FUNCTIONAL ANALYSIS OF SOYBEAN PR-10 GENES

BY

MARIA LAURA MALVINO

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Advisor:

Assistant Professor Steven J. Clough

ABSTRACT

Soybean (*Glycine max)* is the most important oilseed source and one of the most important crops worldwide. In 2014, 249 million metric tons were produced across the world. Even though the worldwide production is increasing every year, scientists, farmers and companies are still struggling to match production increases, to increase in world population growth. The challenge is not simply to produce more, but also to produce more on the same area of land. A main limit to increasing crop yields is the variety of diseases, mainly bacterial and fungal. It is estimated that around 15% of the crop production every year is lost due to biological threats. Understanding how soybean responds in defense to pathogens at a molecular level would help produce innovated seeds and plants that can better withstand biological attacks. This research was designed to characterize a soybean gene family that responds to multiple pathogens within a few hours of infection, the PR-10 gene family. We identified six members of the PR-10 gene family based on expression patterns from in-house microarray studies. Gene-specific PCR primers were designed to clone full-length cDNA of selected PR-10 genes. The cDNA were sequence verified and transferred into an *Agrobacterium* overexpression vector, expression controlled by the CaMV35S constitutive promoter. Four PR-10 family members were transformed into *Arabidopsis thaliana,* and alterations in defense responses were monitored in the PR-10 transformants. Two RNAi constructs were made for future transformation into soybean, to silence these genes and ascertain their function in soybean defense. Additionally, the response of a PR-10 promoter was assayed by studying GFP expression controlled by a PR-10 promoter versus the constitutive promoter GmUbi. These research results increase our understanding of PR-10 function and verify the effectiveness of PR-10 in defense response to pathogen infection, which could potentially lead to the development of markers that are associated with pathogen resistance, and also provide genetic material for basic research and possible development of transgenics with enhanced resistance.

To my parents

ACKNOWLEDGEMENTS

This project was possible thanks to many wonderful people and institutions that supported me through these 2 years. First, I would like to truly thank my advisor Dr. Steven Clough for accepting me in his laboratory and his mentoring during these 2 years. I would like to extend my gratitude to the rest of my committee, Dr. Lila Vodkin and Dr. Yoshie Hanzawa, for their patience and for accepting being part of this endeavor. I would like to give a special thanks to Dave Neece, whose guidance was key to the success of this project. I would like to thank my colleagues and undergraduates for their expertise, assistance and cheerful companionship. Finally, I want to express my gratitude to my husband, Ezequiel Fourcade, for somehow putting aside his career to be with me in the U.S. and for putting up with my "thesis mood" once again. Last but not least, I would like to thank the sponsors for the funding: USDA and Fulbright.

TABLE OF CONTENTS

CHAPTER 1

1. Introduction

Soybean (*Glycine max)* is the most important oilseed source and one of the most important crops worldwide. It is the most important legume and ranks fourth in terms of global crop production, together with rice, wheat and maize. Because of the high oil and protein content, soybean is mainly used as a major ingredient for livestock feed. Even though only a smaller percentage is used to make processed food for human consumption, many countries in the world still depend on soybean as a key source of dietary protein. Soybean is being used to produce soymilk, soy flour, tofu, soy protein and many others food products. It is mainly consumed in Asia, where the local consumption exceeds the production. Soybean is well known for its health benefits, such as a good source of vitamin B, manganese, copper, phosphorus, molybdenum, and omega 3-fatty acids. It has been shown that a daily intake of processed soybean products can help decrease several chronic diseases like cardiovascular diseases, as it lowers LDL cholesterol. A third major use of soybean is not food related, but instead is industrial. Soybeans oils can be used to manufacturer soaps, plastics, cosmetics, inks, resins, solvents, and crayons, among other products. Soybean oil is the primary source of biodiesel in the U.S. In 2014, and 249 million metric tons were produced across the world. Even though the worldwide production is increasing every year, scientists, farmers and companies are still struggling to keep up with the world feed demand, with population growth estimated to increase by another 2 billion people by 2050.

The challenge is not only to produce more, but also to increase the yield on the land currently used for cultivation, as the available world farmland is assumed to be maxed out. It is noteworthy that crops are usually threatened by a variety of diseases such as bacterial and fungal diseases. It is estimated that around 15% of the crop production every year is lost due to biological attacks. Given the current changes in global weather patterns, as higher latitudes warm, the impact of such diseases has increased in such a way in the past few years that attention has to be paid in order to minimize and decrease effects on yield, particularly in crops like soybean which constitute a major part of the global diet. Understanding how soybean responds in defense to pathogens at a molecular level would help produce innovated seeds that can defend better to biological attack. This research is designed to characterize soybean genes that respond to pathogens. We analyzed in-house microarray data from multiple disease studies and identified genes that respond strongly and fairly specifically to multiple pathogens. One such class of genes that responded often in our disease studies was the PR-10 (pathogenesis related 10) gene family. We obtained full-length cDNA of six selected pathogen-responsive PR-10 genes and cloned them

into *Agrobacterium tumefaciens* overexpression vectors for transformation into *Arabidopsis thaliana*. RNAi constructs to PR-10s were also made for future transformation into soybean. The transgenic Arabidopsis plants were then assayed for altered responses to pathogens; the bacterial pathogen *Pseudomonas syringae* and fungal pathogen *Sclerotinia sclerotiorum* were used as experimental pathogen representatives, as their modes of infection dramatically differ. Additionally, the response of a PR-10 promoter was assayed by studying *gfp* expression off of a PR-10 promoter. These research results increase our understanding of promoter function, allowing us to detail specificity of gene expression and verify effectiveness of PR-10 in defense response to pathogen infection, which could potentially lead to the development of markers that are associated with pathogen resistance, and also allow us to generate a collection of soybean promoters for basic research and possible transgenics with enhanced resistance specifically induced only during pathogen attack.

2. Literature review

a. A brief history of soybean

Soybean (*Glycine max)* is a legume belonging to the order Fabales, is an autotetraploid and has 20 pairs of chromosomes (Arumuganathan and Earle, 1991). It is native to East Asia but is now grown across the globe, and is one of the most important oilseed sources and crops worldwide. Regarding its ancestry, it was first thought that *G. max* was derived from *Glycine soja* and domestication occurred about 6,000 to 9,000 years ago. However, there is a second hypothesis which suggests that *G. max* derives from a *G. soja/G. max* complex around 270,000 years ago. So far there is no conclusive evidence to prove either hypothesis (Kim et al., 2010). Moreover, domestication most likely did not occur in only one place, but instead in three. According to recent archeological evidence, soybean domestication occurred in the Eastern Yellow River basin in North China, South Korea, and Japan (Gyoung et al., 2011). The first reports of soybean export came out of China in 1908, when *G. max* was exported from the city of Manchuria to the western hemisphere; initially to England and later to the United States. By the 1930's the United States became a major soybean producer. Afterwards, in the 1970's, South America also became a major producer, led by Brazil. Nowadays, the United States, Brazil and Argentina are the top worldwide soybean producers (Shurtleff and Aoyagui, 2004). Together, in 2014 these three countries produced 232.1 million metric tons, which represented 90% of the worldwide production.

The importance of soybean relies largely on its dietary protein and oil content, but also in its capacity to fix nitrogen during symbiosis with rhizobia, which is important to establish a sustainable agriculture and reduce production costs (Chung et al., 2014).

b. A description of pathogens that affect soybean

The main concern with crop decrease is that yields are often compromised, and any decrease in yield for a major field can translate into large economic losses, as well further stress our ability to feed world populations (Berge et al., 2014). The most common diseases that affect soybean are usually caused by bacteria, viruses, nematodes and fungi. Some of them need a vector to be spread, such as insects, wind or soil movement. Some diseases cause such a yield loss that a disease control has to be implemented. The first sign is the visible evidence; some of them are typical of a disease and other symptoms are shared among diseases. To efficiently manage the disease the best way is to adopt long-term production strategies. Some of the more common diseases of soybean caused by bacteria or fungi are discussed below.

Bacterial blight (caused by *Pseudomonas syringae pv. glycinea):* Bacterial blight was first reported in 1902 as a pathogen of lilac. It was rapidly found to be a disease of many plants, including important crops like soybean. Each *P. syringae* strain is fairly host specific, and therefore the pathover nomenclature system has been adopted to indicate the host plant from where the strain was originally isolated. Early symptoms of bacterial blight are water-soaked leaf spots, followed by chlorosis, and finally necrotic lesions surrounded with yellow halos. This bacterial pathogen enters through the leaf stomata or plant wounds, and once in the apoplast continues to replicate and damage host cells as the bacteria consume host resources and counterbalance host defenses. *P. syringae* pv. *tomato* strain DC3000, which not only infects tomato but also *A. thaliana,* has been used extensively with Arabidopsis as a model system to study molecular fundamentals of plant-bacterial interactions (Engl et al., 2014). Extensive research on *P. syringae* pathovars has led to a detailed understanding of this pathogen. It is a Gram-negative bacterium that utilizes a Type III secretion system to deliver effector proteins into a host in efforts to weaken defenses. This pathogen does not enter the host cells directly but interacts with the host cytoplasm from outside the plant wall, until host cells succumb to infection and eventually lyse. This unusual ability, compared to animal strains, makes it the key to the pathogenicity. In susceptible hosts it elicits foliar chlorotic and necrotic spots. On immune hosts, the plant recognizes one or more of the pathogen-released effector proteins to trigger the hypersensitive response (HR) defense, leading usually to microscopic lesions and a halt to pathogen multiplication. HR often occurs within 24 hours of

infection, while in susceptible host plants the symptoms are usually not observed until about 48-72 hours post infection. For both these types of interactions, one can see the expression of numerous bacterial genes such as: *hrc* (HR and conserved), *hrp* (HR and pathogenicity), *avirulence (avr)* and *Hrp*-dependent outer protein *(hop)* genes. *Hrc* and *hrp* genes encode the Type III secretion system, and *avr* and *hop* genes encode effector proteins injected by the system into plant cells. *Avr* proteins are effectors that get recognized by the host R gene surveillance system, triggering the HR (Collmer et al., 2000). Although bacterial blight is one of the most well-studied plant diseases at a molecular level, it is generally not a serious disease threat in most soybean growing areas. If control is needed, it is advised to plant resistant varieties, use tillage to promote the decay of infected plant residues, and plant pathogen-free seed.

White mold or Sclerotinia stem rot (*Sclerotinia sclerotiorum)*: *S. sclerotiorum* is ubiquitous in temperate arable lands around the globe and infects nearly all dicotyledonous crops of the world, often causing major economic damage to sunflower, canola, common bean and soybean. The severity of white mold disease is highly dependent upon the environment as the pathogen needs near 100% humidity and cool temperatures (below 30ºC) for ideal infection conditions. Therefore, severe disease incidences tend to be sporadic and regional. But when it hits, entire fields can be destroyed. Being a necrotrophic fungus, it needs to start infection on dead or dying tissue; for soybean the initial tissue to be infected are the dead/dying floral tissues. The name comes from the hard, black, melanized sclerotia produced by the fungus. These peppercorn-size particles serve as overwintering structures, as they can survive for years in the soil. The disease cycle starts when the environmental conditions are correct causing the sclerotia to produce apothecia that shoot ascospores into the air. If the ascospores land on senescing flowers they will germinate and infect, completely rotting the flower. If these infected flowers are resting on a soybean leaf, and the temperature is below 30°C and relative humidity is near 100%, then the fungal myclelia will release oxalic acid that will weaken or kill the underlying host tissue, and the fungus will infect using specialized infection 'compound appressoria' (Li et al., 2012) that resemble a collection of short, branching myclelia, and invade the host. As long as the weather remains cool and wet, the disease will progress, eventually colonizing and killing the entire plant. Because soybean does not have good resistance to *S. sclerotiorum*, substantial loses (up to 100%) in isolated areas can occur (Arumuganathan and Earle, 1991). After substantial growth and nutrition becomes limiting, this fungus will produce the sclerotia, that can survive for years in soil, awaiting the next ideal time to germinate and start the disease cycle over again. Even though many quantitative trait loci (QTLs) that control partial resistance have been mapped in the soybean genome, there are still no commercial resistant cultivars. All

resistance in soybean is still partial and therefore, the only control option is the use of fungicides, if it is economically advantageous. Crop rotation with monocots and deep tillage can also reduce the field inoculum. New approaches with genetic engineering, such as introduction of enzymes that degrade oxalic acid, have the potential to give higher resistance in the future

The following diseases will only be mentioned briefly as they are not part of this thesis research. Brown spot (*Septoria glycines)*: Minute spots appear and expand into large and brown ones. The disease expands from lower to upper leaves. This can be prevented with tillage, crop rotation and the use of disease-free seeds.

