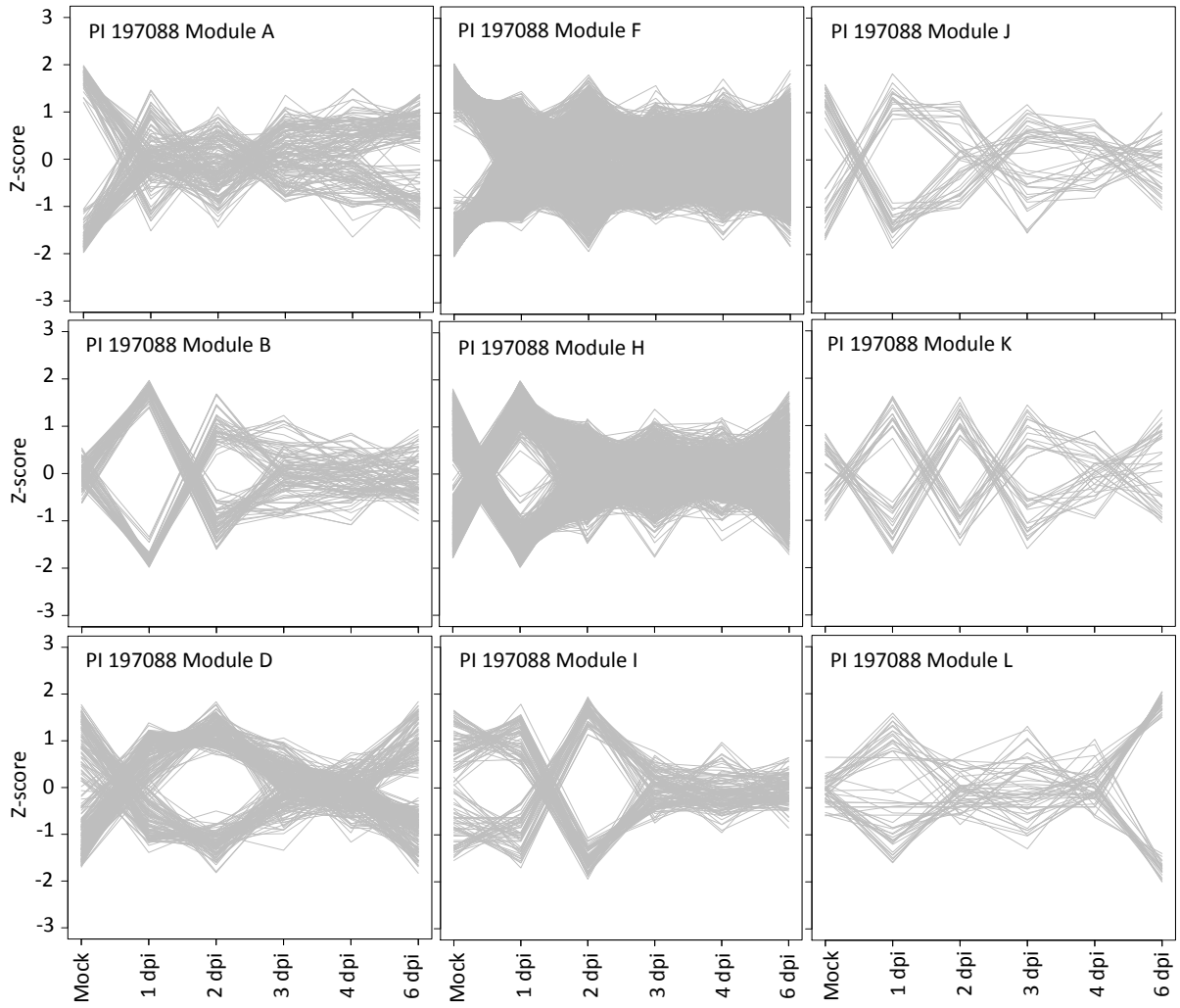
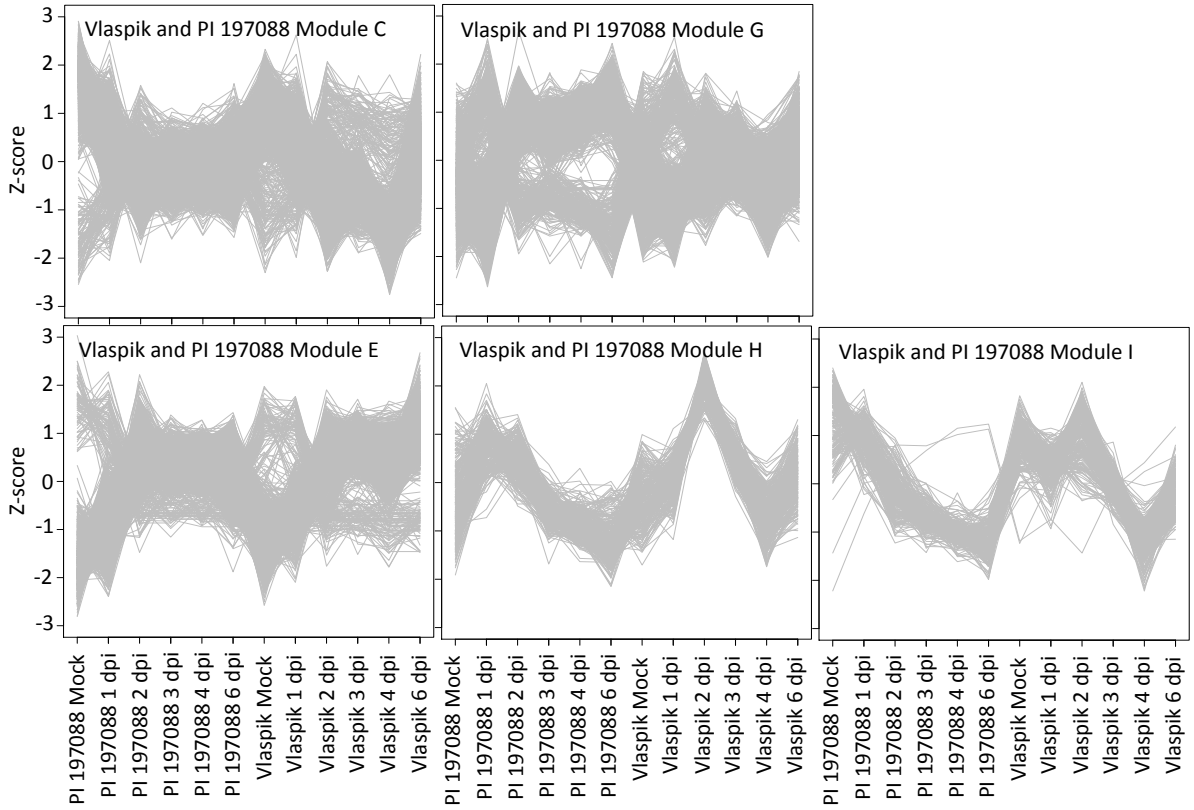


Figure 19 (cont'd)



Co-expression modules in Vlaspik and PI 197088 reveal differing patterns of gene regulation among significantly differentially expressed genes

In PI 197088, several modules consisted of genes with divergent patterns of expression at the mock time point and at early time points during infection (Fig. 17A, Fig. 19). In Module E, enriched GO terms in the genes that were down in the mock time point included DNA replication, phosphorylation, ubiquitination, and peroxidase activity, some of which could be involved in a signaling response to the pathogen (A). In Module C, enriched GO terms among genes there were highly expressed at 1 dpi included translation and glycolysis (Fig. 17A). The genes in Module C that were down at 1 dpi were enriched in terms including several transporters like cation transporters and ion transmembrane transporters. In Module G, genes that were up-regulated at 1 dpi have enriched GO terms that included intracellular transport of proteins, Golgi vesicle transport, and phospholipid transport (Fig. 17A). Finally, genes that were down at 1 dpi in Module G included GO terms enriched for photosynthesis and glycogen biosynthesis; these genes were later expressed at average or slightly above average levels suggesting that the plant's photosynthetic processes are not disrupted by the pathogen at later time points. Furthermore, two genes of interest in Module G involved in thiamine synthesis, Csa3G078800 (thiamine thiazole synthase) and Csa1G528590 (ThiC) are of particular interest as thiamine has been implicated in initiating a resistance response against fungi in rice and cucumber (Ahn et al. 2005) and against grapevine downy mildew (Boubakri et al. 2012). Mechanistically, thiamine has been shown to induce resistance through priming via a systemic acquired resistance (SAR) pathway in

Arabidopsis (Ahn et al. 2007) and is able to manipulate the phenylpropanoid pathway in grapes in resistance to downy mildew (Boubakri et al. 2013).

In Vlaspik, several modules consisted of genes that peaked in expression around 2 or 3 dpi and were then down-regulated, or at the baseline level of expression, before or after that time point (Fig. 17B, Fig. 19). For example, Module G showed patterns of a peak expression at 1 dpi and a tapering off of transcript level throughout the rest of the time course (Fig. 17B). Within Module G, multiple genes encoding transcription factors, including WRKYs and zinc finger proteins were classified. Similarly, in Module D, expression patterns peaked at 2 dpi. Among unique significant GO terms in this module were genes annotated to encode heat shock proteins, which are a group genes that have previously be associated with resistance (Li et al. 2012). Finally, Module K displayed a relative baseline level of transcription, with a decrease at ca. 4 dpi and an increase at 6 dpi (Fig. 17B). In this module, were several kinases, some of which are defense-associated leucine-rich receptor kinases (Antolín-Llovera et al. 2014; Belkhadir et al. 2014).

Co-expression of significant genes across genotypes shows distinct patterns of regulation during *P. cubensis* infection

The modules that combined the expression values of Vlaspik and PI 197088 were especially informative, as genes with opposite patterns of expression were clearly identified. For example, as shown in Figure 18, Modules A and D revealed opposite expression trends, in that in Module A, genes from PI 197088 were up-regulated, while

genes from Vlaspik were down-regulated. The observed trend in Module A was enriched in GO terms including photosynthesis, RNA methyltransferase activity, and amino acid biosynthesis. In contrast, the main trend Module D was enriched in several GO terms including glycolysis, protein catabolism and ubiquitination, serine/threonine kinase, and the endomembrane system. These apparent opposing expression trends support the hypothesis that photosynthesis in PI 197088 is relatively unperturbed, while Vlaspik must metabolize energy stores in response to pathogen growth. Among the genes in Module A were the aforementioned inducer of CBF expression and the N-hydroxycinnamoyl/benzoyltransferase. Additional genes of interest classified within Module A included isochorismate synthase (Csa1G008580) and thioredoxin genes (Fig. 16) (Wildermuth et al. 2001). Similarly, Module D also contained several genes of interest, including a suite of transcription factors, as well as a group of genes encoding proteins of unknown function that displayed strong patterns of differential expression between PI 197088 and Vlaspik. Significant genes in this module with lesser-known functions in resistance were a WAT1-related drug/metabolite transporter protein (Csa2G151580), a cupredoxin (Csa7G432520), and a lipase (Csa6G490910) that were expressed lowly in PI 197088 but was continually upregulated in Vlaspik (Fig. 16).

Additionally, Module B was interesting because it revealed gene expression patterns with extreme opposing patterns of expression (Fig. 18). For example, and of particular note, were two genes that encode protein kinases, Csa6G176420 and Csa6G176430, as well as a gene encoding a receptor-like protein kinase (Csa5G495960). In this module, each of these genes were identified as being highly expressed in PI 197088, and lowly, or not, expressed

in Vlaspik. In contrast, two glutamate receptors, Csa2G363540 and Csa2G363530, were expressed only in Vlaspik; this observation was surprising, as glutamate receptors have been proposed to have a role in pathogen recognition and subsequent resistance signaling (Forde and Roberts 2014). Module F revealed a difference in the timing of response to pathogen infection, as the genes in this module from PI 197088 showed a declining trend of expression at an earlier time point than the same genes in Vlaspik (Fig. 18). For the observed trend of Module F, several GO terms were enriched including intracellular protein transport, glycolipid metabolism, transcription, and protein phosphatase 2C (PP2C). PP2C is relevant because it serves as a negative regulator of ABA responses, and a resistance-associated hypersensitive response to ABA via the removal of a negative ABA regulator has been shown to confer resistance to downy mildew (Hok et al. 2014).

Candidate-based approach to identify resistance-associated genes

In addition to taking a non-directed approach to identify resistance-associated genes, we also employed a literature-informed approach to identify candidates. From the two QTL studies mapping downy mildew resistance, an overlapping QTL, *dm5.3*, was identified in both studies and was associated with a relatively high level to resistance (Yoshioka et al. 2014; Caldwell et al. 2014). As a result, DE genes located within *dm5.3* were evaluated as potential candidates for resistance. The two most differentially expressed genes in this region were two genes of unknown function that were highly expressed in the resistant line. Other common annotations significantly differentially expressed genes in this region included proteases and methyltransferases. Of particular interest within this group, and

among genes with previous annotations previously associated with resistance, were several transcription factors including Csa5G604920 that encodes a homeobox-leucine zipper protein that is homologous to ATHB13, a transcription factor in Arabidopsis that confers broad-spectrum resistance and is up-regulated at some time points in PI 197088 (Fig. 16)(Gao et al. 2014).

A candidate-based approach was also taken through the identification of cucumber homologs of genes previously demonstrated to function in downy mildew resistance signaling in other plant systems (Table 9). Interestingly, and in support of the work presented herein, many of these directed candidates were also identified through our bioinformatics analyses, above. For example, the homologs of two aminotransferase genes previously shown to confer resistance to *P. cubensis* in melon (Taler 2004) were identified as having significantly higher levels of expression in the resistant cucumber line (Fig. 16). Additionally, a lipoxygenase gene, Csa2G028490, had significantly higher expression in PI 197088 compared to Vlaspik, and previous work suggests that lipoxygenases may play a role disease resistance in cucumber (Fig. 16) (Huang et al. 2009). Finally, the thiamine thiazole synthase (*THI1*) gene, Csa3G078800, which showed over a 25-fold increase in PI 1970888 compared to Vlaspik at 6 dpi, represents a potentially interesting candidate, as its rice homolog, *OsDR8*, is an upstream regulator of the rice resistance pathway to both bacterial and fungal pathogens. Briefly, it was shown that reduced expression of this gene resulted in increased susceptibility, and that resistance could be rescued through the application of exogenous thiamine to plants (Wang et al. 2006). Furthermore, increased thiamine levels have been shown to influence SA-induced defense response and up-

regulate the phenylpropanoid pathway (Ahn et al. 2006; Boubakri et al. 2013). This finding is consistent with our data (Fig. 16) showing that isochorismate synthase (Csa1G008580) and Csa4G641000, a homolog of a gene involved in phytoalexin synthesis, are both up-regulated in the resistant line, PI 197088 (Fig. 18).

Table 9. *Candidate cucumber resistance-associated genes.*

Cucumber Gene	Gene ID	Plant Gene	Function and Resistance Association	Cucumber expression pattern	Citation
Csa5G604920	AT1G69780	ATHB13	HD-Zip transcription factor conferring broad-spectrum resistance	Significantly induced in PI 197088 compared to mock at 2-6 dpi, Significantly higher in PI 197088 compared to Vlaspik, Module G for PI 197088 and Vlaspik, located on *dm5.3*	Gao et al. 2014
Csa6G128000	AT4G03550	PMR4/GSL5	Negative regulator of plant defense	Significantly induced in Vlaspik at 4 and 6 dpi compared to mock, significantly higher in Vlaspik compared to PI 197088 at 4 and 6 dpi	Wawrzynska et al. 2010
Csa6G433300	AAL62332.1	AT2	Aminotransferase 2, overexpression in resistant melon	Significantly induced in Vlaspik and PI 197088 compared to mock at 2-6 dpi, Significantly higher in PI 197088 compared to Vlaspik at 2-6 dpi, Module G for PI 197088 and Vlaspik	Taler et al. 2004
Csa1G617370	AAL47679.1	AT1	Aminotransferase 1, overexpression in resistant melon	Significantly higher in PI 197088 compared to Vlaspik at 2-6 dpi, Module A for PI 197088 and Vlaspik	Taler et al. 2004
Csa2G028490	Csa2G028490	LOX	Lipoxygenase	Significantly up in PI 197088 at 6 dpi compared to mock, Significantly up in PI 197088 compared to Vlaspik at 2,4, and 6 dpi, Module A in PI 197088 and Vlaspik	Huang et al. 2009
Csa3G078800	OsDR8	THI1	Thiamine thiazole synthase	Significantly up in PI 197088 from 2-6 dpi compared to Vlaspik, PI 197088 Module G	Wang et al. 2006
Csa1G008580	AT1G74710	ICS1	Isochorismate synthase (SA synthesis)	Significantly up in PI 197088 from 2-6 dpi compared to Vlaspik, Module A for PI 197088 and Vlaspik	Wildermuth et al. 2001
Csa4G641000	AT2G19070	HCBT	Involved in phytoaxlexin synthesis	Significantly up in PI 197088 from 2-6 dpi compared to Vlaspik, Module A for PI 197088 and Vlaspik	Yang et al. 1997

Identification of small RNAs during the *C. sativus* – *P. cubensis* interaction

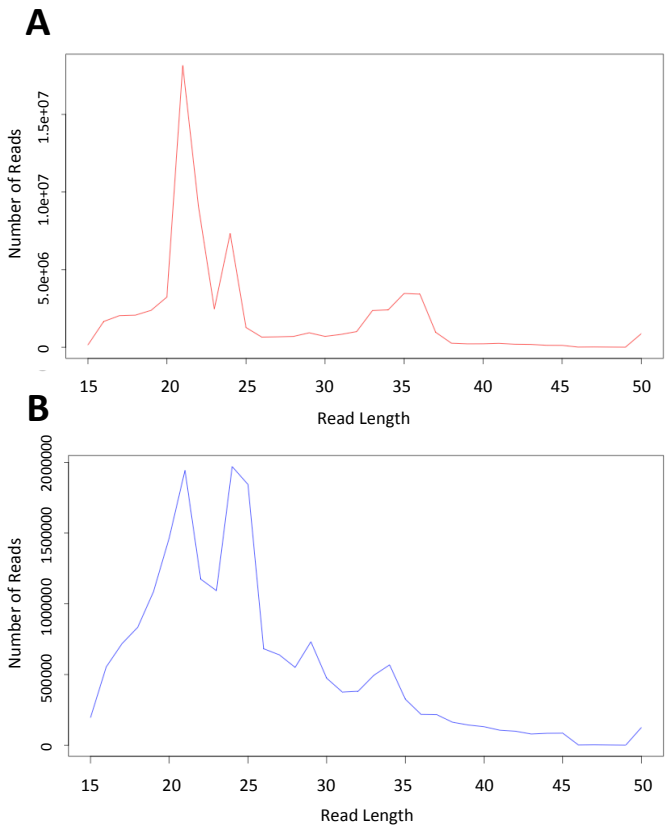
In parallel to the mRNA analyses described above, sRNAs were sequenced (from the same samples from which the mRNA was collected) to establish a foundation for defining the role of sRNAs in the regulation of the cucumber transcriptome during *P. cubensis* infection. The reads from both cucumber and *P. cubensis* were first mapped to the *P. cubensis* and *C. sativus* genomes to remove any reads that map to both species, as previously described (Weiberg et al. 2013). The percentage of reads that uniquely mapped to either *P. cubensis* or cucumber are shown in Table 10. The sample derived from *P. cubensis* sporangia consists of 73% of reads mapping to *P. cubensis.* The remaining 27% of reads mapping to cucumber is due to the contamination of the sporangia with leaf tissue, which is an unavoidable consequence of the collection process. Reads mapping to *P. cubensis* in the mock-inoculated cucumber sample (ca. 2-3%) aligned with chloroplast or plant-derived sequences, as well as some bacterial species, suggesting a low level of plant or bacterial sequence contamination in the previously published *P. cubensis* genome assembly, respectively. In the resistant interaction, the percentage of reads derived from *P. cubensis* is 24% at 1 dpi but decreases to less than 10% at all other time points, which is in agreement with our microscopy and qPCR analyses, indicating that the amount of pathogen decreases in PI 197088 over time. In contrast, the percentage of *P. cubensis* reads from infected Vlaspik samples comprises about 20% of all of the inoculated samples, an observation that is consistent with successful pathogen infection and in agreement with our microscopy and qPCR analyses.

Table 10. *Percentage of trimmed and quality filtered reads mapping to either* P. cubensis *genome or to the cucumber Chinese long genome v. 2 using Bowtie 1.0.0.* Sp. = sporangia.

Sample	*P. cubensis* Uniquely Mapped Reads	Cucumber Uniquely Mapped Reads	Total Uniquely Mapped Reads	Percent *P. cubensis* Reads	Percent Cucumber Reads
P. cubensis Sp.	8742311	3166726	11909037	73.4%	26.6%
PI 197088 Mock	178786	4855957	5034743	3.6%	96.4%
PI 197088 1 dpi	1534438	4804749	6339187	24.2%	75.8%
PI 197088 2 dpi	559919	5435147	5995066	9.3%	90.7%
PI 197088 3 dpi	396422	6117828	6514250	6.1%	93.9%
PI 197088 4 dpi	379417	4552862	4932279	7.7%	92.3%
PI 197088 6 dpi	378731	6018097	6396828	5.9%	94.1%
Vlaspik Mock	173278	7622556	7795834	2.2%	97.8%
Vlaspik 1 dpi	2223510	6793388	9016898	24.7%	75.3%
Vlaspik 2 dpi	1495032	5622572	7117604	21.0%	79.0%
Vlaspik 3 dpi	1361026	7762202	9123228	14.9%	85.1%
Vlaspik 4 dpi	1205324	5658894	6864218	17.6%	82.4%
Vlaspik 6 dpi	1286970	5151986	6438956	20.0%	80.0%

The distribution of all uniquely mapped reads for *P. cubensis* and cucumber shows strong enrichment for reads that are between 20-25 nucleotides (nt), which is consistent with the expected lengths of sRNA (Axtell 2013a) (Fig. 20). In plants, miRNAs are typically 20-24 nt with the mode being 21 nt; this is consistent with the read distribution shown in Figure 20. Other sRNAs within the distribution shown in Figure 20 include small interfering RNAs (siRNAs; 20-24nt), which are derived from dsRNA. Additionally, long siRNA (lsiRNAs; 30-40nt) which are siRNA-like sRNAs, have also been identified as being induced during bacterial infection (Katiyar-Agarwal et al. 2007). It is noteworthy that the enrichment of reads with a length of approximately 35 nt in the read distribution graph (Figure 20) may reflect an enrichment in lsiRNA due to a response to pathogen infection in cucumber. Small RNAs have been identified in other oomycetes and have been found to be enriched in populations of reads that were predominantly 21 or 25 nt, derived from either inverted repeats and miRNA, or transposable elements, respectively (Fahlgren et al. 2013). *Phytophthora infestans* sRNAs were also found to be enriched around 32 nt (Vetukuri et al. 2012). The read length distribution of *P. cubensis* is consistent with that of other oomycete plant pathogens (Fig. 20).

Figure 20. *Read length distribution of reads mapping to cucumber (A) or* P. cubensis *(B).*



Of the reads that uniquely mapped to *P. cubensis* or cucumber, the number of Dicer-derived sRNAs, and specifically miRNAs, were predicted and counted using ShortStack and HTSeq, respectively. Of the reads that uniquely mapped to each species, only a small percentage (less than 10%) mapped to predicted sRNA loci, and even fewer mapped to a predicted miRNA locus (Table 11). This reduction from the number of sequenced reads to the number of reads mapping to sRNA loci has been observed in other studies and (Visser et al. 2014; Martínez et al. 2011; Mao et al. 2012) and the number of reads used to predict and quantify sRNAs in this study was much higher than that used in previous studies identifying sRNAs in cucumber (Martínez et al. 2011; Mao et al. 2012, Li et al. 2013, Jin et al. 2015). In addition, the greater number of *P. cubensis*-derived sRNA and miRNA loci in the inoculated Vlaspik samples compared to the inoculated PI 197088 samples further supports that *P. cubensis* growth is greatly reduced in the resistant plant.

Table 11. *Read processing from total sequenced reads to total reads mapped to miRNA loci in* P. cubensis*, Vlaspik, and PI 197088.*

Category of Reads	Sporangia	Vlaspik Mock	Inoculated Vlaspik	PI 197088 Mock	Inoculated PI 197088
Total Sequenced Reads	28,775,108	22,199,524	118,004,094	15,511,102	102,910,883
Quality-filtered Reads	20,022,634	16,236,220	77,997,535	10,158,836	67,422,425
Reads Unique to Cucumber		7,622,556	30,989,042	4,855,957	26,928,683
Cucumber Dicer-derived sRNA		3,651,841	11,075,300	2,771,988	12,516,322
Cucumber miRNA		1,265,624	2,898,729	820,587	3,082,744
Reads Unique to *P. cubensis*	8,742,311		7,571,862		3,248,927
P. cubensis Dicer-derived sRNA	803,862		390,125		54,608
P. cubensis miRNA	199,165		93,174		7,801

miRNA analysis in *C. sativus* reveals conserved and novel loci

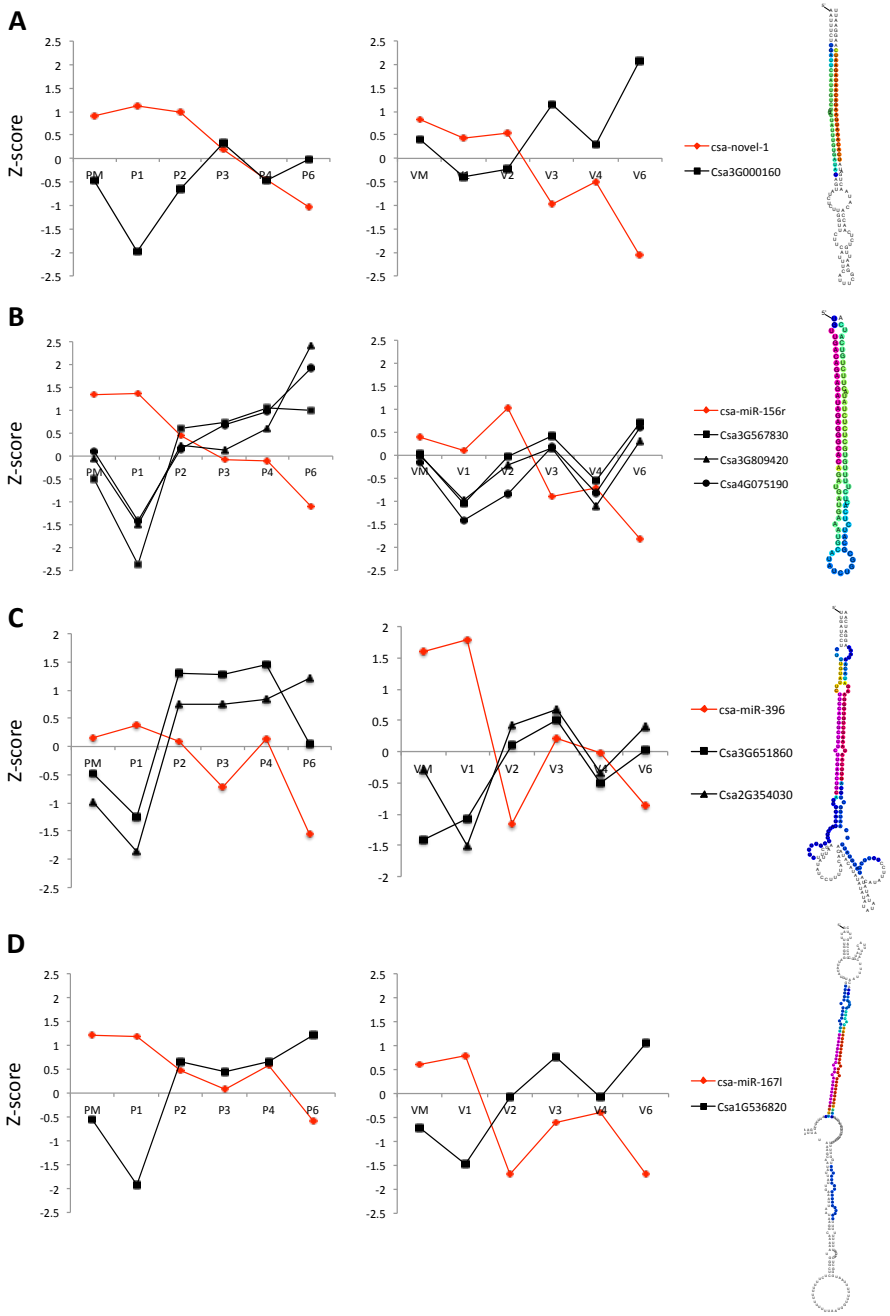
By analyzing pooled reads from both plant lines, novel miRNA and miRNA families previously identified in cucumber and other species were found. As shown in Table 12, 28 conserved miRNA families were identified in this study, including 23 families that were previously identified in cucumber (Mao et al. 2012; Martínez et al. 2011; Li et al. 2013; Jin et al. 2015). Of the remainder, 4 families were previously identified in other species recorded in miRBase, yet not in cucumber. Many of these conserved families, notably miR156, miR166, miR167, and miR171, were all represented by several loci and mature miRNA from the cucumbers analyzed in this study. In addition, many of the conserved miRNA target transcription factors; for example, miR156, miR157, miR169, and miR171 all target squamosa promoter-binding transcription factors, miR159 and miR319 target MYB transcription factors, and miR166 targets Class III homeobox leucine zipper proteins. We posit that miRNA targeting of transcription factors is relevant, as these conserved miRNAs could play a role in signaling pathways involved in resistance during infection. Other notable targets of the identified conserved miRNAs included auxin-related genes, targeted by miR160, the auxin responsive factor, targeted by miR167, and additional auxin signaling genes, targeted by miR393. We hypothesize that this additional layer of regulation is relevant to the host-pathogen interaction, as a number of studies have demonstrated that auxin and SA function in apparent opposing ways to regulate the tradeoff between plant growth and defense (Huot et al. 2014). Finally, and of particular noteworthiness, was the identification of miR482, which was first discovered in cucumber in this study and was predicted to target a CC-NBS-LRR *R*-gene.