Frogeye leaf spot (*Cercospora sojina)*: Leave surface is covered with small spots with dark-brown margin. In humid weather light gray to white spores are produced. Tillage and crop rotation helps prevent it.

Cercospora blight and purple seed stain (*Cercospora kikuchii)*: It affects the leaves becoming reddish-purple and bronzing, pods become black and seeds have purple stains. It is usually prevent with tillage, crop rotation and using treated seeds.

Target spot (*Corynespora cassiicola*): Leaves are covered with round to irregular reddish-brown lesions. Depending on the geographical area, the spray with fungicide may be beneficial.

Anthracnose (*Colletotrichum truncatum*): It affects leaves, pods and stem as brown lesions appear, making the tissue turn brown and have an early senescence. Sometimes the pod fail to produce seeds, but if produced the seed has black lesions. This can be avoided by tillage, seed treatment, crop rotation, and the use of fungicide.

Pod and stem blight (*Phomopsis longicolla*): Infects the seeds causing it to crack. Pods stem and leaves are affected by blight. Tillage, crop rotation, fungicide and seed treatment are recommend to control it.

Bacterial pustule (*Xanthomonas campestris pv. glycines*): It is similar to soybean rust as pustules form in the center of lesions in the leaves. It is prevented with the same treatments as bacterial blight.

Charcoal rot (*Macrophomina phaseolina*): Reddish to brown stains appear on the stem, and fungus can be found in taproots. Irrigation, good soil fertility and rotation can help avoid it (Hartman et al., 2011).

c. Overview of molecular plant-microbe interactions

The interaction between plants and microbes has been estimated to have started around 460 million years ago, ever since plants evolved. Plant-microbe interactions are of increasing importance for agronomists, both for the beneficial aspects (such as in nitrogen fixing bacteria) as well as negative aspects (pathogens). The increase of knowledge in this area has been exponential in recent years as they affect classical areas such breeding and plant physiology, causing a great impact on our understanding.

Because nitrogen is a limiting element it is critical for plant development and growth. Nitrogen is an essential component of chlorophyll, and is the key building block for amino acids, therefore being essential for plant life. Bacteria in the *Rhizobiaceas* family carry out biological nitrogen fixation. Its importance is notorious not only from the cost-effectiveness point of view, but also from the ecological perspective as these bacteria constitute a natural strategy for developing sustainable agricultural strategies compared to the use of chemical fertilizers. Initially, nitrogen fixation occurs because of the coordinated activity of proteins, gene and metabolites which activate transcriptional factors inside bacteroids and activate signal transduction. As a consequence, we can see activation and repression in specific metabolic pathways that produce metabolites necessary for preventing the microenvironmental conditions inside the nodules. Lately, the use of high-throughput technologies has been used to study bacterial nitrogen fixation at a genome scale analysis in to order better understand the biological model of this symbiosis interaction between plant and bacteria. However this analysis is being challenged by large amounts of data which requires efficient and coherent interpretation using on computational modeling of genomic data. (Resendis et al., 2011)

Even though the interaction of Rhizobiaceas is beneficial to plants, many microbes form a negative association with plants. Many microbes are considered economic pathogens because they infect plants, removing nutrients to support their own population growth, resulted in limitation of global food supplies. There are many agents that can cause disease in plants, such as fungi, oomycote, protozoa, bacteria, viruses, phytoplasmas and spiroplasmas (Dickinson, 2000). When pathogens attack a plant, plants defend themselves with both active and passive defense mechanism. Passive defense are pre-existing mechanism that include strategically positioned reservoirs of antimicrobial compounds and structural barriers that prevent the colonization of the tissue. On the other hand, the active defense often involves production of reactive oxygen species, phytoalexins, lignification of the cell wall, and the

activities of pathogenesis related (PR) proteins, of which, PR-10 genes are the subject of this Master's study (Upadhyay et al., 2014).

d. A description of PR genes

Unlike animals, plants do not have acquired immunity or the clonal-selection immune system. Instead, they protect themselves by producing antibiotic compounds like phytoalexins, antibiotic proteins and by hardening their cell walls and limiting easy access to nutrients. Genes associated with these responses can be found expressed in the host plant, induced by the pathogen-derived effectors (proteins) or elicitors (chemicals/molecules).

Important components of the plant defense system are the resistance (R) genes. The plant R gene products play an important role by recognizing pathogens via direct or indirect interaction of pathogen effector proteins. R-gene recognition leads to rapid, robust activation of defenses, called the hypersensitive response (HR), a mechanism that protects plants from many pathogens. If recognition occurs, pathologists have named those specific R-gene recognized effector proteins as 'avirulence' (avr) factors, as their presence in the pathogen renders that pathogen avirulent. When effectors are not recognized by the host, then effectors are often called 'virulence' (vir) factors as these proteins have often been found to enhance virulence for the pathogen. The recognition between an avr and R protein is termed an incompatible reaction, as the HR triggers a multitude of toxic and metabolic disruptions that lead to cell death at the point of infection, preventing the pathogen from spreading to the rest of the plant. As in many plants, R genes in soybean are found in clusters, mostly on chromosomes 2, 7, 8 and 9. Clustering is hypothesized to aid in rapid *de novo* generation of R genes as the abundance of duplicated DNA sequences will enhance recombination, rearrangements, and deletions, corroborating the theory that genetic duplication mechanism is responsible for diversification and evolution of this gene superfamily (Wanderley-Nogueira et al., 2012).

In general, the genes that are activated in plants, whether resistant or susceptible, in response to pathogens tend to be the same, and to be expressed in the same direction (increased or decreased). What seems to make the difference between successful or failed infection is often the timing and robustness of the activation of these defenses (Zou et al., 2005; Tao et al., 2003). More resistant plants tend to hasten the activation of these defenses, but the global gene expression pattern between resistance and susceptible is often very similar, including in soybean (Zou et al., 2005; Calla et al., 2014). One class of genes that is fairly specifically expressed during defense reactions is the PR genes. PR

proteins are produce in response of biotic attack from fungi, bacteria or viruses, and tend to be fairly specific to plants. PR proteins are not only associated to plant defense, but some apparently also play important roles during development given their consistent localization in vacuolar compartment and apoplast and their differential induction by exogenous and endogenous compounds. (Upadhyay et al., 2014).

PR proteins were first associated with tobacco mosaic virus (TMV) infection of tobacco plants. It was later discovered that PR proteins include hydrolytic enzymes and defensins, which specifically act by hydrolyzing the pathogen cell walls and disrupting the membrane as well. So far there are 17 classes of PR proteins recognized in plants (Fernandez et al., 2012). Overexpression of a single individual PR gene might not be better than the coordinated expression of a set of PRs. Not all PRs genes are expressed during pathogenesis attack, the expression can depend on the type of pathogen attacking the plant. For instance, in *A. thaliana* it has been shown that PR-1, PR-2 and PR-5 are expressed more strongly in respond to biotrophic pathogens and are induced by salicylic acid (Spoeland and Dong, 2012). Most PR proteins are known to have antifungal activities, but still their molecular mechanisms are not well understood. PR-1 is the most abundantly expressed PR protein in several plants, such as tomato, *A. thaliana*, tobacco and apple. PR-5 exhibits antifreeze activities in overwintering monocots and has also antifungal activity. Moreover, PR-8, PR-9, PR-14, PR-15 and PR-17 are believed to be involved in plant defense responses but their molecular mechanisms are not known yet. It is important to notice that PR proteins are not always pathogen-specific, but may also be expressed during abiotic stresses such as cold, light and osmotic changes, pollen maturation, leaf senescence and development. This would explain why PR-2 was found to be expressed at a basal level in almost all tissues assayed (Wanderley-Nogueira et al., 2012).

e. PR-10: what's known

The PR protein PR-10 is the focus of this Master's thesis. The PR-10 protein family consists of acidic proteins with a molecular weight of approximately 17 kDa. Some members of this family have been found in different angiosperms like pea, bean, parsley, potato, and soybean, as well as in monocots like rice, lily and asparagus. PR-10 proteins differ from most PR proteins in that they are intracellular, and this was first indicated when no signal peptide was found (van Loon et al., 2006). PR-10 proteins have been sequenced and found to have similarities with a major allergen from birch (*Betula alba*), known as Bet v 1. This allergen causes type I allergic reactions, like bronchial asthma and rhinoconjunctivitis. It is a

major threat to the public in the northern hemisphere where birch is grown, as it affects around 15% of the population. Bet v 1 is a part of the pollen granules and triggers IgE binding in 95% of the cases in allergic patients (Breiteneder et al., 1989). It was found that Bet v 1 has RNAse activity (Albrecht et al., 1996). The complete amino acid sequence of Bet v 1 was first reported in 1989, and was found the high homology to an unknown plant disease resistance response gene of pea, suggesting that this allergen has a role in plant defense responses. (Breiteneder et al., 1989)

Another group of plant genes were found to be expressed strongly during starvation, and to have high homology to Bet v 1. These genes were called Starvation-associated Message 22, or SAM22, and were also found to be allergenic. SAM22 mRNAs were found to be expressed in the roots of soybean seedling, and roots and leaves of mature soybean plants. Moreover, the accumulation of SAM22 was especially high in senescent leaves. At that time it was also found that SAM22 expression is induced in young leaves when exposed to various environmental stresses, such as wounding. This information, all taken together, had at that time suggested that SAM22 genes are induced not only by various stresses, but also by developmental signals. As PR-10's became defined, it became clear that SAM22's should also be classified as PR-10s based on high sequence similarity. It is indeed interesting the finding that SAM22/PR-10 genes also have a role in development, as roots undergo stress physiology as they are wounded during penetration of young lateral roots; moreover, the expression of these genes in leaves occurs during senescent due to programmed cell death, part of the life cycle and different to necrosis (Dring et al., 1992).

Regarding PR-10's ribonuclease activity, apparently its function derives from a conserved C-terminal alpha helix playing an important role during ribonucleic acid degradation, and this function has been conserved during evolution. RNase activity in plants is altered in order to regulate gene expression when exposed to external stimuli such as pathogen attack. It was found in cotton that once the plant interacts with the pathogen, PR-10 degrade RNAs in order to contribute to the hypersensitive reaction. Furthermore, PR-10 proteins may selectively degrade mRNAs during biotic stress, in order to help the plant return to its normal physiological level. This ribonuclease activity was also demonstrated in 1996 in the birch pollen allergen Bet v 1 (Zhou et al., 2002). It has also been demonstrated that PR-10 in hot pepper needs to be phosphorylated in order to have an increased ribonucleolytic activity to cleave invading viral RNAs, which should be of importance in the antiviral pathway during *in vivo* attack. It is still unknown where the PR-10 protein is phosphorylated and by which kinase (Park et al., 2004).