Table 12. *Conserved miRNA families from cucumber or in other plant species in miRBase.* The number of loci and mature miRNA were identified in this study using ShortStack and comparing the predicted mature miRNA to miRBase.

Conserved miRNA family	Loci in Cucumber	Mature miRNA	*Cucumis sativus*	*Arabidopsis thaliana*	*Vitis vinifera*	*Cucumis melo*	*Glycine max*
miR156	6	13	+	+	+	+	+
miR157	1	1	+	+			
miR159	1	3	+	+	+	+	+
miR160	1	1	+	+	+	+	+
miR162	1	3	+	+	+	+	+
miR164	3	3	+	+	+	+	+
miR166	5	10	+	+	+	+	+
miR167	6	12	+	+	+	+	+
miR168	2	5	+	+	+	+	+
miR169	5	5	+	+	+	+	+
miR171	6	10	+	+	+	+	+
miR172	2	3	+	+	+	+	+
miR319	3	8	+	+	+	+	+
miR390	2	2	+	+	+	+	+
miR393	2	4	+	+	+	+	+
miR396	2	6	+	+	+	+	+
miR398	2	4	+	+	+	+	+
miR399	4	6	+	+	+	+	+
miR408	1	3	+	+	+	+	+
miR477	1	1	+		+	+	
miR482	1	1			+		+
miR827	1	3	+	+			
miR1175	1	1					
miR2111	1	1	+	+	+	+	+
miR2950	1	1	+		+		
miR3627	2	2			+		
miR7741	1	1					
miR8210	1	1					

In addition to identifying previously described miRNA families, 57 novel mature miRNAs were identified. This represents a significant expansion in the number of cucumber miRNAs identified to date, and we posit that this increase in the number of miRNAs is likely attributable to the deeper miRNA sequencing undertaken in the current study. Indeed, previous analyses generated 209,331 (Martínez et al. 2011), 4,012,509 (Mao et al. 2012), or 27,765,704 (Jin et al. 2015) high quality reads, whereas we generated over 70 million high quality cucumber reads (Table 11). An additional, or alternative factor, influencing this increased number may also be a reflection of biotic stress induction with both a resistant and susceptible plant (Inal et al. 2014). Among the novel cucumber miRNA found in this study, four were previously identified as novel miRNA or miRNA* sequences in other studies (Martínez et al. 2011, Li et al. 2013, Jin et al. 2015). Among the predicted targets of novel cucumber miRNAs was an mildew resistance locus O (MLO)-like protein, which has a role in mediating powdery mildew defense (Fig. 21) (Acevedo-Garcia et al. 2014). This predicted target suggests a role for newly discovered miRNAs in plant defense. In addition, other miRNAs unique to this study are predicted to target various kinases, some of which may have a role in biotic stress signaling. In addition, many miRNAs have unidentified targets; this could indicate that some of the miRNAs are lineage-specific miRNAs that are typically lower abundance and have poorly defined targets (Ma et al. 2010; Fahlgren et al. 2010; Axtell 2013a).

Figure 21. *Z-score plots of normalized expression values of predicted cucumber miRNA and corresponding predicted cucumber target genes.* (A) Novel cucumber miRNA from csa-novel-1 (Cluster 11080) targeting CsaG000160. (B) csa-miR-156r (Cluster 11464) targeting Csa3G567830, Csa3G809420, and Csa4G075190. (C) csa-miR-396 (Cluster 7011) targeting Csa3G651860 and Csa2G354030. (D) csa-miR-167l (Cluster 6579) targeting Csa1G536820. P1 = PI 197088 at 1 dpi, P2 = PI 197088 at 2 dpi, P3 = PI 197088 at 3 dpi, P4 = PI 197088 at 4 dpi, P6 = PI 197088 at 6 dpi, V1 = Vlaspik at 1 dpi, V2 = Vlaspik at 2 dpi, V3 = Vlaspik at 3 dpi, V4= Vlaspik at 4 dpi, V6 = Vlaspik at 6 dpi; dpi = days post inoculation. The miRNA precursor structure is shown with the red and pink regions indicating the location of the mature miRNA.



miRNA regulation of Vlaspik and PI 197088 targets during *P. cubensis* infection

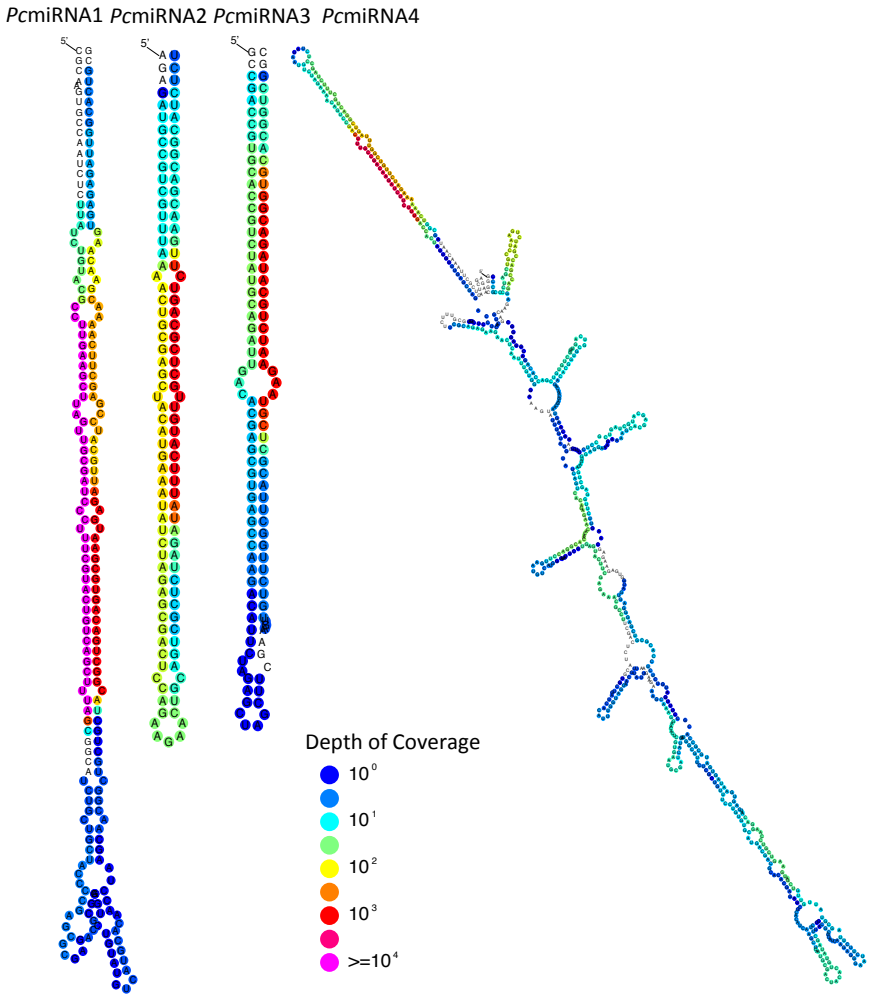
After identifying the suite of conserved and novel miRNAs, described above, we analyzed the miRNA count data for each sample to identify significantly DE miRNAs. 27 miRNA loci were predicted to significantly change in expression over time and ten miRNA loci were predicted to be differentially expressed between the plant lines. For the miRNAs for which targets were predicted, the Z-scores of the normalized expression of the miRNA were compared to the Z-scores of the normalized expression of the target cucumber transcript. Using this information, opposing patterns in gene regulation were used in support of the hypothesis that the miRNA targets were correctly predicted. For example, and as shown in Figure 21, miRNAs, as well as their predicted target gene expression patterns, suggest that the transcripts of the targets may be regulated by miRNA such that when the miRNA level is high, a parallel reduction in the target transcript level is observed. Conversely, when the miRNA level is reduced, an increased transcript level is observed. For example, as shown in Figure 21A, the novel mature miRNA predicted to target Csa3G000160, an MLO-like protein, showed an expression pattern that was oppositely correlated with the expression of the MLO-like transcript. This MLO-like gene has membership in Vlaspik Module K and PI 197088 Module C (Fig. 17). Additionally, csa-miR-156r had an expression pattern that oppositely correlated with expression pattern of the predicted targets encoding squamosa promoter binding proteins (Csa3G567830 and Csa3G809420) (Fig. 21B). These genes were significantly DE genes and were in the Vlaspik-PI 197088 Module G (Fig. 19). Likewise, csa-miR-396 had evidence of targeting Csa3G651860 and Csa2G354030, which are both

growth promoting factors and are also members of Module G from Vlaspik-PI197088 (Fig. 21C, Fig. 19). Finally, csa-miR-167l targeted an auxin response factor encoded by Csa1G536820, which was in the Vlaspik-PI 197088 Module G (Fig. 21D, Fig. 19).

Discovery of miRNA loci in *P. cubensis*

In the Arabidopsis-*Botrytis* interaction, sRNAs from the pathogen have been shown to directly target host defense-related genes (Weiberg et al. 2013). To determine if *P. cubensis*-derived sRNA exist, we sequenced the small RNAome of *P. cubensis* and predicted cucumber target genes. Of the sRNA loci identified in *P. cubensis*, 4 new miRNA loci and 11 mature miRNA were identified that were unique to *P. cubensis*. Similarly, 6 miRNA loci were observed in another oomycete, *P. infestans,* which is consistent with the number of miRNA loci in *P. cubensis* (Vetukuri et al. 2012). The miRNA loci identified in *P. cubensis* all displayed characteristic hairpin loop precursors (Fig. 22). For each predicted miRNA loci, 1-3 mature miRNA were identified and the 5' nucleotide of each miRNA was either a uracil or a cytosine.

Figure 22. *Predicted miRNA precursor structures from* P. cubensis*.* Colored bases indicate the depth of small RNA read coverage mapping to each base in the miRNA model.



In a paradigm-shifting study involving the fungus *Botrytis cinerea* and Arabidopsis, it was shown that fungal miRNA can be transported from the pathogen into the plant to target and silence genes involved in immunity (Weiberg et al. 2013). Based on this observation, potential cucumber targets of *P. cubensis* miRNA were identified (Table 13). Two predicted *P. cubensis* miRNA, *Pc*miRNA2 and *Pc*miRNA4, had predicted cucumber genes targets while the remaining two *P. cubensis* miRNA did not. Interestingly, *Pc*miRNA2 was predicted to target the TIR-NBS-LRR cucumber gene, Csa5G647510. Future work will be needed to confirm the mode of action and specificity of the predicted miRNA and targets.

Table 13. *Sequences of miRNA from* P. cubensis *identified using ShortStack.* Targets from the cucumber genome predicted using TAPIR and psRNATarget.

miRNA	Mature miRNA	Arm	Length	# Reads	Cucumber Target	miRNA*
*Pc*miRNA1	CCUUGAAGCUUAGUUGCGAUC	5'	21	205010	none	UUGCAUCCGAGCUUCAAAACG
	CCUUUCGUACUGUCAGCUUUA	5'	20	28686	none	CGGCUGACAGUGCGAAUGAGA
	CCUUUCGUACUGUCAGCUUU	5'	20	13543	none	GGCUGACAGUGCGAAUGAGA
	CUUUCGUACUGUCAGCUUUA	5'	21	1575	none	CGGCUGACAGUGCGAAUGAG
*Pc*miRNA2	UUUCAUGUUGCUCGCAGUCU	3'	20	480	Csa5G647510 Csa1G537560	ACUGCGAGCUACAUGAAAUA
	UAUUUCAUGUUGCUCGCAGUC	3'	21	186	Csa5G587140	CUGCGAGCUACAUGAAAUAUC
	UUUCAUGUUGCUCGCAGUCUU	3'	21	45	Csa1G537560	AACUGCGAGCUACAUGAAAUA
*Pc*miRNA3	CUGCAUAGACGGUGCACGGUC	3'	21	9	none	CCGUGCACCGUCUAUGCAGAU
*Pc*miRNA4	UUUGCCGACAAGAUCUUCUUU	3'	21	770	none	AAAAGAUCUUGUCGGUUUAGU
	CUUUGCCGACAAGAUCUUCU	3'	20	301	Csa1G600170 Csa2G432220	AAGAUCUUGUCGGUUUAGUC
	UUUGCCGACAAGAUCUUCUU	3'	20	160	none	AAAGAUCUUGUCGGUUUAGU

Conclusions

Based on the sum of the work presented herein, we posit that the timing of the plant response is key to resistance in response to *P. cubensis* infection, as our data show that the resistant line PI 197088 displayed a faster, more robust, response to pathogen infection. For example, the number of genes differentially expressed at early time points and genes grouped in modules indicated that large-scale change in transcription occurred more rapidly in PI 197088 in response to *P. cubensis.* In addition, our observed regulation of several host gene groups and pathways suggested that resistance in PI 197088 might be attributed to a combination of hormones and nutrient regulation. For example, genes associated with SA and vitamin B_1 (thiamine) synthesis were strongly differentially regulated between the resistant and susceptible plant over the time course of infection. However, the work from this study focused only on transcript data and was unable to identify which genes may be a cause or effect of the resistance response in PI 197088. Future studies are necessary to investigate the *in planta* protein and metabolite levels of genes of interest to further test the hypotheses developed in this study. In addition to sequencing transcripts, this work also sequenced sRNA and predicted miRNAs in cucumber and *P. cubensis* that contribute to the complex interactions between the host and pathogen. Many miRNAs had predicted cucumber targets that could mediate signal transduction pathways within cucumber. In total, this work provides a foundation for future studies aimed at defining the functional role and activity for the core set of candidate resistance-associated genes and miRNAs uncovered through this study.

Acknowledgements

AB is supported by a Michigan State University Graduate Student Distinguished Fellowship. Research in the laboratory of BD is supported by funding from the Michigan State University Rackham Foundation, Michigan State University Project GREEEN (GR13-007), and a grant from the United States Department of Agriculture Specialty Crops Research Initiative (2011-51181-30661). We would like to thank Kevin Childs and Robin Buell for valuable discussions at the onset and throughout the research presented herein.

Author-recommended supplemental files

http://www.daylab.plp.msu.edu/aburkardt-supporting-files

Username: data

Password: ABManuscriptFiles

REFERENCES

REFERENCES

Acevedo-Garcia, J., Kusch, S., and Panstruga, R. 2014. Magical mystery tour: MLO proteins in plant immunity and beyond. *New Phytol*. 204:273–281.

Adhikari, B. N., Hamilton, J. P., Zerillo, M. M., Tisserat, N., Lévesque, C. A., and Buell, C. R. 2013. Comparative genomics reveals insight into virulence strategies of plant pathogenic oomycetes. *PLoS ONE* 8:e75072. Published online.

Adhikari, B. N., Savory, E. A., Vaillancourt, B., Childs, K. L., Hamilton, J. P., Day, B., and Buell, C. R. 2012a. Expression profiling of *Cucumis sativus* in response to infection by *Pseudoperonospora cubensis*. *PLoS ONE* 7:e34954. Published online.

Ahn, I-P., Kim, S., and Lee, Y-H. 2005. Vitamin B_1 functions as an activator of plant disease resistance. *Plant Phys.* 138:1505-1515.

Ahn, I-P., Kim, S., Lee, Y-H., and Suh, S-C. 2007. Vitamin B_1-induced priming is dependent on hydrogen peroxide and the *NPR1* gene in Arabidopsis. *Plant Phys.* 143:838-848.

Anders, S., McCarthy, D. J., Chen, Y., Okoniewski, M., Smyth, G. K., Huber, W., and Robinson, M. D. 2013. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc*. 8:1765–1786.

Anders, S., Theodor, P., and Huber, W. 2014. HTSeq - a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166-169.

Antolín-Llovera, M., Petutsching, E. K., Ried, M. K., Lipka, V., Nürnberger, T., Robatzek, S., and Parniske, M. 2014. Knowing your friends and foes - plant receptor-like kinases as initiators of symbiosis or defence. *New Phytol.* 204:791–802.

Axtell, M. J. 2013a. Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* 64:137–159.

Axtell, M. J. 2013b. ShortStack: Comprehensive annotation and quantification of small RNA genes. *RNA* 19:740–751.

Belkhadir, Y., Yang, L., Hetzel, J., Dangl, J. L., and Chory, J. 2014. The growth–defense pivot: crisis management in plants mediated by LRR-RK surface receptors. *Trends Biochem. Sci.* 39:447-456.

Bonnet, E., He, Y., Billiau, K., and Van de Peer, Y. TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. *Bioinformatics* 26:1566-1568.

Boubakri, H., Wahab, M.A., Chong, J., Bertsch, C., Mliki, A., and Soustre-Gacougnolle, I. 2012. Thiamine induced resistance to *Plasmopara viticola* in grapevine and elicited host-defense responses, including HR like-cell death. *Plant Phys. Biochem*. 57:120-133.

Boubakri, H., Poutaraud, A., Wahab, M.A., Clayeux, C., Baltenweck-Guyot, R., Steyer, D., Marcic, C., Mliki, A., Soustre-Gacougnolle, I. 2013. Thiamine modulates metabolism of the phenylpropanoid pathway leading to enhanced resistance to *Plasmopara viticola* in grapevine. *BMC Plant Biol.* 13:31. Published online.

Burkhardt, A., Buchanan, A., Cumbie, J. S., Savory, E. A., Chang, J. H., and Day, B. 2015. Alternative splicing in the obligate biotrophic oomycete pathogen *Pseudoperonospora cubensis*. *Mol. Plant-Microbe Interact.* 28:298-309.

Caldwell, D., Chan, E., de Vries, J., Joobeur, T., King, J., Reina, A., and Shetty, N. V. 2014. Methods and compositions for identifying downy mildew resistance cucumber plants. US Patent 8809622. Date issued: August 19.

Call, A. D., Criswell, A. D., Wehner, T. C., Ando, K., and Grumet, R. 2012a. Resistance of cucumber cultivars to a new strain of cucurbit downy mildew. *HortScience* 47:171-178.

Call, A. D., Criswell, A. D., Wehner, T. C., Klosinska, U., and Kozik, E. U. 2012b. Screening cucumber for resistance to downy mildew caused by (Berk. and Curt.) Rostov. *Crop Sci.* 52:577-592.

Childs, K. L., Davidson, R. M., and Buell, C. R. 2011. Gene coexpression network analysis as a source of functional annotation for rice genes. *PLoS ONE* 6(7):e22196. Published online.

Coates, M. E., and Beynon, J. L. 2010. *Hyaloperonospora arabidopsidis* as a pathogen model. A*nnu. Rev. Phytopathol.* 48:329-345.

Coca, M., and San Segundo, B. 2010. AtCPK1 calcium-dependent protein kinase mediates pathogen resistance in Arabidopsis. *Plant J.* 63:526–540.

Dai, X. and Zhao, P.X. 2011. psRNATarget: a plant small RNA target analysis server. *Nucl. Acids Res.* 39:W155-9. Published online.

Fahlgren, N., Bollmann, S. R., Kasschau, K. D., Cuperus, J. T., Press, C. M., Sullivan, C. M., Chapman, E. J., Hoyer, J. S., Gilbert, K. B., Grunwald, N. J., and Carrington, J. C. 2013. *Phytophthora* have distinct endogenous small RNA populations that include short interfering and microRNAs. *PLoS ONE* 8(10):e77181. Published online.

Fahlgren, N., Jogdeo, S., Kasschau, K. D., Sullivan, C. M., Chapman, E. J., Laubinger, S., Smith, L. M., Dasenko, M., Givan, S. A., Weigel, D., and Carrington, J. C. 2010. MicroRNA gene evolution in *Arabidopsis lyrata* and *Arabidopsis thaliana*. *Plant Cell* 22:1074–1089.

Forde, B. G., and Roberts, M. R. 2014a. Glutamate receptor-like channels in plants: A role as amino acid sensors in plant defence. *F1000Prime Rep*. 6:37. Published online.

Furniss, J. J., and Spoel, S. H. 2015. Cullin-RING ubiquitin ligases in salicylic acid-mediated plant immune signaling. *Front. Plant Sci.* 6:154. Published online.

Gao, D., Appiano, M., Huibers, R. P., Chen, X., Loonen, A. E. H. M., Visser, R. G. F., Wolters, A. M. A., and Bai, Y. 2014. Activation tagging of ATHB13 in *Arabidopsis thaliana* confers broad-spectrum disease resistance. *Plant Mol. Biol.* 86:641-653.

Hammerschmidt, R. 1999. Phytoalexins: What have we learned after 60 years? *Annu. Rev. Phytopathol.* 37:285-306.

Hok, S., Allasia, V., Andrio, E., Naessens, E., Ribes, E., Panabieres, F., Attard, A., Ris, N., Clement, M., Barlet, X., Marco, Y., Grill, E., Eichmann, R., Weis, C., Huckelhoven, R., Ammon, A., Ludwig-Muller, J., Voll, L. M., and Keller, H. 2014. The receptor kinase IMPAIRED OOMYCETE SUSCEPTIBILITY1 attenuates abscisic acid responses in Arabidopsis. *Plant Physiol.* 166:1506-1518.

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

Hu, J., Sun, L., Zhu, Z., Zheng, Y., Xiong, W., and Ding, Y. 2014. Characterization of conserved microRNAs from. *Biochimie.* 102:137-144.

Hu, Y., Jiang, L., Wang, F., and Yu, D. 2013. Jasmonate regulates the INDUCER OF CBF EXPRESSION-C-REPEAT BINDING FACTOR/DRE BINDING FACTOR1 cascade and freezing tolerance in Arabidopsis. *Plant Cell* 25:2907-2924.

Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., Lucas, W. J., Wang, X., Xie, B., Ni, P., Ren, Y., Zhu, H., Li, J., Lin, K., Jin, W., Fei, Z., Li, G., Staub, J., Kilian, A., van der Vossen, E. A. G., Wu, Y., Guo, J., He, J., Jia, Z., Ren, Y., Tian, G., Lu, Y., Ruan, J., Qian, W., Wang, M., Huang, Q., Li, B., Xuan, Z., Cao, J., Asan, Wu, Z., Zhang, J., Cai, Q., Bai, Y., Zhao, B., Han, Y., Li, Y., Li, X., Wang, S., Shi, Q., Liu, S., Cho, W. K., Kim, J.-Y., Xu, Y., Heller-Uszynska, K., Miao, H., Cheng, Z., Zhang, S., Wu, J., Yang, Y., Kang, H., Li, M., Liang, H., Ren, X., Shi, Z., Wen, M., Jian, M., Yang, H., Zhang, G., Yang, Z., Chen, R., Liu, S., Li, J., Ma, L., Liu, H., Zhou, Y., Zhao, J., Fang, X., Li, G., Fang, L., Li, Y., Liu, D., Zheng, H., Zhang, Y., Qin, N., Li, Z., Yang, G., Yang, S., Bolund, L., Kristiansen, K., Zheng, H., Li, S., Zhang, X., Yang, H., Wang, J., Sun, R., Zhang, B., Jiang, S., Wang, J., Du, Y., and Li, S. 2009. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41:1275-1281.

Huot, B., Yao, J., Montgomery, B. L., and He, S. Y. 2014. Growth-defense tradeoffs in plants: A balancing act to optimize fitness. *Mol. Plant* 7:1267-1287.

Inal, B., Türktas, M., Erem, H., Ilhan, E. , Okay, S., Atak, M., Erayman, M., and Unver, T. 2014. Genome-wide fungal stress responsive miRNA expression in wheat. *Planta.* 240:1287-1298.

Jin, W. and Wu, F. 2015. Identification and characterization of cucumber microRNAs in response to *Pseudoperonospora cubensis* infection. *Gene* 569: 225-232.

Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A., and Jin, H. 2007. A novel class of bacteria-induced small RNAs in Arabidopsis. *Genes Dev.* 21:3123-3134.

Knight, H. 2004. Abscisic acid induces CBF gene transcription and subsequent induction of cold-regulated genes via the CRT promoter element. *Plant Physiol.* 135:1710-1717.

Krasileva, K. V., Zheng, C., Leonelli, L., Goritschnig, S., Dahlbeck, D., and Staskawicz, B. J. 2011. Global analysis of Arabidopsis/downy mildew interactions reveals prevalence of incomplete resistance and rapid evolution of pathogen recognition. *PLoS ONE* 6:e28765. Published online.

Langfelder, P., and Horvath, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. Published online.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Gen. Bio.* 10:R25. Published online.

Lapin, D., Meyer, R. C., Takahashi, H., Bechtold, U., and Van den Ackerveken, G. 2012. Broad-spectrum resistance of Arabidopsis C24 to downy mildew is mediated by different combinations of isolate-specific loci. *New Phytol.* 196:1171-1181.

Li, C., Li, Y., Bai, L., Zhang, T., He, C., Yan, Y., and Yu, X. 2015. Grafting-responsive miRNAs in cucumber and pumpkin seedlings identified by high-throughput sequencing at whole genome level. *Physiologia Platarum* 151:406-422.

Li, H., Handsaker, B., Wysoker, A., Fennel, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078-2079.

Li, J., Zhang, H., Hu, J., Liu, J., and Liu, K. 2012. A heat shock protein gene, CsHsp45.9, involved in the response to diverse stresses in cucumber. *Biochem. Genet*. 50:565–578.

Love, M.I., Huber, W., and Simon, A. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Gen. Bio*. 15:550. Published online.

Lu, Y-J., Schornack, S,, Spallek, T., Geldner, N., Chory, J., Schellmann, S., Schumacher, K., Kamoun, S., and Robatzek, S. 2012. Patterns of plant subcellular responses to successful oomycete infections reveal differences in host cell reprogramming and endocytic trafficking. *Cell. Microbiol.* 14: 682-697.

Ma, Z., Coruh, C., and Axtell, M. J. 2010. *Arabidopsis lyrata* small RNAs: Transient miRNA and small interfering RNA loci within the Arabidopsis genus. *Plant Cell* 22:1090-1103.

Mao, W., Li, Z., Xia, X., Li, Y., and Yu, J. 2012. A combined approach of high-throughput sequencing and degradome analysis reveals tissue specific expression of microRNAs and their targets in cucumber. *PLoS ONE* 7:e33040. Published online.