Recent studies showed that, in pepper (*Capsicum annuum)*, that PR-10 proteins physically interact with

leucine-rich repeat 1 (LRR1) proteins to assist in recognition of pathogen attack and to activate defense genes. In pepper, the PR10/LRR1 complex is excreted from the cytoplasm to the apoplast. The formation of this complex in the cytoplasm prior to export is required for cell death induction. It is noteworthy that the localization of the PR-10/LRR1 complex to the apoplast may be a consequence of cell-wall degradation, rather than a cause for cell death. Moreover, LRR1 also phosphorylates PR-10 and promotes its ribonuclease activity, enhancing the cell-death signaling. The importance of PR-10 in the HR was supported by silencing PR-10 via the virus induce gene silencing technique. Pepper plants were first silenced for PR-10 and then PR-10/LRR1, and afterwards infected with an avirulent strain of *Xanthomonas campestri* pv. *vesicatoria*. In the PR-10 silenced plants it was observed a significantly greater inhibition of electrolyte leakage, lower accumulation levels of H_2O_2, and downregulation of defense response genes like PR-1, *Defensin1* (DEF1), *SYSTEMIC ACQUIRED RESISTANCE8.2* (SAR82) and *PEROXIDASE2* (PO2). All these data together supports the fact that PR-10 is required for defense response and HR, including ROS burst, salicylic acid accumulation, and defense-related gene induction. Moreover, the over expression of PR10/LRR1 in transgenics *A. thaliana* lines infected with *P. syringae* showed a reduction of necrotic disease symptoms in leaves, and a less significant growth of both *P. syringae,* with or without avrRPM1. In these leaves it was also observed an increase of H_2O_2 compared to the wild type (Choi et al., 2012).

f. Use of Arabidopsis transgenics for functional characterization of a soybean gene

A good way to study individually the importance of a gene is by transforming *A. thaliana* and studying the subsequent transgenic plants. Even though *A. thaliana* is not economically important as a crop, its basic plant biochemistry is highly conserved among other plants, giving this plant a lot of importance from the genetic, physiological and biochemical point of view. *A. thaliana* is convenient to study, given that it only needs water, air, light and few minerals to grow, and its life cycle is fast. Moreover, the Floral Dip method (Clough and Bent, 1998) a technique that delivers Agrobacterium to the interior of the developing gynoecium prior to locule closure for efficient transformation (Desfeux et al., 2000), provides an easy, robust and efficient way of transforming *A. thaliana*. Additionally, it self-pollinates, is easy to cross-pollinate, grows well under artificial lighting and one plant can produce thousands of offspring. These features makes it easier and quicker to make hypothesis and draw conclusions in *A. thaliana*, that can then be used to make important decisions on how to best study plants that are economically important (Sussman et al., 2013).

Important aspects of gene regulation are promoters, and little is known so far about soybean promoters. In this study GmPath21, a PR-10 promoter, was studied. For the purpose of the research, the green fluorescent protein (GFP) was used as a reporter of the promoter activity. GFP is a small protein of 238 aminoacids and it is naturally found in jellyfish *Aequorea victoria*. Because of its property of exhibiting bright green fluorescence to light in the blue to ultraviolet range, it is used as a molecular marker inside living organisms as it can be dynamically imaged.

A GFP gene was cloned into a pCambia plasmid, regulated by a PR-10 promoter, GmPath21. As a control, a second construct was built but with a constitute soybean promoter, GmUbi. Even though soybean transformation was first reported in 1988 (Hinchee et al., 1988), it is still an inefficient method, making it difficult to study promoters in soybean. Therefore, transforming other plants like *A. thaliana* is a good heterologous plant expression system to use as an alternative.

g. Functional gene analysis in soybean using RNAi

Another good tool used to assess the importance of the action of a gene or gene family, is RNAi which will knock down the expression of the target gene(s). Fragments of 21-23 nucleotides guide the mRNA degradation. These are small interfering (si) RNA that initiate the silencing when binding the mRNA target and cleaving it. On the other hand, endogenous miRNAs do not only degrade mRNA but also inhibit mRNA translation. Moreover, there is a third mechanism for endogenous miRNAs that involves silencing of heterochromatin but it is still not well understood.

Double-stranded RNA (dsRNA) molecules, like the ones constructed in this work, initiate RNAi by converting long dsRNA into smaller 21-23-nt siRNAi by the enzyme Dicer, which is an RNAse III-type enzyme. This step creates RNAs with a phosphate group at 5' ends and 2-nt overhangs at 3'. Even though post-Dicer cleaved products (siRNAs) can be introduced exogenously into cells to induce RNAi, there is evidence that shows long dsRNA has a more potent triggering effect in silencing than siRNAs. siRNAs have two strands with sense and antisense configuration with respect to the target mRNA. This antisense strand is also called the guide strand because it serves as a guide for silencing the target. The sense strand is called passenger strand.

siRNA is assembled into an RNA-induced silencing complex (RISC) forming an RNA-protein complex, this last one known as siRISC because it incorporates the guide strand into RISC. This guide strand then binds the target RNA and cleaves it, therefore silencing gene expression (Tariq, 2007).

3. Research Justification

It has been estimated by the Food and Agriculture Organization of the United Nations that by 2030 the global food demand will increase by 50%. Therefore, scientist, farmers and companies are facing one of the hardest challenges of the 21st century. Not only is it needed to increase the production of crops but also to reduce the yield loss due to pathogen attack, which represent a heavy load in agriculture. Plant pathogens cause diseases which decrease crop yields resulting in large economic loses and, more important, threatening global food security. Therefore, it is of great importance to better understand the mechanistic of the plant-pathogen interaction and to use this knowledge to make crops more resistant and improve crop management. Pathogenesis related genes encode proteins produced by plants in the event of pathogen attack. They act against viruses, fungi or bacteria and are part of the active defense of the plant. It is known that PR-10 genes are induced by microbial attack, fungal elicitors and wounding stress; however, their specific function in plant immunity and cell death signaling In soybean are still not clear. Thus, this study aims to begin shedding some light on the specific function of four soybean PR-10 candidate genes.

CHAPTER 2

1 Introduction

Soybean yield is restricted by disease. To maximize yields, seed producers need to minimize the negative factors, including pathogens. Plants can withstand invasion of pathogens with both preformed and inducible mechanisms. To assist producers in ensuring that defenses to restrict pathogens from spreading and developing, it would be beneficial to understand which genes are responsible for producing the most effective defense-related proteins. Previous research was found that PR-10 genes are strongly induced in response to *Pseudomonas syringae* infection (Zhou et al., 2002) as well as in response to *Sclerotinia sclerotiorum* (Calla et al., 2014.) and *Fusarium virguliforme* (Radwan et al., 2012).

The PR-10 protein family consists of proteins that are acidic, with molecular weights of approximately 17 kDa. Some members of this family have been found in different angiosperms like pea, bean, parsley, potato, and soybean, as well as in monocots like rice, lily and asparagus. PR-10 proteins differ from most PR proteins in that they are intracellular, and this was first indicated when no signal peptide was found (van Loon et al., 2006). PR-10 proteins have been sequenced and found to have similarities with a major pollen-associated allergen from birch (*Betula alba*), known as Bet v 1, which causes type I allergic reactions such as bronchial asthma and rhinoconjunctivitis. Bet v 1 was found to have RNAse activity (Albrecht et al., 1996) and triggers IgE binding in 95% of the cases in allergic patients (Breiteneder et al., 1989). Another class of proteins was also found to have high homology to Bet v 1, Starvation-Associated Message 22, or SAM22, and these proteins were also found to be allergenic. As PR-10's became defined, it became clear that SAM22's should also be classified as PR-10s based on high sequence similarity. It is indeed interesting the finding that SAM22/PR-10 genes also have a role in development, as roots undergo stress physiology as they are wounded during penetration of young lateral roots; moreover, the expression of these genes in leaves occurs during senescent due to programmed cell death, part of the life cycle and different to necrosis (Dring et al., 1992).

Recent studies showed that, in pepper (*Capsicum annuum)*, that PR-10 proteins physically interact with the leucine-rich repeat 1 (LRR1) protein , to assist in recognition of pathogen attack and activation of defense genes. In pepper, the PR10/LRR1 complex is excreted from the cytoplasm to the apoplast. The formation of this complex in the cytoplasm prior to export, is required for cell death induction. It is noteworthy that the localization of the PR-10/LRR1 complex to the apoplast may be a consequence of cell-wall degradation, rather than a cause for cell death. Moreover, LRR1 also phosphorylates PR-10 and

promotes its ribonuclease activity, enhancing the cell-death signaling. The importance of PR-10 in the HR was supported by silencing PR-10 via the virus induce gene silencing technique. Pepper plants were first silenced for PR-10 and then PR-10/LRR1, and afterwards infected with an avirulent strain of *Xanthomonas campestri* pv. *vesicatoria*. In the PR-10 silenced plants it was observed a significantly greater inhibition of electrolyte leakage, lower accumulation levels of H_2O_2, and downregulation of defense response genes like PR-1, *Defensin1* (DEF1), *SYSTEMIC ACQUIRED RESISTANCE8.2* (SAR82) and *PEROXIDASE2* (PO2). All these data together supports the fact that PR-10 is required for defense response and HR, including ROS burst, salicylic acid accumulation, and defense-related gene induction. Moreover, the overexpression of PR10/LRR1 in transgenics *Arabidopsis thaliana* lines infected with *P. syringae* showed a reduction of necrotic disease symptoms in leaves, and a less significant growth of both *P. syringae,* with or without avrRPM1. In these leaves it was also observed an increase of H_2O_2 compared to the wild type (Choi et al., 2012).

To have a better understanding of the role of individual PR-10 genes in defense, we cloned cDNA to obtain the coding sequence minus introns corresponding to the candidate PR-10 genes. Candidate genes were cloned into an Agrobacterium transformation vector, expressed by a constitutive CaMV 35S promoter and transformed into *Arabidopsis thaliana*. Transgenic lines were then exposed to pathogen attack to assess the function of the individual PR-10 candidate genes. To observe expression from a PR-10 promoter, the promoter from one PR-10 gene was cloned in front of the *gfp* gene and expression analyses were conducted in Arabidopsis, showing pathogen-induced expression.

To facilitate future functional studies of PR-10s in soybean, a region of homology between putative PR-10 genes was cloned and used to make an RNAi construct with the goal of silencing multiple PR-10 paralogs to test the role of PR-10 in soybean defense to pathogens and pests. A Gateway cloning vector (from Chris Taylor, The Ohio State University) was used for this purpose. A vector containing a sense and antisense insert was constructed such that the resulting RNA molecule would form a 150 bp dsRNA hairpin that should trigger RNAi against PR-10 genes, producing small interfering (si) RNA that initiate the silencing when binding the mRNA target and triggering its cleavage. Furthermore, a second RNAi construct was built using part of the coding sequence of candidate gene #1 (Glyma07g37240.2= Glyma.07G243500.2) to determine how the silencing of this individual gene impacts defense in soybean.

2. Material and Methods

a. Gene Selection

To search for pathogen-inducible genes in soybean, we used our in-house database, Soybean Gene Expression Database (SGED), to search for genes that have the key term SAM22 or PR-10.

The Soybean Gene Expression Database (SGED: http://sged.cropsci.illinois.edu/) is an in-house gene information database that contains information of soybean microarrays (cDNA, oligo, Affymetrix) and RNA-seq experiments.

Once we obtained a list of candidate genes that matched the key terms, poorly matched genes were removed. The remaining GlymaIDs were clustered across different experiments (symbiotic, pathogenic, chemical and others). A final cluster of six genes was obtained, all of which appear to be responsive to pathogen infection (See Fig. 1).

Figure 1 depicts a heat map where expression ranges from yellow to blue. Yellow color indicates $\log_2$ ratios above zero, and blue color below zero; colors were more intense as they increased their distance from zero. In this research we focused on candidates genes that showed over expression in response to *Pseudomonas syringae* and *Sclerotinia sclerotiorum*. Six candidate genes were identified that be the focus of this study:

Internal annotation	**Transcript name**		**Location**	**Base pairs (bp)**
	Phytozome v9 ID's	**Gmax_275_Wm82.a2.v1**		
Gene #1	Glyma07g37240.2	Glyma.07G243500.2	Chr07:42331425..42332476 reverse	628
Gene #2	Glyma17g03340.1	Glyma.17G030100.1	Chr17:2210399..2211606 reverse	729
Gene #3	Glyma07g37270.1	Glyma.07G243600.1	Chr07:42336266..42345196 forward	660
Gene #4	Glyma17g03350.1	Glyma.17G030200.1	Chr17:2215796..2216870 reverse	688
Gene #5	Glyma09g04520.1	Glyma.09G040500.1	Chr09:3370431..3371819 forward	951
Gene #6	Glyma07g37240.3	Glyma.07G243500.3	Chr07:42331425..42332476 reverse	859

An alignment of the six candidate genes can be seen in Figure 2.