Marchive, C., Léon, C., Kappel, C., Coutos-Thévenot, P., Corio-Costet, M.-F., Delrot, S., and Lauvergeat, V. 2013. Over-expression of VvWRKY1 in grapevines induces expression of jasmonic acid pathway-related genes and confers higher tolerance to the downy mildew. *PLoS ONE* 8:e54185. Published online.

Martínez, G., Forment, J., Llave, C., Pallás, V., and Gómez, G. 2011. High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs. *PLoS ONE* 6:e19523. Published online.

Merz, P. R., Moser, T., Höll, J., Kortekamp, A., Buchholz, G., Zyprian, E., and Bogs, J. 2014. The transcription factor VvWRKY33 is involved in the regulation of grapevine (*Vitis vinifera*) defense against the oomycete pathogen *Plasmopara viticola*. *Physiol. Plantarum* 153:365-80.

Mohr, T. J., Mammarella, N. D., Hoff, T., Woffenden, B. J., Jelesko, J. G., and McDowell, J. M. 2010. The Arabidopsis downy mildew resistance gene RPP8 is induced by pathogens and salicylic acid and is regulated by W box cis elements. *Mol. Plant Microbe Interact.* 23:1303-1315.

Nemri, A., Atwell, S., Tarone, A. M., Huang, Y. S., Zhao, K., Studholme, D. J., Nordborg, M., and Jones, J. D. G. 2010. Genome-wide survey of Arabidopsis natural variation in downy mildew resistance using combined association and linkage mapping. *Proc. Natl. Acad. Sci. U. S. A.* 107:10302–10307.

Porter, K., Shimono, M., Tian, M., and Day, B. 2012. Arabidopsis actin-depolymerizing factor-4 links pathogen perception, defense activation and transcription to cytoskeletal dynamics. *PLoS Pathog.* 8:e1003006. Published online.

Qiao, Y., Liu, L., Xiong, Q., Flores, C., Wong, J., Shi, J., Wang, X., Liu, X., Xiang, Q., Jiang, S., Zhang, F., Wang, Y., Judelson, H. S., Chen, X., and Ma, W. 2013. Oomycete pathogens encode RNA silencing suppressors. *Nat. Gen*. 45:330-333.

Savory, E. A., Granke, L. L., Quesada-Ocampo, L. M., Varbanova, M., Hausbeck, M. K., and Day, B. 2010. The cucurbit downy mildew pathogen *Pseudoperonospora cubensis*. *Mol. Plant Pathol.* 12: 217-226.

Savory, E. A., Adhikari, B. N., Hamilton, J. P., Vaillancourt, B., Buell, C. R., and Day, B. 2012a. mRNA-Seq analysis of the *Pseudoperonospora cubensis* transcriptome during cucumber (*Cucumis sativus L.*) infection. *PLoS ONE* 7:e35796. Published online.

Savory, E. A., Zou, C., Adhikari, B. N., Hamilton, J. P., Buell, C. R., Shiu, S.-H., and Day, B. 2012b. Alternative splicing of a multi-drug transporter from *Pseudoperonospora cubensis* generates an RXLR effector protein that elicits a rapid cell death. *PLoS ONE* 7:e34701. Published online.

Seyednasrollah, F., Laiho, A., and Elo, L. L. 2013. Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief. Bioinform.* 16:59-70.

Shahid, S., and Axtell, M. J. 2013. Identification and annotation of small RNA genes using ShortStack. *Methods* 67:20-27.

Stassen, J. H. M. and Van den Ackerveken, G. 2011. How do oomycete effectors interfere with plant life? *Curr. Opin. Plant. Biol.* 14:407-414.

Taler, D. , Galperin, M., Benjamin, I., Cohen, Y., and Kenigsbuch, D. 2004. Plant eR Genes that encode photorespiratory enzymes confer resistance against disease. *Plant Cell* 16:172-184.

Tian, M., Win, J., Savory, E., Burkhardt, A., Held, M., Brandizzi, F., and Day, B. 2011. 454 genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Mol. Plant-Microbe Interact.* 24:543-553.

Trapnell, C., Pachter, L., and Salzberg, S. L. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25:1105-1111.

van Damme, M., Huibers, R. P., Elberse, J., and Van den Ackerveken, G. 2008. Arabidopsis DMR6 encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew. *Plant J.* 54:785-793.

Vetukuri, R. R., Åsman, A. K. M., Tellgren-Roth, C., Jahan, S. N., Reimegård, J., Fogelqvist, J., Savenkov, E., Söderbom, F., Avrova, A. O., Whisson, S. C., and Dixelius, C. 2012. Evidence for small RNAs homologous to effector-encoding genes and transposable elements in the oomycete *Phytophthora infestans*. *PLoS ONE* 7:e51399. Published online.

Visser, M., van der Walt, A. P., Maree, H. J., Rees, D. J. G., and Burger, J. T. 2014. Extending the sRNAome of apple by next-generation sequencing. *PLoS ONE* 9:e95782. Published onilne.

Wang, G., Ding, X., Yuan, M., Qui, D., Li, X., Xu, C., and Wang, S. 2006. Dual function of rice *OsDR8* gene in disease resistance and thiamine accumulation. *Plant Mol. Biol.* 60:437-449.

Wawrzynska, A., Rodibaugh, N. L., and Innes, R. W. 2010. Synergistic activation of defense responses in Arabidopsis by simultaneous loss of the GSL5 callose synthase and the EDR1 protein kinase. *Mol Plant Microbe Interact.* 23:578-584.

Weiberg, A., Wang, M., Bellinger, M., and Jin, H. 2014. Small RNAs: A new paradigm in plant-microbe interactions. *Annu. Rev. Phytopathol.* 52:495-516.

Weiberg, A., Wang, M., Lin, F. M., Zhao, H., and Zhang, Z. 2013. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* 342:118-123.

Wildermuth, M. C., Dewdney, J., Wu, G., and Ausubel, F.M. 2001. Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature* 414: 562-571.

Yang, Q., Reinhard, K., Schiltz, E., and Matern, U. 1997. Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA: anthranilate N-hydroxycinnamoyl/benzoyltransferase from elicited cell cultures of carnation, *Dianthus caryophyllus* L. *Plant Mol. Biol*. 35:777-789.

Yoshioka, Y., Sakata, Y., Sugiyama, M., and Fukino, N. 2014. Identification of quantitative trait loci for downy mildew resistance in cucumber (*Cucumis sativus* L.). *Euphytica* 198:265-276.

Zdobnov, E. M., and Apweiler, R. 2001. InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17:847-848.

Zhao, H.-C., Li, G.-J., and Wang, J.-B. 2005. The accumulation of phytoalexin in cucumber plant after stress. *Colloids and Surfaces B: Biointerfaces* 43:187-193.

Zhao, M., Cai, C., Zhai, J., Lin, F., Li, L., Shreve, J., Thimmapuram, J., Hughes, T. J., Meyers, B. C., and Ma, J. 2015. Coordination of microRNAs, phasiRNAs, and NB-LRR Genes in response to a plant pathogen: Insights from analyses of a set of soybean Rps gene near-isogenic lines. *Plant Genome* 8:1-13.

CHAPTER 4

Conclusions and future perspectives

Summary of dissertation

When I began my dissertation research in 2010, the popularity and availability of next-generation sequencing and bioinformatics were still in the early stages. The genome of cucumber (Huang et al. 2009) and the obligate downy mildew pathogen *Pseudoperonospora cubensis* (Tian et al. 2011) had recently been sequenced, and our lab was in the process of a collaborative project to sequence the transcriptome of the susceptible interaction between cucumber and *P. cubensis* (Savory et al. 2012a; Adhikari et al. 2012). Although I began my dissertation with the intent of using only molecular biology and microscopy-based techniques to answer questions in the field of oomycete-plant interactions, I soon realized that the field was quickly moving in the direction of "omics" biology, and I shifted my research focus as a result. During the past five years, advances in sequencing technology and bioinformatics have led to the curation of vast amounts of genomic and transcriptomic data that have collectively advanced the field of oomycete plant pathology by providing genomics resources and by allowing for changes in entire transcriptomes to be analyzed under specific conditions (Savory et al. 2012a; Adhikari et al. 2012; Chen et al. 2013; Lamour et al. 2012; Asai et al. 2014; Kunjeti et al. 2011). In addition, the sequencing of multiple plant-pathogenic oomycetes has allowed for comparative genomics to identify common characteristics of obligate oomycetes (Baxter et al. 2010) and to improve genome annotations (Stam et al. 2013).

Within the field of plant-oomycete interactions, the work of this dissertation focuses on the interaction between cucumber and *P. cubensis* from a mostly basic science perspective. Previous work largely focused on screening cucumber lines to find a genetic source of

downy mildew resistance for breeding (Call et al. 2012b; 2012a) since the pathogen re-emerged in the United States in 2004 (Holmes et al. 2006). While resistant cucumber lines were identified, several had poor fruit quality, and the inheritance patterns of downy mildew still remained unclear (Call et al. 2012a; 2012b; Zhang et al. 2013; Yoshioka et al. 2014). In parallel to breeding work done with cucumbers, the sequencing of the cucumber genome in 2009 greatly improved the genetic resources available to study cucumber and allowed for easier analysis of the transcriptome and identification of SNPs and other genetic variations (Huang et al. 2009). Previous work from our lab contributed to the genetic understanding of the interaction between cucumber and *P. cubensis* by sequencing the transcriptome of a susceptible interaction over a time course of downy mildew infection (Adhikari et al. 2012; Savory et al. 2012a). Through this work, several patterns of gene expression were identified in both *P. cubensis* and in cucumber; however, this work did not identify genes contributing to resistance, which was the focus of work in Chapter 3 of this dissertation. In addition, the sequenced transcriptome of *P. cubensis* led to the identification of the first identified alternatively spliced gene in *P. cubensis* which occurred when a gene annotated as a multi-drug transporter became an effector when the transcript was alternatively spliced (Savory et al. 2012b). This first identification of alternative splicing in *P. cubensis* prompted the work that became Chapter 2 of this dissertation, which investigated alternative splicing in *P. cubensis* on a transcriptome-wide scale.

In Chapter 2, we resequenced the *P. cubensis* transcriptome and used the resultant data to reannotate the genome and to identify alternatively spliced genes. The genome reannotation process improved the annotation of nearly half of the previously annotated

genes and identified nearly 4,000 new genes, some of which were experimentally validated. In addition, alternative splicing was studied on a transcriptome-wide scale and was found to be more prevalent than previously thought. Through our work described in Chapter 2, we found that 24% of the expressed genome had evidence for alternative splicing, which is in contrast to oomycete research that was completed with less deep sequencing but is consistent with more recent research using RNA-Seq (McGuire et al. 2008; Shen et al. 2011). In addition, the most frequent mechanism of alternative splicing found was intron retention, which is consistent with other oomycetes and plants (McGuire et al. 2008). We also found evidence for the usage of alternative 3' and 5' ends of exons, but no data supported the use of alternative exons, which is more common in animals. The mechanisms and frequency of alternative splicing was most similar to that found in plants, which is consistent with the placement of oomycetes closer to plants than to fungi on the evolutionary tree (Simpson and Roger 2004; Burki et al. 2012). This work contributed to the field of oomycete biology in that it provided evidence for the increased prevalence of alternative splicing among multiple types of genes, including those including proteins relevant to host infection like secreted proteins and putative effectors. Among the predicted alternatively spliced *P. cubensis* transcripts, one gene even showed evidence for alternative splicing during the sporangia to zoospore transition, which parallels other developmental studies in fern (Boothby et al. 2013).

In Chapter 3, the focus of my research shifted towards defining the host processes that mediate the plant-pathogen interaction, again utilizing next-generation sequencing methods to answer relevant questions about this pathosystem. Similar to previous work

(Adhikari et al. 2012), a time course of infection of cucumber with the MSU-1 isolate of *P. cubensis* was completed, yet in this work, the novelty lies in the inclusion, and subsequent analysis of gene expression in a resistant cucumber line, PI 197088. Through the comparative analyses between 2 complete biological replicates of a time course of infection between Vlaspik and PI 197088, I was able to identify groups of genes that were coordinately and oppositely regulated in a resistant and susceptible line. By using transcriptomic analyses guided by DESeq and weighted genome correlation network analysis (WGCNA), I found that the resistant plant had an earlier response to the pathogen. In addition, I found that several genes involved in signaling hormone responses, including some defense-associated hormones like salicylic acid (SA), were differentially regulated between the resistant and susceptible lines. Beyond hormone changes, I identified potential changes in metabolites within the plant that could be contributing to resistance. Notably, the levels of genes associated with thiamine synthesis, a B-vitamin, were strong induced in the resistant plant, which is significant given that high thiamine levels in plants have previously been associated with resistance (Wang et al., 2006, Ahn et al. 2007; Boubakri et al. 2012; Boubakri et al. 2013) Finally, through coordination with data from other plant-oomycete interactions, I identified differentially expressed transcription factors that were predicted to have a role in mediating resistance, including one transcription factor that is also found within a QTL that is involved in downy mildew resistance (Gao et al. 2014a; Yoshioka et al. 2014; Caldwell et al. 2011).

Beyond just sequencing the transcriptome, I sequenced the small RNAs in *P. cubensis* sporangia an in mock-inoculated and inoculated Vlaspik and PI 197088 leaves.

Bioinformatic analysis of these data revealed four newly predicted miRNA in *P. cubensis* and several previously discovered miRNA and newly predicted miRNA in cucumber. Furthermore, a handful of miRNA were differentially expressed over the time course or between plant lines, and the expression patterns of some of these miRNA oppositely correlated with the expression pattern of their predicted targets. This supports our hypothesis that predicted miRNA regulate the transcript levels of their predicted targets. In total, the work presented in Chapter 3 describes the transcriptome and small RNAome changes that occur in a resistant and susceptible interaction between cucumber and *P. cubensis* and provides a list of candidate genes that can be further analyzed.