The following are the protein sequences expressed by each candidate gene:

Glyma07g37240.2 (158 residues):
MGVFTFEDEINSPVAPATLYKALVTDADNVIPKALDSFKSVENVEGNGGPGTIKKITFLEDGETKFVLHKIESIDEANLGYS
YSVVGGAALPDTAEKITFDSKLVAGPNGGSAGKLTVKYETKGDAEPNQDELKTGKAKADALFKAIEAYLLAHPDYN*

Glyma07g37240.3 (146 residues):
MGVFTFEDEINSPVAPATLYKALDSFKSVENVEGNGGPGTIKKITFLEDGETKFVLHKIESIDEANLGYSYSVVGGAALPD
TAEKITFDSKLVAGPNGGSAGKLTVKYETKGDAEPNQDELKTGKAKADALFKAIEAYLLAHPDYN*

Glyma17g03340.1 [157 residues]:
MGVFTSESEHVSPVSAAKLYKAIVLDASNVFPKALPNFIKSVETIEGDGGPGTIKKLTLAEGLGYVKHHVDAIDTENYVYN
YSVIEGSALSEPLEKICYEYKLVATPDGGSIVKSTSKYYTKGDEQLAEEYVKTGKERSAGFTKAIEDFIQANPDYN*

Glyma07g37270.1 [158 residues]:
MGVFTFEDETTSPVAPATLYKALVTDADNVIPKAVDAFRSVENVEGNGGPGTIKKITFLEDGETKFVLHKIEAIDEANLGY
SYSVVGGDGLPDTVEKITFECKLAAGANGGSAGKLTVKYQTKGDAQPNQDDLKIGKAKSDALFKAVEAYLLAHPDYN*

Glyma17g03350.1 [158 residues]:
MGIFTFEDETTSPVAPATLYKALVTDADNVIPKAVEAFRSVENLEGNGGPGTIKKITFVEDGESKFVLHKIESVDEANLGY
SYSVVGGVGLPDTVEKITFECKLAAGANGGSAGKLTVKYQTKGDAQPNPDDLKIGKVKSDALFKAVEAYLLANPHYN*

Glyma09g04520.1 [157 residues]:
MGVVTQIYDTPAAVPPTRLFKAMTLDFHNLFPKLVDSIHSIVFTQGNGGPGTIKKITTIEGDKTKYVLHRVDAIDEANFVY
NFSITEGTALADTLEKVSFESQLVEAPNGGSIRKVSVQFFTKGDATLSEEELTANKAKIQGLVKLVEGYLLANPDY*

b. Over-expression in Arabidopsis

Once the target candidate genes were identified, the next step was to construct a vector with the candidate gene of interest and a constitutive promoter. The pBIN-T Cloning Protocol of Clough Lab was used (Neece, 2014) for this purpose. The pBIN-T is a modified version of the pBIN-m-gfp5-ER vector, in which the native AhdI site was removed, and an AhdI cloning cassette was inserted allowing for direct

T/A cloning of PCR products. With the methodology used in the current research, time was saved as the candidate genes were directly cloned in the vector that was used to transform Arabidopsis.

RNA was extracted from *Glycine max*, genotype Williams 82, which was previously exposed to *Pseudomonas syringae*. It was chosen to use cDNA from these infected soybean leaves instead of cDNA from tissue exposed to *Sclerotinia sclerotiorum,* as plants exposed to the two pathogens not only share a very similar expression pattern, but the overexpression was induced stronger by *Pseudomonas syringae* (Fig. 1). RNA was extracted following TRIzol@ Reagent (Invitrogen, Carlsbad, CA, USA) with Phase Lock Gel-Heavy ™ (Brinkmann Instruments, Inc., Westbury, NY, USA). Afterwards, potential DNA contamination was removed using DNase I (Ambion). cDNA was synthesized using SuperScript III (Invitrogen) and was used as a template to obtain clones of the six candidate genes (Neece, 2013).

To amplify the candidate genes, conventional PCR was performed using the following program:

1 cycle	Initial denaturation	95ºC	30 seconds
35 cycles	Denaturation	95ºC	30 seconds
	Annealing	55ºC	30 seconds
	Elongation	68ºC	2 minutes
1 cycle	Extension	68ºC	5 minutes

The primers for the candidate genes were designed using Primer 3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) (See Table 1). Once amplified candidate cDNA were cloned, the PCR products were purified using Quiaquick PCR clean up (Quiagen #28106), and ligated into pBIN-T.

pBIN-T was first prepared by cutting it with AhdI to yield cloning sites for PCR fragments. AhdI cuts twice, resulting in two fragments: 440 bp and 12,900 bp. These fragments were separated by running a 1% LMP agarose gel. The upper band, 12,900 bp, was cut and the agarose digested with agarase in order to obtain the cut plasmid (See Fig. 3).

Ligation of PCR product into pBIN-T vector was performed on the same day as the ligation for maximum efficiency. This reaction was done using NEB Quick Ligation Kit. In this procedure it was important to use concentrated DNA ligase for the inefficient T/A cloning of PCR products. Moreover, to maximize the efficiency of the reaction, additional ATP was added.

NEB Turbo competent *E. coli* (High Efficiency) was transformed by heat shock, the cells spread on LB plus 50 ug/ml kanamycin agar plates, and incubated overnight at 37ºC.

E. coli colonies that grew were assumed to be successfully transformed with a ligated vector plus insert, and individual colonies were randomly selected and grown in liquid LB media containing kanamycin (50 ug/ml) for correct insert verification. Five ml of culture were grown overnight at 37ºC and plasmid DNA was extracted and purified using the QIAprep Spin Miniprep Kit (Qiagen).

DNA is extracted from *E.coli* to check that the candidate gene was inserted in the proper orientation. This was verified by running agarose gels using the cloned gene's specific forward primer, and the M13 For-21 primer as a reverse primer for PCR. The plasmid with the gene inserted in the proper orientation yielded a band of the candidate gene size plus 290 bp. If the candidate gene was inserted in the wrong direction it would not yield a PCR product. Afterwards, in order to check the full sequence, the PCR product was purified and sequenced (See Fig. 4).Once the right colony with the proper orientation was chosen, its plasmid was extracted and used to transform *Agrobacterium tumefaciens* (See Fig. 5).

Electroporation was the method used for transformation (Protocol from Steve Farrand's Lab, 1996). Floral dip method proceeded afterwards to transform *Arabidopsis thaliana* (Clough and Bent, 1998).

Arabidopsis thaliana was grown in a chamber at 23ºC with a photoperiod of 16 hours (8 hours of dark), light intensity of 130 umol/m^2sec and 60% of humidity. Plants were allowed to grow for around seven weeks to allow them to reach maturity and seeds pods became yellow and dried completely, to finally shatter and release the seeds in order to harvest them.

Seed selection

Seeds were selected on agar plates (0.8% Bacto Agar, ½X Murashige and Skoog salts, 50 ug/ml of kanamycin). Seeds that were not successfully transformed died, arrested the root development and/or lost chlorophyll. Seeds that were successfully transformed were transplanted to soil, and grown to seed. Three generations of transformants were achieved, in order to obtain plants that were homozygous for the pathogenesis related candidate genes. Third generation (T3) seeds were used to conduct diseases studies. T3 seeds were selected in plates where only when 100% of them grew, indicating the homozygous condition (See Fig. 6)

Disease assays

In order to characterize the resistance response of the four candidate genes, the different lines were inoculated with both *Pseudomonas syringae* and *Sclerotinia sclerotiorum*, each one-at-a-time in different plants. The objective was to monitor not only macroscopic disease symptoms for *Sclerotinia sclerotiorum* but also to count the bacterial population growth for *Pseudomonas syringae*.

For *Pseudomonas syringae,* disease assays were conducted on three leaves in eight different plants, in order to get consistent results. The inoculations (overnight culture used to make suspension of approximately 10e4 cells/ml) were done on day 0 using a syringe of 1 cc to hand infiltrate individual leaves, and samples of leaves were taken with a 9 mm hole punch at 24 h, 48 h and 96 h post infiltration in order to measure the growths of the bacteria, in both four transgenic lines and in the control Col-0. Leaf discs were ground in cold water, and dilution plated to obtain colony counts.

For *Sclerotinia sclerotiorum,* five leaves were taken from each plant (four lines and control) and were inoculated with *Sclerotinia sclerotiorum* after one day of growing the fungus at 28ºC in Difco™ potato dextrose agar plates. Leaves were incubated at room temperature in humid Petri dishes for 3 days with a 3 mm plug of the agar inoculated with *Sclerotinia sclerotiorum*. In order to evaluate the damage, pictures were taken at 24, 36 and 48 hours and the damage was measured using *AnalyzingDigitalImages@* software.

qRT-PCR

In order to confirm the presence of the transgene in each line, RNA was extracted from infected leaves with RNeasy plant mini kit (Qiagen) and measured with NanoDrop ND-1000 spectrophotometer. cDNA was synthesized using an SuperScript III cDNA synthesis kit (Bio-Rad,Hercules,CA,USA). The candidate genes were then amplified with the primers designed for each of them and using conventional PCR. A positive result was confirmed running agarose 1% gels and checking the band size.

c. Silencing of PR-10 paralogs

With the Gateway cloning system, one fragment was targeted for PCR cloning as its sequence was conserved among the six candidate PR-10 genes (from 271 to 438 bp in the consensus identity alignment). The amplified PCR fragment was later cloned into the entry vector CGT11050. *E. coli* was transformed with the ligation reaction. *E.coli* was then plated on LB plates containing chloramphenicol (50 ug/ml) to grow overnight at 37ºC. Individual colonies were selected to grow in liquid LB containing

chloramphenicol (50 ug/ml) to extract the plasmid to later verify entry vectors by PCR and sequencing. Once confirmed, The RNAi entry vector was recombined with the destination vector CGT11017A, using the LR Clonase reaction (Neece, 2014). The destination vector contained the terminator NOS, the FAD2 intron and the FMV promoter. These allowed to later verifying the correct insertion of the sense and antisense vector.

A second RNAi was constructed in order to only silence candidate gene #1 (Glyma07g37240.2= Glyma.07G243500.2). A region inside the coding sequence was amplified with PCR and cloned into an entry vector in order to the then be transferred into a destination vector. The procedure followed was the same as described for the consensus identity alignment.

d. Evaluation of a PR-10 promoter

For the purpose of this study, GFP was used as a molecular marker. Transformation in Arabidopsis with pGmPath21 (PR-10) and pGmubi promoters was done. The plasmids were constructed by John Finer's lab at The Ohio University (See Fig. 7 and 8) and sent to our lab for analysis. GmPath21 is a good candidate to direct specific pathogen-induced expression in hairy root assays, as according to its differential GFP expression in soybean hairy roots and RNA expression data for its gene, it is expressed weakly in roots and leaves, and is strongly induced by pathogens.

Each plasmid was first transformed in *Agrobacterium tumefaciens* and plated on LB. Colonies successfully transformed were selected with kanamycin and used for floral dip method to transform *Arabidopsis thaliana*, selecting for hygromycin resistant plantlets. Three generations of transgenic plants were grown in order to obtain homozygous plants to study the GFP expression driven by the GmPath21 promoter (the promoter to Glyma.17G030200).

T3 plants were inoculated with *Pseudomonas syringae* (OD_{600}= 0.14 in 10 mM $MgCl_2$) and were observed under a fluorescent dissecting microscope with a GFP filter every hour, during the initial four hours following leaf infiltration. Pictures were taken for the plant of study and the control plants: Columbia-0 and plants transformed with the GFP reporter fusions.