In summary, the work of this dissertation contributes the field of plant-oomycete interactions by adding knowledge about the changes in the transcriptome and small RNAome that occur in both the plant (cucumber) and the pathogen (*P. cubensis*) throughout a time course of infection. Prior to this work, relatively little was known about the frequency and type of alternative splicing across the transcriptome of *P. cubensis*. Now, we know that alternative splicing occurs relatively frequently and that intron retention is the most commonly used mechanism. Next-generation sequencing data collected to analyze the breadth of alternatively splicing was also valuable in reannotating the *P. cubensis* genome and in predicting new genes. On the plant side, next-generation sequencing was also used to determine the transcriptional differences between a resistant and susceptible cucumber and to generate a list of candidate resistance-associated genes. Additional work identified novel miRNA in both the plant and the pathogen and predicted host targets for each miRNA. In conclusion, the work described in this dissertation utilized advances in

sequencing technology and bioinformatics to increase the knowledge about the cucumber-*P. cubensis* pathosystem and provided hypotheses regarding gene regulation and resistance that can be tested through the future work proposed below.

Proposed future work

As described above, progress has been made through the completion of this research to broaden our understanding of the *P. cubensis*-cucumber interaction. In this section, I will provide an overview of what I believe are critical areas of future research that need to be undertaken to continue to move this research forward. The future studies proposed below could serve to increase the quality of knowledge about this system and to validate some of the bioinformatically predicted functions of genes and small RNAs. For example, resequencing the *P. cubensis* genome will serve to improve the prediction of small RNAs and will also improve the overall gene annotation. Further bioinformatics work on the cucumber side could be done to identify single nucleotide polymorphisms between the resistant (PI 197088) and susceptible (Vlaspik) transcripts. Beyond "omics" work involving sequenced nucleotides, additional proteome and metabolome profiling is proposed to provide a stronger understanding of the physiological change that is prompted in the plant, as a change in transcript level is not always reflective of the protein or metabolite levels in the plant. Finally, candidate genes and miRNAs should be validated using both Arabidopsis and cucumber.

Resequencing of the *P. cubensis* genome

Since the initiation of *P. cubensis* genome sequencing by our lab in 2009 (Tian et al. 2011), genome sequencing technologies have greatly improved, with a concomitant decrease in cost (Buermans and Dunnen 2014). In the originally published genome sequence from Tian et al. 2011, only about 600,000 reads were assembled into 42,799 contigs with the average contig length only being 863 nucleotides. The genome coverage and assembly could be improved, as the sequenced reads are hypothesized to cover only about 14% of the total predicted genome size, and the average contig length is very short (Tian et al. 2011). Improving the genome is important because it could improve gene annotation and prediction and small RNA prediction (Martin et al. 2013). Notably, some types of sRNAs derived from oomycetes have been shown to be derived from transposable elements (Fahlgren et al. 2013). Recently, newer sequencing technology like SMRT Cells from Pacific Biosciences allow for better genome assembly because they have the ability to produce very long reads (over 14,000 base pairs). Future work to improve the *P. cubensis* genome assembly could make use of the SMRT cells and other sequencing advances, effectively enabling a deeper coverage of the genome with longer reads (Buermans and Dunnen 2014).

SNP analysis between Vlaspik and PI 197088

Further bioinformatics analyses can be done with the existing data to analyze sequenced transcripts for single nucleotide polymorphisms (SNPs) that exist between the resistant

and susceptible cucumber. Differences in transcript expression levels were thoroughly analyzed in Chapter 3; however, all of the reads were aligned to the well-sequenced and annotated reference genome of the Chinese Long 9930 cucumber (Huang et al. 2009). In addition to using the sequenced reads to measure changes in transcript levels, these reads could also be used to identify SNPs that might exist between each plant line and the reference genome and between Vlaspik and PI 197088. Examining SNPs between Vlaspik and PI 197088 is a relevant part of understanding the complete picture of host resistance, as transcript expression levels do not reflect the predicted functionality of a protein. For example, a SNP could lead to a change in an amino acid or a premature stop codon which may result in a protein with an altered function or a dysfunctional protein; this change might not be reflected in transcript/protein abundance. In support of this, mutations in some genes have already been shown to confer resistance to downy mildew pathogens, as well as *Fusarium* spp. (van Damme et al. 2008; Brewer et al. 2014). In addition, it is possible that some SNPs within a transcript may disrupt the binding site for a predicted miRNA. In plants, miRNA and targets typically tolerate very few mismatches in the 2-13 nucleotide position (Axtell 2013); thus, it is possible that transcript abundance could be affected by a SNP that causes a disruption in an miRNA target site. In order to identify SNPs, a program like SAMTools could be used (Li et al. 2009). Following SNP identification, a separate program called SIFT could be used to predict if a changed nucleotide would result in a change in gene function (Kumar et al. 2009).

Proteomics and metabolomics

Although the transcriptomics and small RNAome research presented in Chapter 3 provided valuable insight and a foundation for further investigation into the mechanisms of resistance in PI 197088, additional proteomic and metabolomic data would improve our understanding of the processes that are occurring in the resistant plant (e.g., Feussner and Polle 2015). Proteomics and metabolomics are important because changes in the transcriptome might not be reflected by changes in the proteome or metabolome, as some transcripts might not be translated or some proteins might be quickly degraded and or might not produce the expected metabolites. For example, in a study involving a compatible and incompatible reaction between *P. infestans* and potato, about half of the differentially abundant apoplastic proteins had a corresponding change in transcript level, which provides evidence that factors other than transcript levels can have a role in affecting the level of protein at any given time (Ali et al. 2014). In addition to quantifying the level of protein, proteomics can also shed light on the functionality of a protein by identifying post-translational modifications to a protein such as palmitoylation, phosphorylation, glycosylation, and acetylation (Hemsley et al. 2012; Melo-Braga et al. 2012). Another advantage of proteomics is the ability to identify and quantify proteins from unique spaces within the plant, such as the apoplast, which is an important interaction zone between plant and pathogens where proteins like pathogen effectors and plant defense-associated enzymes reside (Delaunois et al. 2014). In the *P. cubensis*-cucumber system, proteomics studies of the apoplastic space could be particularly useful to identify apoplastic effectors from *P. cubensis.*

Beyond using proteomics to better define mechanisms of resistance, metabolomics could also help us to better understand the complexity of the *P. cubensis*-cucumber interaction. For example, metabolomics-based studies have previously been used to discover resistance mechanisms against the fungal pathogen *F. graminearum* in wheat, showing that the accumulation of resistance-related metabolites in the phenylpropanoid pathway was key to resistance in a previously poorly understood quantitatively inherited resistance (Gunnaiah and Kushalappa 2014). In addition, a separate metabolomics-based study in Arabidopsis showed that phenylpropanoids confer resistance to *V. longisporum* (König et al. 2014). In cucumber, phytoalexins have been shown to be important in defense responses to other pathogens (Hammerschmidt 1999), and data in Chapter 3 suggests that increased phytoalexin synthesis might have a role in conferring resistance in PI 197088. Finally, similar methods could be used to evaluate plant hormones on a large-scale (Kojima et al. 2009), which would elucidate the differences in salicylic acid, jasmonic acid, ethylene, and auxin that were highlighted in different stages during the resistant and susceptible *P. cubensis*-cucumber interaction in Chapter 3. Together with proteomics and transcriptomics, metabolomics would offer valuable insight into the complex processes regulating resistance to *P. cubensis* in cucumber, and this combination of "omics" technologies has already proven valuable in Brassicaceae-fungal interactions (Floerl et al. 2012). Additionally, a different study in wheat coupled transcriptome and metabolite profiling to study the infection mechanisms and plant response to the hemibiotrophic fungal pathogen, *Zymoseptoria tritici* (Rudd et al. 2015).

Validation of resistance gene candidates using SALK lines

From the reduced candidate list created through the combined "omics" approach described above, putative resistance-associated genes could be evaluated using one or more of the methods described below. For genes that were not previously identified to have a role in downy mildew resistance in Arabidopsis as described in the previously cited literature (Chapters 1 and 3), we will order mutant SALK lines for the closest Arabidopsis homolog of the predicted resistance-associated candidate in cucumber. Once homozygosity of the mutant lines has been established, we will evaluate the mutants for susceptibility or resistance to the Arabidopsis downy mildew, *H. arabidopsidis.* Although this system is not a perfect recapitulation of the cucumber – downy mildew system, it does incorporate key aspects of the cucumber – *P. cubensis* pathosystem in that *H. arabidopsidis* is an obligate oomycete pathogen of leaves and has infection strategies and nutritional needs similar to that of *P. cubensis.* In addition to screening the SALK lines for resistance and susceptibility to *P. cubensis,* we will include the Arabidopsis *pad4* mutant (Col-0 background) as a susceptible control, and WT Col-0 as a resistant control. By using these mutant lines, we hope to validate genes that are predicted to contribute to resistance in PI 197088 when highly expressed, and thus when silenced in a similar plant-oomycete pathosystem would lead to susceptibility of the plant. Alternatively, resistance gene candidates that are lowly expressed in the resistant line but highly expressed in the susceptible line may lead to resistance to *H. arabidopsidis* when their homologs are mutated or knocked out in Arabidopsis. Although this data has limitations, in that the species used for validation are different than the species being researched, the evidence provided through this study can

be used to corroborate or provide negative evidence toward putative resistance-associated genes. In addition, given the abundance of genetic tools and resources available in Arabidopsis, plant lines with genes supporting the hypotheses developed in cucumber could be used in future studies to determine a mechanism of action for the gene involved in mediating resistance. For example, interaction studies could be performed in Arabidopsis using co-immunoprecipitation assays to identify host proteins that interact with the putative resistant gene product (Feys et al. 2001; Chinchilla et al. 2007). Additionally, known resistance-associated genes, such as *PR* genes, could be measured through qPCR and known MAPK pathways involved in resistance could be quantified (Yi et al. 2015; Meng and Zhang 2013; Seyfferth and Tsuda 2014; Oide et al. 2013)

Validation of resistance gene candidates using viral induced gene silencing

In addition to validating genes using Arabidopsis, mutants of candidate genes could be made using one or both of the following methods in cucumber. First, viral induced gene silencing, or VIGS, could be used to knock out putative resistance (*R*) genes in either cucumber plant line (Vlaspik or PI 197088) depending on if resistance is predicted to be associated with up- or down-regulated of a particular gene in PI 197088 compared to Vlaspik. VIGS has previously been used in cucumber with the apple latent spherical virus (Igarashi et al. 2009), and while it has limitations of being labor and time intensive and expression can be inconsistent, it is a valuable method because it can be used for cucumber (Senthil-Kumar and Mysore 2011). In recent years, VIGS has successfully been used in several crops to validate gene function in regard to abiotic and biotic stress responses

(Ramegowda et al. 2014; Wang et al. 2015; Hou et al. 2014; Xu et al. 2014; Gao et al. 2014b; Wang et al. 2012). For example, the involvement of a wheat polygalacturonase-inhibiting protein in mediating resistance to *Fusarium graminearum* was supported when a plant expressing a VIGS construct to knock down expression of the gene was more susceptible than unaltered control (Hou et al. 2014). On the other hand, VIGS has been used to identify a negative regulator of plant immunity by knocking down a WRKY transcription factor in pepper that was determined to be a negative regulator of plant immunity (Wang et al. 2012). Similar strategies could be used to investigate genes regulating susceptibility and resistance in cucumber by knocking down negative regulators of resistance with the expectation of increased resistance or by knocking down positive regulators of resistance with the expectation of increased susceptibility.

Validation of candidate genes using amiRNA

In addition to the approaches described above, an alternative, or additional, method that can be used in cucumber is using artificial miRNA (amiRNA) constructs designed to target specific genes of interest. Artificial miRNA constructs can be designed to target a specific candidate resistance gene and knock down expression of that gene in a transformed cucumber plant. Specifically, cucumber amiRNA could be designed using the WMD3 – Web MicroRNA Designer (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi). These amiRNA could then be inserted into the pCAMBIA1300-STTM vector and transformed into the *Agrobacterium rhizogenes* strain K599; this vector and bacteria have previously been used for hairy root transformation of cucumber in our lab by Dr. Patricia Santos (unpublished).

Transformation of cucumber with *A. rhizogenes* was first shown in 1986 by Trulson et al. (Trulson et al. 1986). To generate composite transgenic cucumber lines, the root systems of 10-day-old cucumber plantlets will be removed, and the shoots placed in small cubes of rock wool soaked in a solution of *A. rhizogenes* strain K599 expressing the pCAMBIA1300-STTM and amiRNA constructs. After incubation for two days, the rock wool cubes containing the plantlets will be allowed to dry until plants wilt (5 to 6 hours, maximum). Plantlets will be watered and kept moist in covered trays at room temperature until root development growing from the teratoma is observed (approximately 35-45 days). After 10-15 days, roots growing outside of the teratoma will be removed and the composite plants moved to a hydroponic system for further growth and development. With this protocol, ~100% of the infected plantlets form transgenic hairy roots. The transformation of the cucumber roots will be confirmed using PCR to confirm the presence of the transgene and the expression of the miRNA will be confirmed using northern blot and or qPCR.

Once transgenic roots have been established, grafting can be used to determine if the amiRNA is able to knock down gene expression and confer resistance or susceptibility to the scion. The grafting approach has strong potential to offer a faster way to knockdown a gene in a whole plant, as cucumber is amenable to grafting and it has been shown that some miRNA are capable of traveling long distances (Pant et al. 2008). Some work in cucumber has found that some miRNA are responsive to grafting; however, the data did not provide conclusive evidence on the ability of specific miRNA to travel between the rootstock and the scion (Li et al. 2013). The proposed work with the transgenic rootstock and non-transgenic scion could provide useful data regarding the ability of miRNA to travel

within cucumber and could be a relatively fast way to test gene function using amiRNA. If the amiRNA are able to successfully travel and silence target genes in the leaves, then resistance or susceptibility conferred by knocking down candidate resistance genes could be evaluated through downy mildew inoculation of the grafted plants.

Validation of genes using genome editing

Over the past few years, new genome editing techniques, including zinc finger nucleases (ZFN), transcription activator-like effector nucleases (TALEN) and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 have all been developed and have been used to edit plant genomes (Osakabe and Osakabe 2015). These genome editing methods are similar in that they all work by creating nicks or break in the genome and use DNA repair mechanisms to induce mutations or insertions in the genome; they differ in the targeting of the DNA breakage and thus in their specificity (Osakabe and Osakabe 2015). The TALEN method is newer and more specific than the ZFN in that specific combinations of amino acids have been identified to target individual nucleotides (Cermak et al. 2011). TALEN been used in several plant species including Arabidopsis as well as in multiple crops (Li et al. 2014; Wendt et al. 2013; Christian et al. 2013; Curtin et al. 2012). Most recently, the CRIPSR/Cas9 method has emerged as a popular method with high potential to direct very specific mutations in the genome with its RNA-based targeting method that allows for precise genome alterations (Bortesi and Fischer 2015). Even within the past couple of years, the popularity of the CRISPR/Cas9 system has grown in plant genome editing and has been used in Arabidopsis, rice, tobacco, sorghum, and other plants (Miao et al. 2013;

Jiang et al. 2013; Fauser et al. 2014; Xie and Yang 2014). These rapidly improving technologies in plant genome editing provide promising prospects for the ability to edit the cucumber genome to test the function of candidate resistance genes and perhaps to make stable transgenic cucumbers that are resistant to downy mildew. Currently, the limiting step in using many of these modern genome altering techniques is the difficulty in using *Agrobacterium tumefaciens* to transform cucumber, but recent research has improved the efficiency of this process (Nanasato et al. 2012). Future work can focus on further improvement of cucumber transformation with the goal of using quickly developing genome editing technology to make stable transgenic cucumbers that are resistant to downy mildew.

Validation of miRNA expression and targets

The presence and relative abundance of cucumber and *P. cubensis* miRNAs during infection of resistant and susceptible cultivars can be confirmed by northern blot and or qPCR (Mao et al. 2012; Martínez et al. 2011). In Chapter 4, the some of the predicted targets of miRNA candidates were supported with oppositely coordinated patterns of target gene expression. A similar approach of using correlating gene expression patterns to miRNA expression has been previously used to provide support for target predictions (Li et al. 2013). In addition, an further approaches can be taken to validate the targets of predicted miRNA by sequencing the degradome, which is a high-throughput approach that has previously been successful in cucumber (Mao et al. 2012). To determine if miRNA reduces the protein of targets, Western blot analyses could be performed (Sun et al. 2014). Finally, another

approach to confirm the target of a miRNA is to co-express a fluorescently labeled target and the miRNA transiently in tobacco. This approach would provide evidence that the target is correctly predicted if the level of fluorescence is reduced in a plant expressing the miRNA compared to a plant not expressing the miRNA. This method has been successfully used to identify MAPK as a target for *Botrytis cinerea* small RNA (Weiberg et al. 2013).

Functional analyses of miRNA

The biological function of miRNAs can be evaluated by transforming constructs designed to knock down or overexpress a specific miRNA using the proposed plant transformation methods described above. Previous research in the oomycete *Phytophthora sojae* has shown that using the short tandem target mimic (STTM) technique is effective in functionally knocking down an miRNA by expressing binding sites to which the miRNA of interest can bind but not degrade (Wong et al. 2014). This method could be used in cucumber using the root transformation with *A. rhizogenes* to knock down the expression of an miRNA. In contrast, miRNAs could be overexpressed in a plant through the use of a construct with a strong promoter. In either the case of a knock down or overexpression of a miRNA, the level of miRNA and the target transcript could be evaluated using either Northern blot or qPCR. Finally, the biological function of each miRNA in regard to downy mildew resistance could be tested by inoculating each plant with either the knocked down or overexpressed miRNA with *P. cubensis* and measuring the level of infection. In this way, it might be determined which, if any, miRNA has a critical role in regulating the transcripts or protein levels that are contributing to downy mildew resistance in cucumber.

Future perspectives

In the field of oomycete-plant interactions, research advances have led to numerous debates and have opened up doors for future research that could reach the ultimate goal of improving plant health. In terms of the pathogen, broad areas of current research focus on the mechanisms by which the pathogen interacts with the host on a molecular level, with a particular emphasis on effectors and nutrient exchange in the haustoria. In particular, the entry of effectors has been recently debated topic in oomycete research and remains to be fully agreed upon (Dou et al. 2008, Kale et al. 2010, Ellis and Dodds 2011, Kale 2011, Yaeno et al. 2011). This topic is of particular relevance because if the entry of effectors into the host were understood, the goal of blocking effector entry would be closer to being achieved. Thus, future work should focus on resolving this conflict and reaching a consensus about the mode of effector entry. In addition, the functional role of a few effectors is known, but many effectors still have an unknown biological function, which could be a topic of future work in the field. Finally, some effectors are recognized by known R proteins, but more *R* genes remain to be discovered in yet un-researched plant lines. In some cases, effector-directed breeding could be used to test multiple plant lines against a set of conserved effectors for which *R* genes are more likely to remain durable. In summary, future work regarding effectors and *R* genes should focus on questions of function and entry while using the knowledge gained to direct the development of oomycete-resistant plants.

Another question that arises around the interface of the plant and pathogen is better understanding the process by which nutrients and other small molecules, including small RNAs, are able to travel between the host and the pathogen. This area of research is of particular interest to obligate oomycetes like the downy mildews, as they must obtain essential nutrients from the plant in order to survive. Future work should focus on better understanding the specific nutritional needs of obligate microbes and the mechanisms by which they acquire nutrition from the plant. In addition, the movement of regulatory small RNAs traveling from pathogen to host or vice versa could open up an interesting additional layer of regulation in the interaction between host and pathogen. Thus far, it has been found that small RNA can travel from fungus to plant (Weiberg et al. 2013) and that small RNAs are important in regulating the host response during an oomycete infection (Qiao et al. 2013; Wong et al. 2014). Future research should focus on better understanding how small RNAs are involved in regulating the interaction between the plant and the pathogen.

Finally, the ultimate goal of a large part of research in plant pathology is to improve crop health and increase food security. Currently, a large part of the more basic research in plant pathology is focused on generating and analyzing large data sets, like the ones described in this dissertation, which are helpful in detecting trends and forming hypotheses. However, future work will need to focus on combining large data sets from different types of experiments (i.e., proteomics and transcriptomics) and on using systems biology to gain a more complete understanding of the plant's response to a particular pathogen. In order to transition from the lab to the field, we will need to begin to apply our research and focus more on researching crop pathosystems. For example, breeders could use resistance genes

that have been discovered in controlled studies in the lab to specifically select for genes or traits in the field. In addition, previously discussed research has opened up new and exciting opportunities in genome editing through TALEN and CRISPR/Cas9, and these approaches could be translated into crops. A significant goal for future research will be to translate the data generated by quickly developing "omics" research into validated gene candidates and ultimately into healthier crops.

REFERENCES

REFERENCES

Adhikari, B. N., Savory, E. A., Vaillancourt, B., Childs, K. L., Hamilton, J. P., Day, B., and Buell, C. R. 2012. Expression profiling of *Cucumis sativus* in response to infection by *Pseudoperonospora cubensis*. *PLoS ONE* 7:e34954. Published online.

Ahn, I-P., Kim, S., Lee, Y-H., and Suh, S-C. 2007. Vitamin B_1-induced priming is dependent on hydrogen peroxide and the *NPR1* gene in Arabidopsis. *Plant Phys.* 143:838-848.

Ali, A., Alexandersson, E., Sandin, M., Resj, S., Lenman, M., Hedley, P., Levander, F., and Andreasson, E. 2014. Quantitative proteomics and transcriptomics of potato in response to *Phytophthora infestans* incompatible and incompatible interactions. *BMC Genomics* 15:497. Published online.

Asai, S., Rallapalli, G., Piquerez, S. J. M., Caillaud, M.-C., Furzer, O. J., Ishaque, N., Wirthmueller, L., Fabro, G., Shirasu, K., and Jones, J. D. G. 2014. Expression profiling during Arabidopsis/downy mildew interaction reveals a highly-expressed effector that attenuates responses to salicylic acid. *PLoS Pathog*. 10:e1004443. Published online.

Axtell, M. J. 2013. Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol*. 64:137–159.

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

Boothby, T. C., Zipper, R. S., van der Weele, C. M., and Wolniak, S. M. 2013. Removal of retained introns regulates translation in the rapidly developing gametophyte of *Marsilea vestita*. *Dev. Cell* 24:517–529.

Bortesi, L., and Fischer, R. 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotech. Adv.* 33:41–52.

Boubakri, H., Wahab, M.A., Chong, J., Bertsch, C., Mliki, A., and Soustre-Gacougnolle, I. 2012. Thiamine induced resistance to *Plasmopara viticola* in grapevine and elicited host-defense responses, including HR like-cell death. *Plant Phys. Biochem.* 57:120-133.

Boubakri, H., Poutaraud, A., Wahab, M.A., Clayeux, C., Baltenweck-Guyot, R., Steyer, D., Marcic, C., Mliki, A., Soustre-Gacougnolle, I. 2013. Thiamine modulates metabolism of the phenylpropanoid pathway leading to enhanced resistance to *Plasmopara viticola* in grapevine. *BMC Plant Biol.* 13:31. Published online.

Brewer, H. C., Hawkins, N. D., and Hammond-Kosack, K. E. 2014. Mutations in the Arabidopsis homoserine kinase gene DMR1 confer enhanced resistance to *Fusarium culmorum* and *F. graminearum*. *BMC Plant Biol.* 14:317. Published online.

Buermans, H. P. J., and Dunnen, den, J. T. 2014. Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta* 1842:1932-1941.

Burki, F., Okamoto, N., Pombert, J. F., and Keeling, P. J. 2012. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *Proc. Biol. Sci.* 279:2246–2254.

Caldwell, D., Chan, E., de Vries, J., Joobeur, T., King, J., Reina, A., and Shetty, N. V. 2011. Methods and compositions for identifying downy mildew resistance cucumber plants. US Patent 20110123609 A1. Date issued: May 26.

Call, A. D., Criswell, A. D., Wehner, T. C., Ando, K., and Grumet, R. 2012a. Resistance of cucumber cultivars to a new strain of Cucurbit downy mildew. *HortScience* 47:171-178.

Call, A. D., Criswell, A. D., Wehner, T. C., Klosinska, U., and Kozik, E. U. 2012b. Screening cucumber for resistance to downy mildew caused by (Berk. and Curt.) Rostov. *Crop Sci.* 52:577-592.

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J., and Voytas, D. F. 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucl. Acids Res*. 39:e82–e82. Published online.

Chen, X.-R., Xing, Y.-P., Li, Y.-P., Tong, Y.-H., and Xu, J.-Y. 2013. RNA-Seq reveals infection-related gene expression changes in *Phytophthora capsici. PLoS ONE* 8:e74588. Published online.

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

Christian, M., Qi, Y., Zhang, Y., and Voytas, D. F. 2013. Targeted mutagenesis of *Arabidopsis thaliana* using engineered TAL effector nucleases. *G3-Gene Genom. Genet.* 3:1697-1705.

Curtin, S. J., Voytas, D. F., and Stupar, R. M. 2012. Genome engineering of crops with designer nucleases. *Plant Gen*. 5:42-50.

Delaunois, B., Jeandet, P., Clément, C., Baillieul, F., Dorey, S., and Cordelier, S. 2014. Uncovering plant-pathogen crosstalk through apoplastic proteomic studies. *Front. Plant Sci.* 5:1-18. Published online.

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

Ellis, J. G. and Dodds, P. N. Showdown at the RXLR motif: Serious differences of opinion in how effector proteins from filamentous eukaryotic pathogens enter plant cells. *Proc. Nat. Acad. Sci.* 108:14381-14382.

Fahlgren, N., Bollmann, S. R., Kasschau, K. D., Cuperus, J. T., Press, C. M., Sullivan, C. M., Chapman, E. J., Hoyer, J. S., Gilbert, K. B., Grunwald, N. J., and Carrington, J. C. 2013. *Phytophthora* have distinct endogenous small RNA populations that include short interfering and microRNAs. *PLoS ONE* 8:e77181. Published online.

Fauser, F., Schiml, S., and Puchta, H. 2014. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J.* 79:348–359.

Feussner, I., and Polle, A. 2015. What the transcriptome does not tell - proteomics and metabolomics are closer to the plants' patho-phenotype. *Curr. Opin. Plant Biol.* 26:26–31.

Feys, B. J., Moisan, L. J., and Newman, M. A. 2001. Direct interaction between the Arabidopsis disease resistance signaling proteins, EDS1 and PAD4. *The EMBO J.* 20:5400-5411.

Floerl, S., Majcherczyk, A., Possienke, M., Feussner, K., Tappe, H., Gatz, C., Feussner, I., Kues, U., and Polle, A. 2012. *Verticillium longisporum* infection affects the leaf apoplastic proteome, metabolome, and cell wall properties in *Arabidopsis thaliana. PLoS ONE* 7:e31435. Published online.

Gao, D., Appiano, M., Huibers, R. P., Chen, X., Loonen, A. E. H. M., Visser, R. G. F., Wolters, A. M. A., and Bai, Y. 2014a. Activation tagging of ATHB13 in *Arabidopsis thaliana* confers broad-spectrum disease resistance. *Plant Mol. Biol*. 86:641–653.

Gao, D., Huibers, R. P., and Loonen, A. 2014b. Down-regulation of *acetolactate synthase* compromises *Ol-1*-mediated resistance to powdery mildew in tomato. *BMC Plant Biol.* 14:32. Published online.

Gunnaiah, R., and Kushalappa, A. C. 2014. Metabolomics deciphers the host resistance mechanisms in wheat culitvar Sumai-3, against trichothecene producing and non-producing isolates of *Fusarium graminearum. Plant Phys. Biochem*. 83:40–50.

Hammerschmidt, R. 1999. Phytoalexins: What have we learned after 60 years? *Annu. Rev. Phytopathol.* 37:285–306.

Hemsley, P. A., Weimar, T., Lilley, K. S., Dupree, P., and Grierson, C. S. 2012. A proteomic approach identifies many novel palmitoylated proteins in Arabidopsis. *New Phytol*. 197:805–814.

Holmes, G. J., Wehner, T. C., and Thornton, A. 2006. An old enemy re-emerges. *Dis. Manag.* 2:14–15.

Hou, W., Mu, J., Li, A., Wang, H., and Kong, L. 2014. Identification of a wheat polygalacturonase-inhibiting protein involved in *Fusarium* head blight resistance. *Eur. J. Plant Pathol.* 141:731–745.

Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., Lucas, W. J., Wang, X., Xie, B., Ni, P., Ren, Y., Zhu, H., Li, J., Lin, K., Jin, W., Fei, Z., Li, G., Staub, J., Kilian, A., van der Vossen, E. A. G., Wu, Y., Guo, J., He, J., Jia, Z., Ren, Y., Tian, G., Lu, Y., Ruan, J., Qian, W., Wang, M., Huang, Q., Li, B., Xuan, Z., Cao, J., Asan, Wu, Z., Zhang, J., Cai, Q., Bai, Y., Zhao, B., Han, Y., Li, Y., Li, X., Wang, S., Shi, Q., Liu, S., Cho, W. K., Kim, J.-Y., Xu, Y., Heller-Uszynska, K., Miao, H., Cheng, Z., Zhang, S., Wu, J., Yang, Y., Kang, H., Li, M., Liang, H., Ren, X., Shi, Z., Wen, M., Jian, M., Yang, H., Zhang, G., Yang, Z., Chen, R., Liu, S., Li, J., Ma, L., Liu, H., Zhou, Y., Zhao, J., Fang, X., Li, G., Fang, L., Li, Y., Liu, D., Zheng, H., Zhang, Y., Qin, N., Li, Z., Yang, G., Yang, S., Bolund, L., Kristiansen, K., Zheng, H., Li, S., Zhang, X., Yang, H., Wang, J., Sun, R., Zhang, B., Jiang, S., Wang, J., Du, Y., and Li, S. 2009. The genome of the cucumber, *Cucumis sativus,* L. *Nature Gen*. 41:1275–1281.

Igarashi, A., Yamagata, K., Sugai, T., Takahashi, Y., Sugawara, E., Tamura, A., Yaegashi, H., Yamagishi, N., Takahashi, T., Isogai, M., Takahashi, H., and Yoshikawa, N. 2009. Apple latent spherical virus vectors for reliable and effective virus-induced gene silencing among a broad range of plants including tobacco, tomato, *Arabidopsis thaliana*, cucurbits, and legumes. *Virology* 386:407–416.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., and Weeks, D. P. 2013. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucl. Acids Res*. 41:e188–e188. Published online.

Kale, S. D., Gu, B., Capelluto, D. G. S., Dou, D., Feldman, E., Rumore, A., Arredondo, F. D., Hanlon, R., Fudal, I., Rouxel, T., Lawrence, C. D., Shan, W., and Tyler, B. M. 2010. External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell* 14:284-295.

Kale , S. D. 2011. Oomycete and fungal effector entry, a microbial Trojan horse. *New Phytol.* 193:874-881.

Kojima, M., Kamada-Nobusada, T., Komatsu, H., Takei, K., Kuroha, T., Mizutani, M., Ashikari, M., Ueguchi-Tanaka, M., Matsuoka, M., Suzuki, K., and Sakakibara, H. 2009. Highly sensitive and high-throughput analysis of plant hormones using MS-probe modification and liquid chromatography-tandem mass spectrometry: An application for hormone profiling in *Oryza sativa*. *Plant Cell Physiol*. 50:1201–1214.

König, S., Feussner, K., Kaever, A., Landesfeind, M., Thurow, C., Karlovsky, P., Gatz, C., Polle, A., and Feussner, I. 2014. Soluble phenylpropanoids are involved in the defense response of Arabidopsis against *Verticillium longisporum*. *New Phytol*. 202:823–837.

Kumar, P., Henikoff, S., and Ng, P. C. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4:1073–1081.

Kunjeti, S. G., Evans, T. A., Marsh, A. G., Gregory, N. F., Kunjeti, S., Meyers, B. C., Kalavacharla, V. S., and Donofrio, N. M. 2011. RNA-Seq reveals infection-related global gene changes in *Phytophthora phaseoli*, the causal agent of lima bean downy mildew. *Mol. Plant Pathol*. 13:454–466.

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., Miller, N. A., Rice, B. J., Raffaele, S., Cano, L. M., Bharti, A. K., Donahoo, R. S., Finley, S., Huitema, E., Hulvey, J., Platt, D., Salamov, A., Savidor, A., Sharma, R., Stam, R., Storey, D., Thines, M., Win, J., Haas, B. J., Dinwiddie, D. L., Jenkins, J., Knight, J. R., Affourtit, J. P., Han, C. S., Chertkov, O., Lindquist, E. A., Detter, C., Grigoriev, I. V., Kamoun, S., and Kingsmore, S. F. 2012. Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. *Mol. Plant Microbe Interact.* 25:1350–1360.

Li, C., Li, Y., Bai, L., Zhang, T., He, C., Yan, Y., and Yu, X. 2013. Grafting-responsive miRNAs in cucumber and pumpkin seedlings identified by high-throughput sequencing at whole genome level. *Physiol. Plantarum* 151:406–422.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.

Li, T., Liu, B., Chen, C. Y., and Yang, B. 2014. TALEN utilization in rice genome modifications. *Methods* 69:9–16.

Mao, W., Li, Z., Xia, X., Li, Y., and Yu, J. 2012. A combined approach of high-throughput sequencing and degradome analysis reveals tissue specific expression of microRNAs and their targets in cucumber. *PLoS ONE* 7:e33040. Published online.

Martin, L. B. B., Fei, Z., Giovannoni, J. J., and Rose, J. K. C. 2013. Catalyzing plant science research with RNA-Seq. *Front. Plant. Sci.* 4:1-10.

Martínez, G., Forment, J., Llave, C., Pallás, V., and Gómez, G. 2011. High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs. *PLoS ONE* 6:e19523. Published online.

McGuire, A. M., Pearson, M. D., Neafsey, D. E., and Galagan, J. E. 2008. Cross-kingdom patterns of alternative splicing and splice recognition. *Gen. Biol.* 9:R50. Published online.

Melo-Braga, M. N., Verano-Braga, T., León, I. R., Antonacci, D., Nogueira, F. C. S., Thelen, J. J., Larsen, M. R., and Palmisano, G. 2012. Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (*Vitis vinifera*) mesocarp and exocarp owing to *Lobesia botran*a infection. *Mol. Cell Proteomics* 11:945–956.

Meng, X., and Zhang, S. 2013. MAPK cascades in plant disease resistance signaling. *Annu. Rev. Phytopathol.* 51:245–266.

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J., Gu, H., and Qu, L.-J. 2013. Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* 23:1233–1236.

Nanasato, Y., Konagaya, K.-I., Okuzaki, A., Tsuda, M., and Tabei, Y. 2012. Improvement of *Agrobacterium*-mediated transformation of cucumber (*Cucumis sativus* L.) by combination of vacuum infiltration and co-cultivation on filter paper wicks. *Plant Biotechnol Rep.* 7:267–276.

Oide, S., Bejai, S., Staal, J., Guan, N., Kaliff, M., and Dixelius, C. 2013. A novel role of PR2 in abscisic acid (ABA) mediated, pathogen-induced callose deposition in *Arabidopsis thaliana. New Phytol*. 200:1187–1199.

Osakabe, Y., and Osakabe, K. 2015. Genome editing with engineered nucleases in plants. *Plant and Cell Physiol*. 56:389–400.

Pant, B. D., Buhtz, A., Kehr, J., and Scheible, W. R. 2008. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant J*. 53:731–738.

Qiao, Y., Liu, L., Xiong, Q., Flores, C., Wong, J., Shi, J., Wang, X., Liu, X., Xiang, Q., Jiang, S., Zhang, F., Wang, Y., Judelson, H. S., Chen, X., and Ma, W. 2013. Oomycete pathogens encode RNA silencing suppressors. *Nature Gen.* 45:330-333.

Ramegowda, V., Mysore, K. S., and Senthil-Kumar, M. 2014. Virus-induced gene silencing is a versatile tool for unraveling the functional relevance of multiple abiotic-stress-responsive genes in crop plants. *Front. Plant. Sci.* 5:1-12.

Rudd, J. J., Kanyuka, K., Hassani-Pak, K., Derbyshire, M., Andongabo, A., Devonshire, J., Lysenko, A., Saqi, M., Desai, N. M., Powers, S. J., Hooper, J., Ambroso, L., Bharti, A., Farmer, A., Hammond-Kosack, K. E., Dietrich, R. A., and Courbot, M. 2015. Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritician* wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. *Plant Physiol.* 167:1158–1185.

Savory, E. A., Adhikari, B. N., Hamilton, J. P., Vaillancourt, B., Buell, C. R., and Day, B. 2012a. mRNA-Seq analysis of the *Pseudoperonospora cubensis* transcriptome during cucumber (*Cucumis sativus* L.) infection. *PLoS ONE* 7:e35796. Published online.

Savory, E. A., Zou, C., Adhikari, B. N., Hamilton, J. P., Buell, C. R., Shiu, S.-H., and Day, B. 2012b. Alternative splicing of a multi-drug transporter from *Pseudoperonospora cubensis* generates an RXLR effector protein that elicits a rapid cell death. *PLoS ONE* 7:e34701. Published online.

Senthil-Kumar, M. and Mysore, K. S. 2011. New dimensions for VIGS in plant functional genomics. *Trends Plant Sci*. 16:656–665.

Seyfferth, C. and Tsuda, K. 2014. Salicylic acid signal transduction: the initiation of biosynthesis, perception and transcriptional reprogramming. *Front Plant Sci.* 5:1-10.

Shen, D., Ye, W., Dong, S., Wang, Y., and Dou, D. 2011. Characterization of intronic structures and alternative splicing in *Phytophthora sojae* by comparative analysis of expressed sequence tags and genomic sequences. *Can. J. Microbiol.* 57:84–90.

Simpson, A. G. B., and Roger, A. J. 2004. The real "kingdoms" of eukaryotes. *Curr. Biol.* 14:R693–R696.

Stam, R., Jupe, J., Howden, A. J. M., Morris, J. A., Boevink, P. C., Hedley, P. E., and Huitema, E. 2013. Identification and characterisation CRN effectors in *Phytophthora capsici* shows modularity and functional diversity. *PLoS ONE* 8:e59517. Published online.

Sun, X., Zhang, Y., Zhu, X., Korir, N. K., Tao, R., Wang, C., and Fang, J. 2014. Advances in identification and validation of plant microRNAs and their target genes. *Physiol. Plantarum* 152:203–218.

Tian, M., Win, J., Savory, E., Burkhardt, A., Held, M., Brandizzi, F., and Day, B. 2011. 454 genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. *Mol. Plant Microbe Interact*. 24:543–553.

Trulson, A. J., Simpson, R. B., and Shahin, E. A. 1986. Transformation of cucumber (*Cucumis sativus* L.) plants with *Agrobacterium rhizogenes*. *Theor. Appl. Genet.* 73:11–15.

van Damme, M., Huibers, R. P., Elberse, J., and Van den Ackerveken, G. 2008. Arabidopsis DMR6 encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew. *Plant J.* 54:785–793.

Wang, G., Ding, X., Yuan, M., Qui, D., Li, X., Xu, C., and Wang, S. 2006. Dual function of rice *OsDR8* gene in disease resistance and thiamine accumulation. *Plant Mol. Biol.* 60:437-449.

Wang, F., Lin, R., Feng, J., Chen, W., Qiu, D., and Xu, S. 2015. TaNAC1 acts as a negative regulator of stripe rust resistance in wheat, enhances susceptibility to *Pseudomonas syringae*, and promotes lateral root development in transgenic *Arabidopsis thaliana*. *Front. Plant Sci.* 6:1–17.

Wang, Y., Dang, F., Liu, Z., Wang, X., Eulgem, T., Lai, Y., Yu, L., She, J., Shi, Y., Lin, J., Chen, C., Guan, D., Qiu, A., and He, S. 2012. CaWRKY58, encoding a group I WRKY transcription factor of *Capsicum annuum*, negatively regulates resistance to *Ralstonia solanacearum* infection. *Mol. Plant Pathol.* 14:131–144.

Weiberg, A., Wang, M., Lin, F. M., Zhao, H., and Zhang, Z. 2013. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* 342:118-122.

Wendt, T., Holm, P. B., Starker, C. G., Christian, M., Voytas, D. F., Brinch-Pedersen, H., and Holme, I. B. 2013. TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. *Plant Mol. Biol.* 83:279–285.

Wong, J., Gao, L., Yang, Y., Zhai, J., Arikit, S., Yu, Y., Duan, S., Chan, V., Xiong, Q., Yan, J., Li, S., Liu, R., Wang, Y., Tang, G., Meyers, B. C., Chen, X., and Ma, W. 2014. Roles of small RNAs in soybean defense against *Phytophthora sojae* infection. *Plant J.* 79:928-940.

Xie, K., and Yang, Y. 2014. RNA-guided genome editing in plants using a CRISPR–Cas system. *Mol. Plant* 6:1975–1983.

Xu, L., Zhang, W., He, X., Liu, M., Zhang, K., Shaban, M., Sun, L., Zhu, J., Luo, Y., Yuan, D., Zhang, X., and Zhu, L. 2014. Functional characterization of cotton genes responsive to Verticillium dahliae through bioinformatics and reverse genetics strategies. *J Exp. Bot.* 65:6679–6692.

Yaeno, T., Li, H., Chaparro-Garcia, A., Schornack, S., Koshiba, S., Watanabe, S., Kigawa, T., Kamoun, S., and Shirasu, K. 2011. Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector Avr3a is required for its stability in host cells to modulate plant immunity. *Proc. Nat. Acad. Sci.* 108:14682-14687.

Yi, S. Y., Min, S. R., and Kwon, S.-Y. 2015. NPR1 is instrumental in priming for the enhanced flg22-induced MPK3 and MPK6 activation. *Plant Pathol. J.* 31:192–194.

Yoshioka, Y., Sakata, Y., Sugiyama, M., and Fukino, N. 2014. Identification of quantitative trait loci for downy mildew resistance in cucumber (*Cucumis sativus* L.). *Euphytica*. 198:265–276.

Zhang, S. P., Liu, M. M., Miao, H., Zhang, S. Q., Yang, Y. H., Xie, B. Y., Wehner, T. C., and Gu, X. F. 2013. Chromosomal mapping and QTL analysis of resistance to downy mildew in *Cucumis sativus*. *Plant Dis.* 97:245–251.