3. Results and discussion

a. Gene selection

Genes were selected using the key terms 'SAM22' or 'PR-10' in the Soybean Gene Expression Database (SGED: http://sged.cropsci.illinois.edu/). Even though initially six candidate genes were selected, only four of them could be successfully cloned and transformed in *Arabidopsis thaliana*.

The following genes could not be cloned or transformed:

Internal annotation	Transcript name		Location
	Phytozome v9 ID's	Gmax_275_Wm82.a2.v1	
Gene #3	Glyma07g37270.1	Glyma.07G243600.1	Chr07:42336266..42345196 forward
Gene #6	Glyma07g37240.3	Glyma.07G243500.3	Chr07:42336266..42345196 forward

Candidate gene #6 could not be cloned presumably because its transcript was not present in the RNA used to make the cDNA. Primers for both candidate gene #1 (Glyma07g37240.2= Glyma.07G243500.2) and #6 (Glyma07g37240.3= Glyma.07G243500.3) were the same, as these are variants of the same gene, presumably different alleles based on their nucleotide alignments (see APPENDIX A). Apparently, the transcript corresponding to Glyma07g37240.3 was of too low abundance relative to Glyma07g37240.2, to be cloned.

Candidate gene #3 (Glyma07g37270.1= Glyma.07G243600.1) was amplified by PCR, but the PCR product could never be cloned into the pBIN-T vector. Given the fact that primers and buffer concentrations, temperatures and Mg level were correct, there following could be a possibility for not achieving the ligation:

1- Candidate gene ##3 (Glyma07g37270.1= Glyma.07G243600.1) is toxic for *E. coli*: this might be solved by incubating plates at lower temperature (25 – 30°C), or by carrying out transformation using a strain that exerts tighter transcriptional control of the DNA fragment of interest (e.g., NEB 5-alpha F´ I^q Competent *E. coli*)

b. Over expression in Arabidopsis

Plants were transformed successfully with the following candidate genes:

Internal annotation	Transcript name		Location
	Phytozome v9 ID's	Gmax_275_Wm82.a2.v1	
Gene #1	Glyma07g37240.2	Glyma.07G243500.2	Chr07:42331425..42332476 reverse
Gene #2	Glyma17g03340.1	Glyma.17G030100.1	Chr17:2210399..2211606 reverse
Gene #4	Glyma17g03350.1	Glyma.17G030200.1	Chr17:2215796..2216870 reverse
Gene #5	Glyma09g04520.1	Glyma.09G040500.1	Chr09:3370431..3371819 forward

As every construct containing each of the candidate genes also contains a gene resistant to kanamycin, seeds were selected by growing them on agar plates containing kanamycin. Only plants that were successfully transformed grew normally and were transplanted, paying special attention to transplant those whose roots were fully growing inside the agar. This selection was done during each generation, in order to obtain the third generation of homozygous transgenic plants. Once the third generation was achieved, disease studies were conducted on those plants The number of plants that successfully grew for each generation and each candidate gene can be seen in Figure 9.*Pseudomonas syringae* study

Plants were grown until just prior to bolting, leaves inoculated, and bacteria counted by dilution plating at 24, 48 and 96 hours. Three replicates were taken for each plant; and up to five dilutions were done on each replicate in order to achieve an accurate count. The results can be seen in table 2 and figures 10 and 11.

It can be seen that the addition of soybean PR-10 gene #4 into *Arabidopsis thaliana* had an inhibitory effect on *P. syringae* during the first 48 hours, as the bacteria count decreased compared to the control Columbia-0. It cannot readily be said that the role of this gene alone makes a difference at 96 hours as standard errors are overlapping at this point with the control.

It is noteworthy that, inoculated with the HR strain, transgenic line #4 had a similar pattern as Columbia-0, except at 96 hours, where the counting even decreased more than transgenic line #1 (Glyma07g37240.2= Glyma.07G243500.2).

Bacteria growth monitoring in plants transformed with PR-10 gene #1 (Glyma07g37240.2= Glyma.07G243500.2), termed transgenic line #1, suggests an enhance resistance against both the HR and virulence strains of *P. syringae*. Transgenic line #1 showed a reduced bacterial count for both the virulence and HR strain at time points. These results for transgenic line #1 (Glyma07g37240.2=

Glyma.07G243500.2) support that soybean PR-10 can function in Arabidopsis, and over-expression enhances resistance to *P. syringae*.

Sclerotinia sclerotiorum study

Transgenic *Arabidopsis thaliana* plants expressing soybean PR-10 genes were also assayed for changes in their defense responses to infection by *Sclerotinia sclerotiorum*. Lesion on the leaves was measure at three time points with *AnalyzingDigitalImages*[@] software. Four leaves from each plant were taken and photographed at 24, 36 and 48 hours. In Table 3 and Figure 12 the results can be seen. The extended data (pictures and tables) can be seen in APPENDIX B.

As standard errors are overlapping between the control and the transgenic lines, even though we see a qualitative difference, we cannot conclude that there is a significant difference between the control and the overexpression of the different candidate genes. These results suggest that PR-10 is not enhancing resistance to *Sclerotinia sclerotiorum* in *Arabidopsis thaliana*. As PR-10 has been shown to enhance programmed cell death, the results are not surprising since *Sclerotinia sclerotiorum* is a necrotrophic fungal which means that it infects and kills host tissue and extract nutrients from the dead host cells, and would benefit from enhanced cell death, where as P. syringae would not.

qRT-PCR

Although transgenic plants were selected by gain of kanamycin resistance, it was also necessary to confirm that the transgene was actually being expressed. To determine gene expression, we assayed the transgenic plants via quantitative reverse-transcribed PCR (qRT-PCR). In transgenic line #1, th soybean PR-10 gene was confirmed to be transcribed by extracting RNA, synthesizing cDNA, performing PCR with primers design for this candidate gene, running the PCR reaction on a 1% agarose gel and visualizing a band of the expected size (628 bp) of the targeted region (Fig. 13).

Phenotypical findings

The addition of a foreign gene under constitutive expression into a host might have deleterious effects, as these genes are normally only expressed under certain stress and developing conditions in their native host. Therefore, at the beginning of the study, tranformants were visually observed for abnormalities. The observations were encouraging, as no major deformations were observed in any of the transgenic lines; however, the transgenic lines are a little bit smaller compared to the control, Columbia-0, while growing in the kanamycin agar plates during seed selection. Nevertheless, after

transplanted to soil, the plants showed not only the same size as the wild type, but were even slightly bigger for most of the plants of line #1 (Glyma07g37240.2= Glyma.07G243500.2). Moreover, some of the plants of line #1 (Glyma07g37240.2= Glyma.07G243500.2) showed not only a greater quantity of leaves, but also additional propagation characteristics. New rosettes appeared on the fluorescent, suggesting a possible 'runner type' propagation similar to strawberries; however, when these 'aerial rosettes´ where forced to be in continual contact with soil, no new roots appeared by 22 days. More stems were also produced in the plants with these characteristics (see Figures in APPENDIX C).

These finding are definitely encouraging because of the amount of biomass and seed produced. In future studies we should analyze in detail how the candidate gene caused the phenotypic differences. Knowing the number of insertions and place of insertion might be different compared to the few plants from transgenic line that didn't show this difference.

c. Silencing of PR-10 paralogs

In order to verify the role of PR-10 candidate genes in soybean, it was intended to silence as many candidate PR-10 genes as possible. For this reason one vector carrying a consensus fragment (168 bp) as an inverted repeat was built. Primers were designed accordingly to amplify those regions that belong to part of the coding sequence.

The consensus fragment was decided on the alignment of all the six PR-10 candidate genes. Geneious Basic 5.6.7 software was used for this purpose (Figure 14). The best matching area was used to design the primers that cloned the fragment which was later inserted in the different vectors.

The fragment cloned was sequenced and was as follows:

TGGGTGTTTTCACATTCGAGGATGAAACCAACTCCCCTGTGGCTCCTGCTACTCGTTACAAGGCCCTAGTTACAGAT
GCCGACAACGTCATCCCAAAGGCTCTTGATTCCTTCAAGAGTGTTGAAAACGTTGAGGGAAATGGTGGCCCAGGA
ACCATCAAGAAGA

Secondly, it was also intended to silence one candidate gene which had been shown to have an impact during this study. For this reason, candidate gene #1 (Glyma07g37240.2= Glyma.07G243500.2) was chosen. A fragment inside the coding sequence was amplified to clone into the RNAi vector, and to later transform into soybean in attempt to knock down expression of the gene. The region targeted for amplication and cloning is shown in Figure 15. Both of the RNAi fragments were blasted in Phytozome to check that they do not match any other region of the soybean genome to ensure that only the candidate

genes are being targeted. After blasting the first query sequence against the glycine max genome, not only most of the candidate genes from this studied matched but also other pathogenesis related genes (Fig. 16). After blasting with Phytozome 10.2, it was confirmed that, in theory, three out of six of the candidate genes would be knock down using the fragment matching the consensus alignment: candidate gene #1 (Glyma07g37240.2=Glyma.07G243500.2), candidate gene # 3 (Glyma07g37270.1=Glyma.07G243600.1) and candidate gene #4 (Glyma17g03350.1=Glyma.17G030200.1). Additionally, other pathogenesis related genes not included in this study would be knocked out: Glyma.17G03400.1, Glyma.17G030300.1, Glyma.15G145600.1, Glyma15G145900.2 and Glyma.09G040400.1. These (candidate) genes are located in chromosomes 17, 15 and 9 and are also PR-10 genes (Fig. 16 and APPENDIX D). On the other hand, when blasting the fragment belonging to CDS sequence of gene #1 (Glyma07g37240.2= Glyma.07G243500.2), not only gene #1 matches but also another PR-10 gene located adjacent, Glyma.07G243600.1 (Fig. 17 and APPENDIX D).

Transformation vectors containing the RNAi construct will be shipped to Canada to the Laboratory of Daina Simmonds (Agriculture and AgriFood, Ontario). Once there, they will be transformed into soybean using particle bombardment; then the transformed soybean will be shipped to the Clough lab on the University of Illinois campus. Further experiments will be assayed on these plants to analyze the impact of knocking down genes for the effects of exposure to pathogens *Pseudomonas syringae* and *Sclerotinia sclerotiorum*. Additioinally, *Agrobacterium tumefaciens* was transformed with these vectors in order to eventually transform soybean with the in-house transformation method, which is in the process of being developed.

Evaluation of a PR-10 promoter

Leaves from Columbia-0 (negative control), *Arabidopsis thaliana* transformed with Gmubi::GFP/pCAMBIA1300 (positive control) and *Arabidopsis thaliana* transformed with GmPath21::GFP/pCAMBIA1300 (subject of study) were inoculated with *Psuedomonas syringae* and photographed under exposure to UV light every hour after inoculation until the fourth hour with a fluorescent dissecting microscope (Fig. 18, 19 and 20). Results show that soybean GmPath21 PR-10 promoter functioned in Arabidopsis in a pathogen-responsive manner, showing expression of GFP after the first hour of inoculation, getting stronger by the third and fourth hour of the inoculation. The fluorescence was stronger in the petiole, mid and secondary veins, becoming fader to absent in the rest of the leave. This could be attributed to the structure of the vascular system: the ratio of the tertiary and

quaternary veins and veinlets is so small that even though GFP protein might be present it is hard to visualize, or to nature of infiltration of liquid into an Arabidopsis leaf, such that more bacteria were present near veins.

This finding is very important as GmPath21 promoter can be included in the toolbox of available promoters to use in plant biotechnology, which nowadays is narrow, therefore limiting the development of resistant transgenic crops. However, further studies should be conducted, for instance: 1) the regulation of the promoter when the plant is exposed to other pathogens, 2) the expression of GFP by this promoter in other parts of the plants like shoots, stems and roots, 3) the expression of this promoter together with a PR-10 gene to evaluate the resistance to pathogens compared to what was reported here using a constitutive promoter. An ideal situation would be to find that pGmPath21 is not only rapid activated during a wide spectrum of pathogens, but also stays inactive or not expressed during disease-free conditions. This would be a great advantage for candidate genes that might cause a toxic effect for their constitutive expression. Moreover, it is pending the study of cis-regulatory elements for this promoter in order to gain a better understanding of gene regulation and plant signaling during biotic stress conditions. Additionally, the specific external factors that trigger the regulation of pGmPath21 could be determined in future studies in order to deliberately regulate the expression of specific genes in experimental setting and eventually in large agricultural scale.

Concluding remarks and future directions

Bacterial growth monitored in transgenic line #1 (Glyma07g37240.2=Glyma.07G243500.2) and #4 (Glyma17g03350.1= Glyma.17G030200.1) suggest an enhance resistance against both the HR and virulence strains of *Pseudomonas. syringae* during the first 48 hours. However, results suggest that PR-10 is not enhancing resistance to *Sclerotinia sclerotiorum* in Arabidopsis.

Results show that soybean GmPath21 PR-10 promoter functioned in Arabidopsis in a pathogen-responsive manner, showing enhanced expression of GFP after the first hour of inoculation.

Phenotypic findings transgenic line #1 (Glyma07g37240.2=Glyma.07G243500.2): it would be really important to study how the candidate gene causes the phenotypic differences (number of insertions and place of insertion would be a start for the study). Moreover, it would be remarkable to conduct studies with '#1' transgenic to determine if yield and biomass are enhanced compared to wild type if grown under the same conditions.

Even though in this research only *Pseudomonas syringae* and *Sclerotinia sclerotiorum* were studied, it would be of value to challenge the transgenics with additional pathogens.

RNAi construct will be transformed in soybean in both the Clough's Lab and in Laboratory of Daina Simmonds (Agriculture and AgriFood, Ontario) to ascertain their function in soybean defense.

Tables and Figures

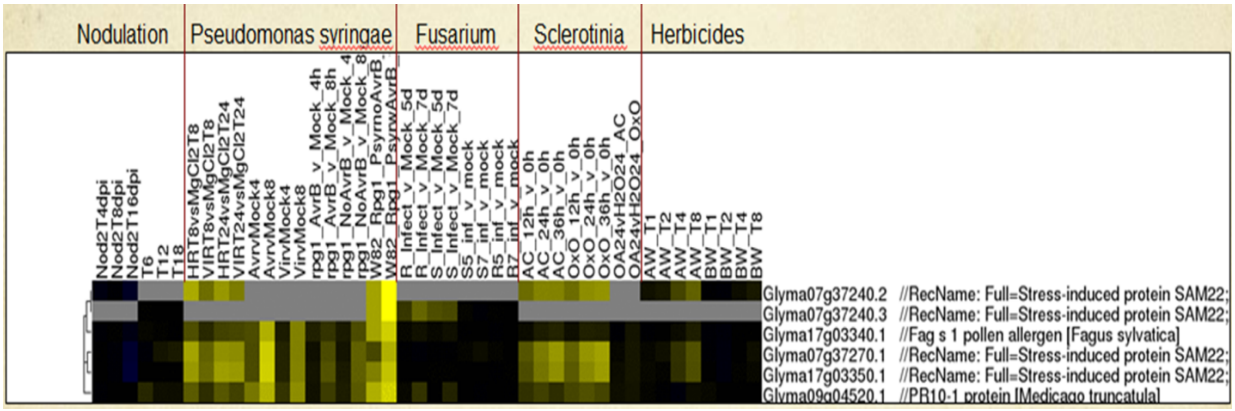

Figure 1. Heat map cluster of differentially expressed PR-10 genes. Yellow indicates increased expression of treatment versus control in each of the 44 treatments.

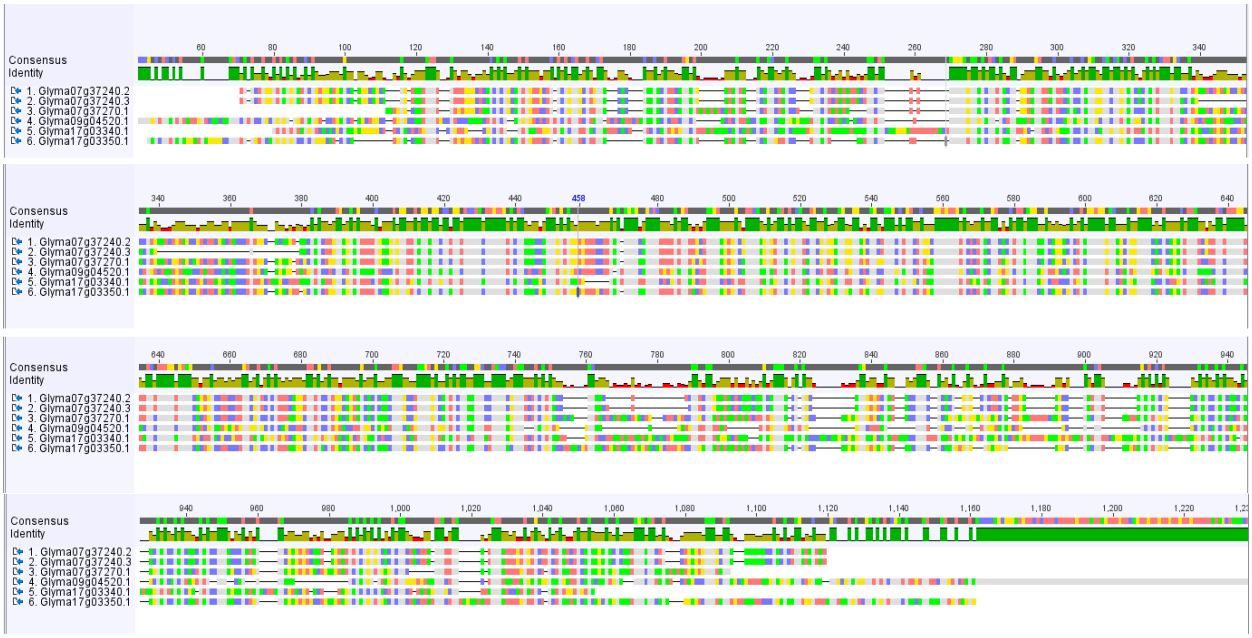

Figure 2. Alignment of six candidate PR-10 genes.

Table 1. Primers designed to clone candidate genes.

Candidate gene	Forward primer	Reverse Primer
#1	ATGGATCCAAAGAAGCACACAGCAGCAA	ATTGTACATCCAAGCAAGAAAAAGCAAGA
#2	ATGGATCCAAAGAAGCACACAGCAGCAA	ATTGTACATCCAAGCAAGAAAAAGCAAGA
#3	ATGGATCCAAATCGAACTCTCGCGTTGT	ATTGTACAAGCAACACACGACAAGAAGG
#4	ATGGATCCCCACACAGCAGCAAACATCT	ATTGTACAGCACACACCACACACAGTGA
#5	ATGGATCCCTAGAGTGACCCCAGGGCTA	ATTGTACACACTCCACACAAGGCAAGAA
#6	ATGGATCCATGCAGGCCAGTGCCTTAT	ATTGTACAAAGAAAGCCTTGGTGCTGAG

The nucleotides in red (AT) are added as "sticky ends". The blue nucleotides correspond to the restriction sites BamHI for the Forward primers and SacI for the Reverse primers. The nucleotides in black correspond to the design for the candidate genes done with Primer 3 software. Restriction sites were added as a backup option in case the T/A cloning failed.

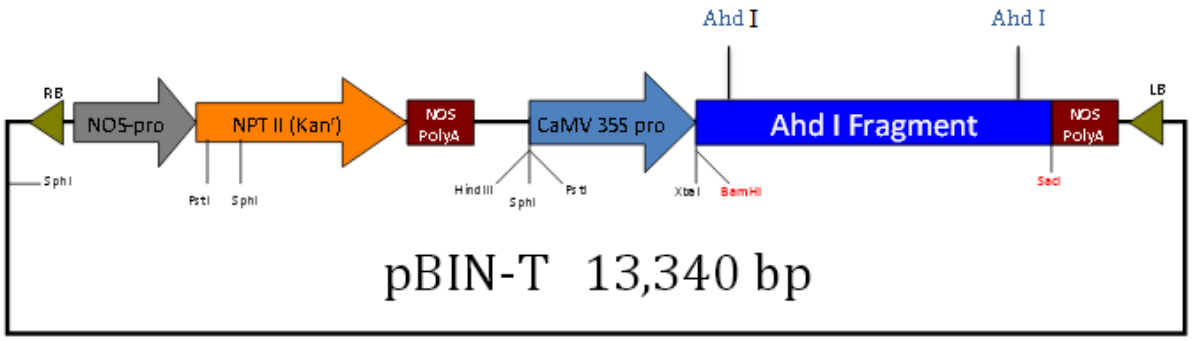

Figure 3. pBIN-T vector with the AhdI fragment before it is cut with AhdI enzyme. (Taken from pBIN-T Cloning Protocol-The Clough Lab: http://clough.cropsci.illinois.edu/pages/protocols.html)

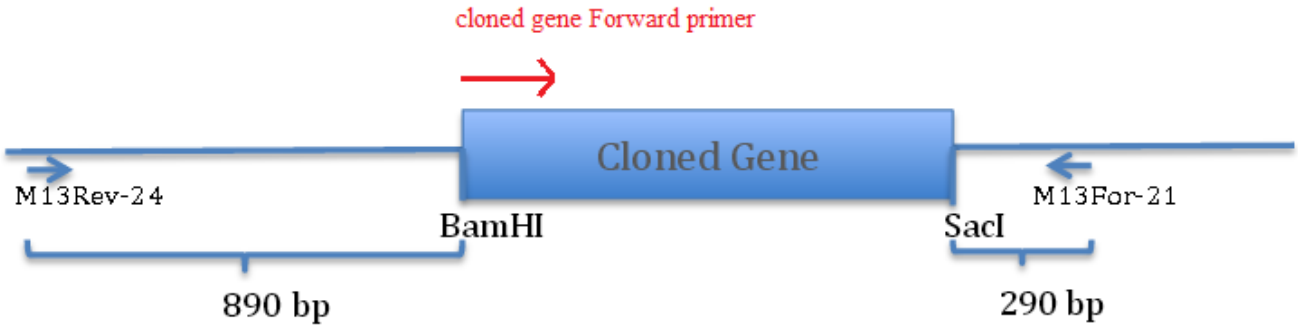

Figure 4. The cloned gene region in the pBIN-T plasmid. The cloned gene Forward primer and the M13 Forward-21 primer are used to check the proper orientation of the insertion. (Taken from pBIN-T Cloning Protocol-The Clough Lab: http://clough.cropsci.illinois.edu/pages/protocols.html)

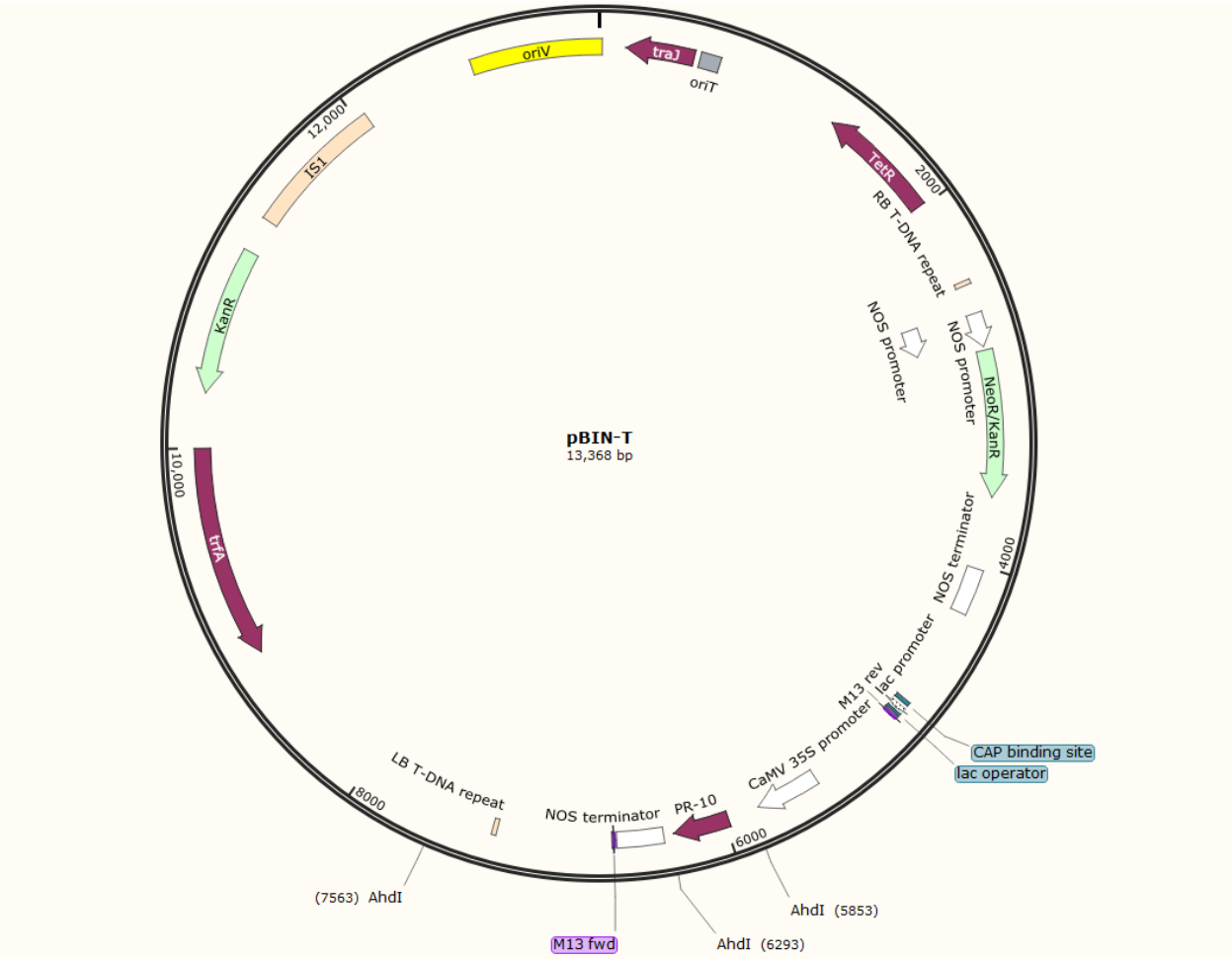

Figure 5. pBIN-T construct showing alignment of a candidate PR-10 gene (Image provided by Dave Neece).

T0 Parental gametes A a

T1 Aa aa

T2

	A	a	
A	AA	Aa	75%
a	aA	aa	

T3

	A	A	
A	AA	AA	100%
A	AA	AA	

	A	a	
A	AA	Aa	75%
a	aA	aa	

Figure 6. How homozygosity of the transgene was determined. T0 generation occurs when Arabidopsis thaliana is transformed. About 1% of ovules will be transgenic, shown here as ´A´. The non transformants are represented with ´a´. The T1 generation can have two options: Aa and aa. Aa will be selected and aa will die. In the T2, three options are possible: AA, aA and Aa (aa will not be selected). Each T2 plant is individually selected and will produce the T3 generation. Depending on which genotype is chosen, the outcome for the T3 can be 100% homozygous (AA) which is the desirable genotype to conduct disease studies, or it can be a mix of homozygous (AA) and heterozygous plants (Aa and aA) which will be identified by only 75% of plants growing. This follows Mendel´s first law of genetics.

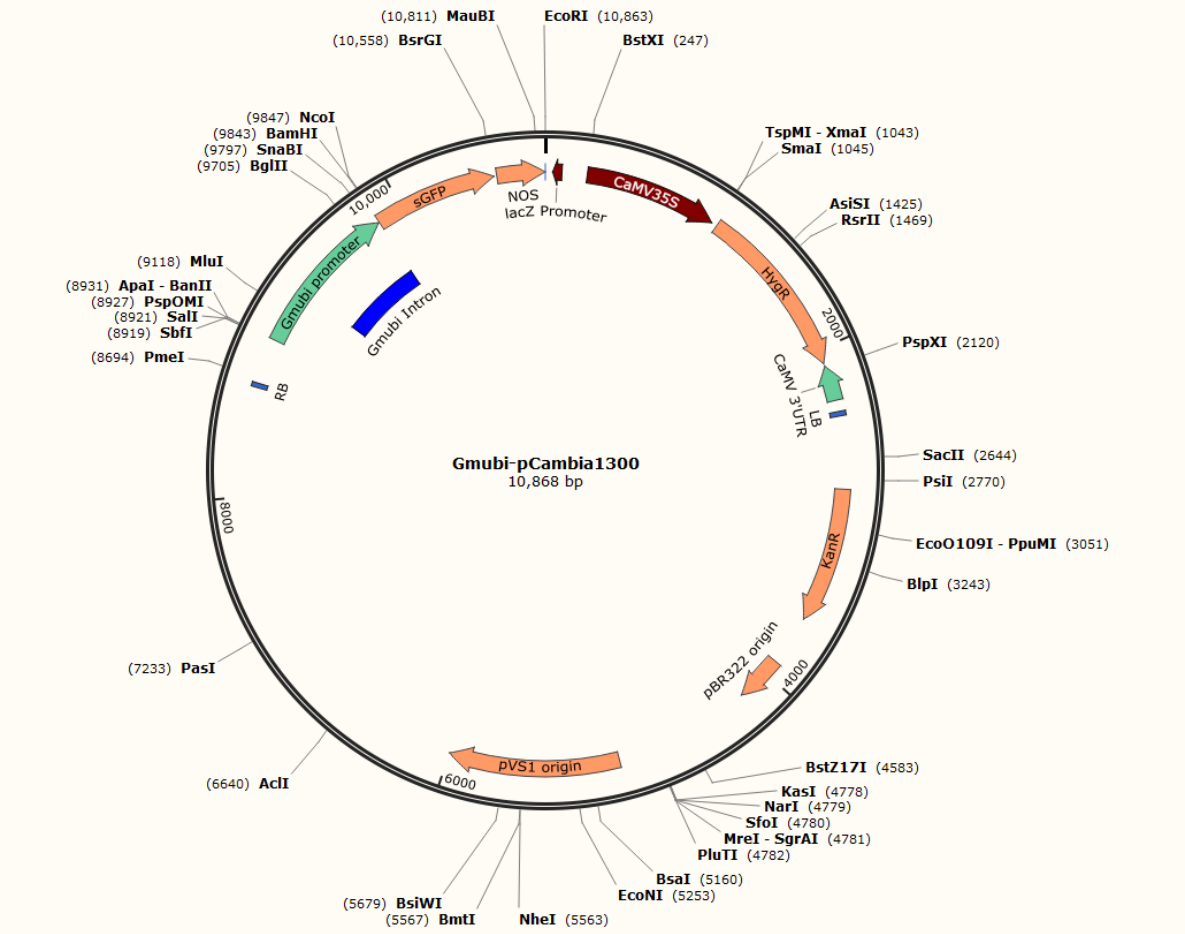

Figure 7. Gmubi-p-Cambia1300. GFP is regulated by the constitutive soybean promoter, Gmubi. This is used as a control and was constructed by the Finer Lab at The Ohio State University.

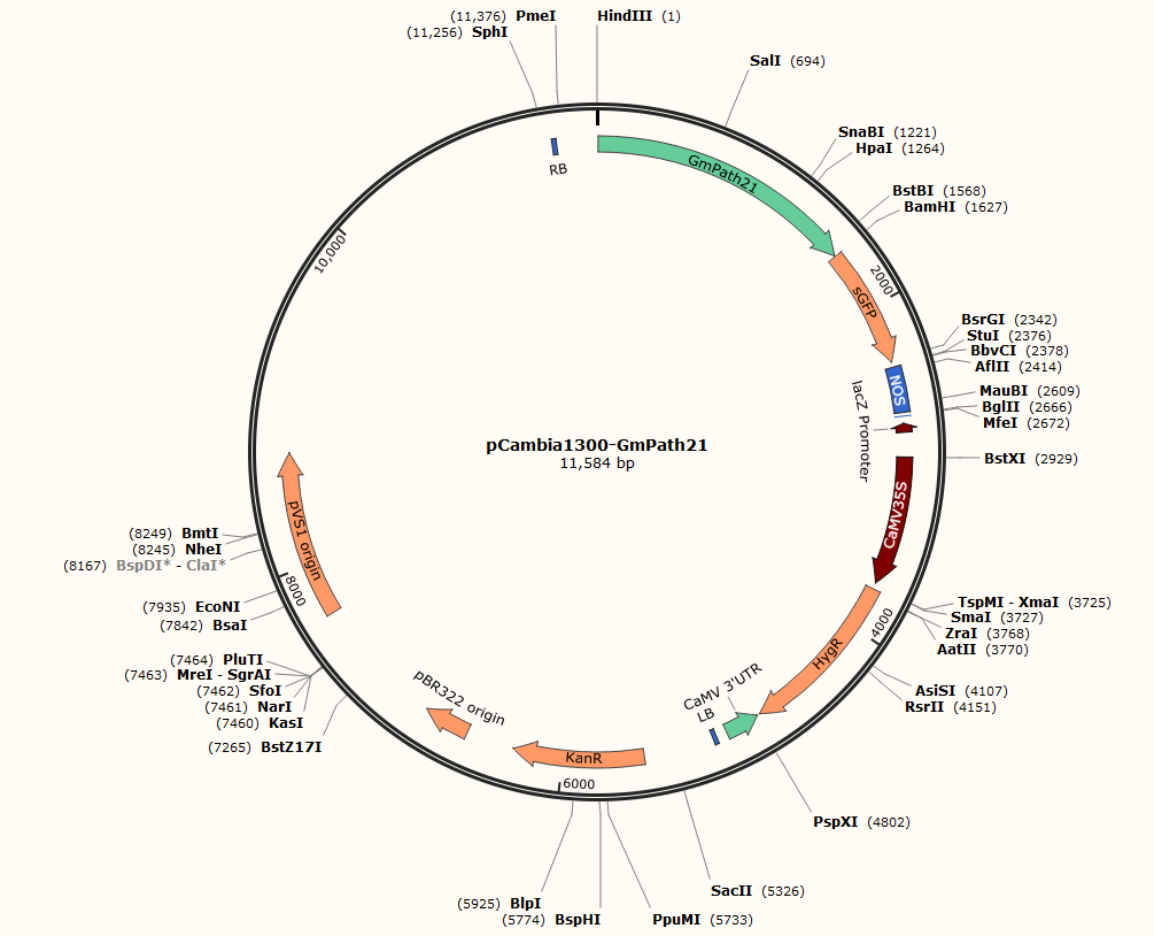

Figure 8. GmPath21-pCambia1300. GFP is regulated by the PR-10 promoter, GmPath21. This was constructed by the Finer Lab at The Ohio State University.

Table 2. Population growth of *Pseudomonas syringae* pv. *tomato* DC3000 with *avrRpt2* (avirulent HR-inducing strain) or without *avrRpt2* (virulent strain) taken from inoculated leaves of Columbia-0, and the transgenic lines, gene #1 and gene #4.

		hours	**Result***	**std error**
Col-0	HR	24	2.8887	0.0600
		48	2.6304	0.0476
		96	2.4728	0.2820
	vir	24	2.5051	0.1288
		48	3.6678	0.1008
		96	5.0391	0.1426
Gene #1	HR	24	2.6064	0.0645
		48	2.3139	0.1628
		96	2.2945	0.2556
	vir	24	1.8921	0.2226
		48	2.9624	0.0753
		96	4.2527	0.2380
Gene #4	HR	24	2.9479	0.0793
		48	2.6405	0.0993
		96	2.8048	0.2313
	vir	24	2.5224	0.1134
		48	3.6275	0.1214
		96	3.7347	0.0909

* results are expressed in log base 10 of the CPU/ul

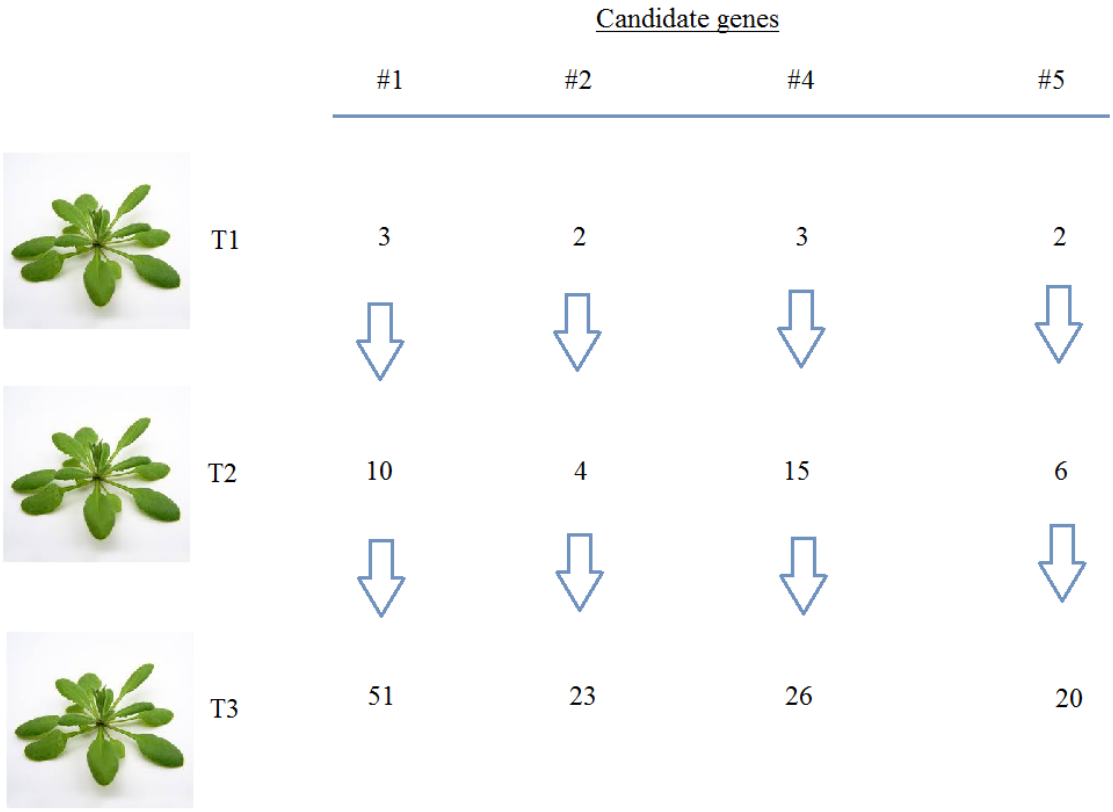

Figure 9. Number of plants that were advanced to the next generation are displayed for each generation and each candidate gene.

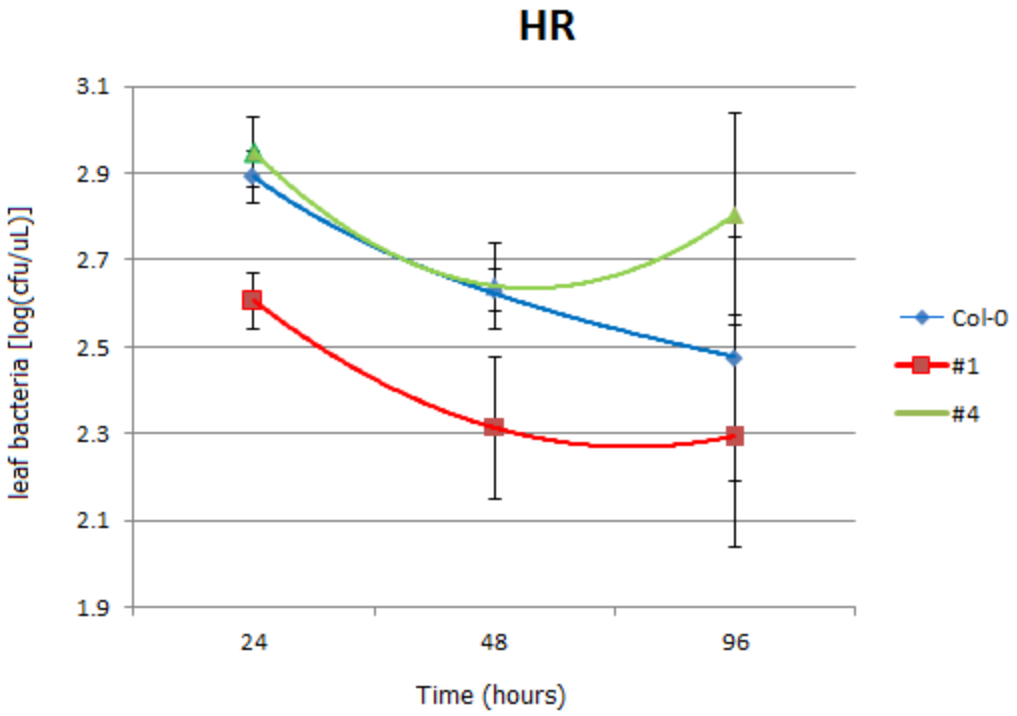

Figure 10. Population growth of *Pseudomonas syringae* pv. *tomato* DC3000 carrying *avrRpt2* (avirulent HR-inducing strain) taken from inoculated leaves of in Arabidopsis Col-0 and transgenics lines #1 and #4.

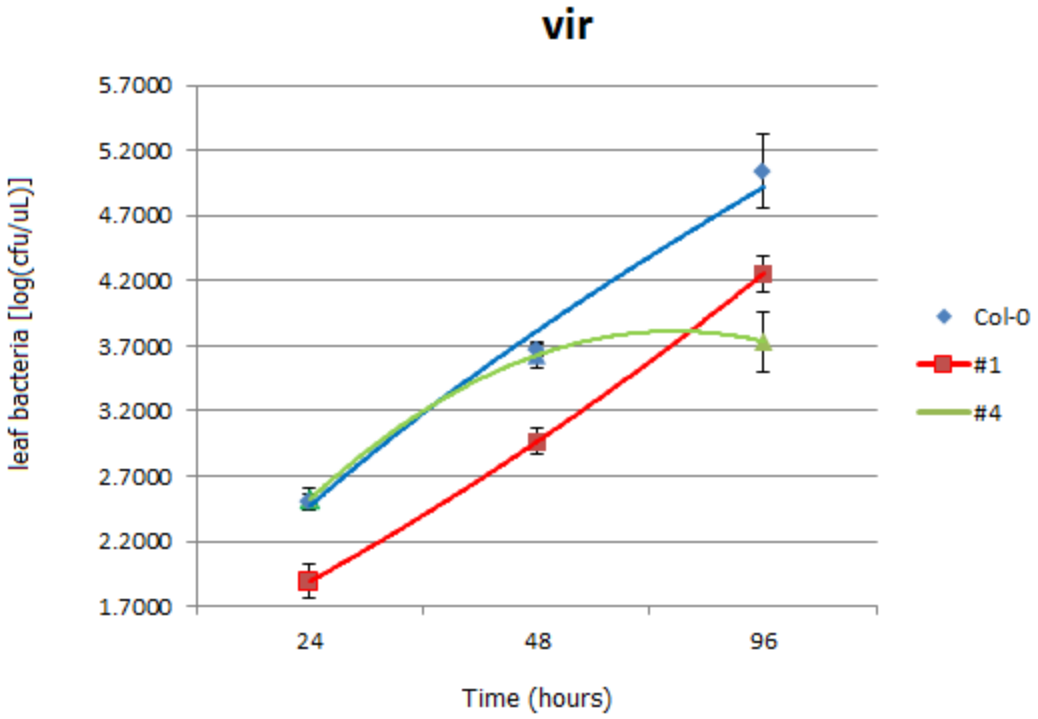

Figure 11. Population growth of *Pseudomonas syringae* pv. *tomato* DC3000 (virulent strain) taken from inoculated leaves of Arabidopsis Col-0 and transgenics lines #1 and #4.

Table 3. Area of lesion in Columbia and transgenic lines caused by *Sclerotinia sclerotiorum* inoculation.

	Columbia-0		Gene #1		Gene #4	
	Average*	Std error	Average*	Std error	Average*	Std error
24 hs	0.51	0.02	0.46	0.03	0.62	0.03
36 hs	1.30	0.04	1.08	0.08	1.31	0.06
48 hs	2.01	0.07	1.74	0.10	2.29	0.09

* Area of lesion is measured in mm^2.

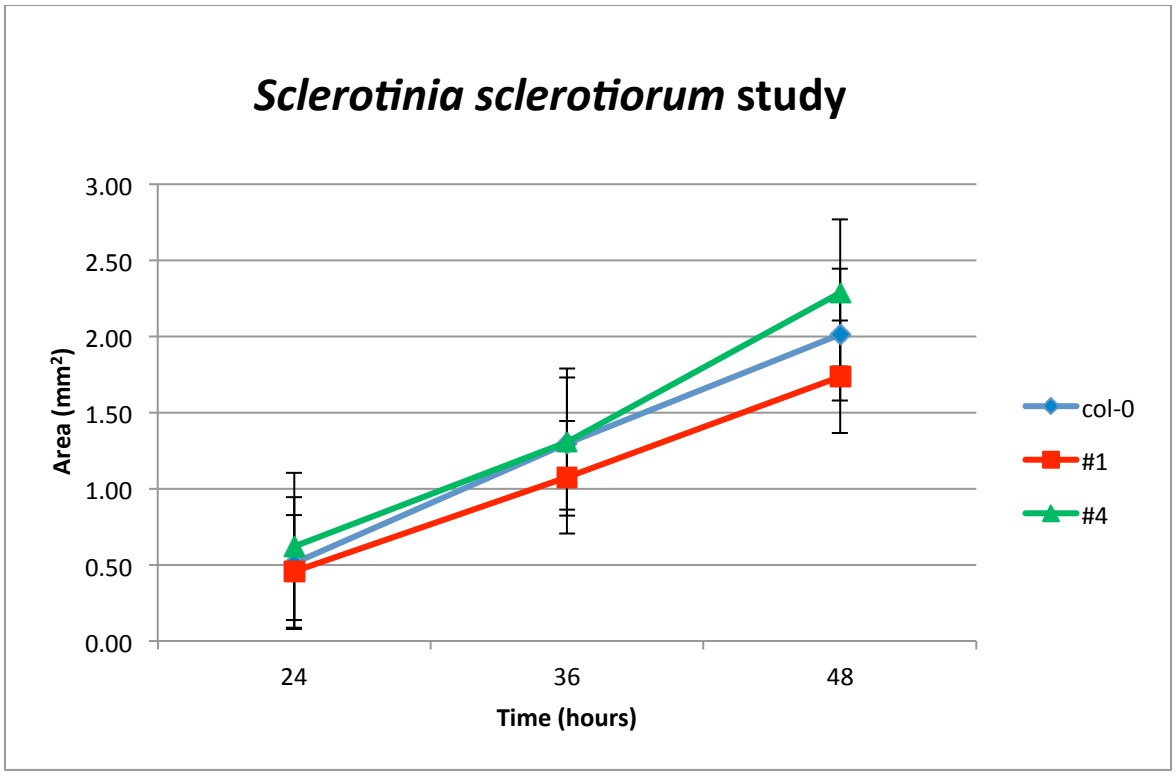

Figure 12. Area of lesion for Columbia-0 and transgenic lines

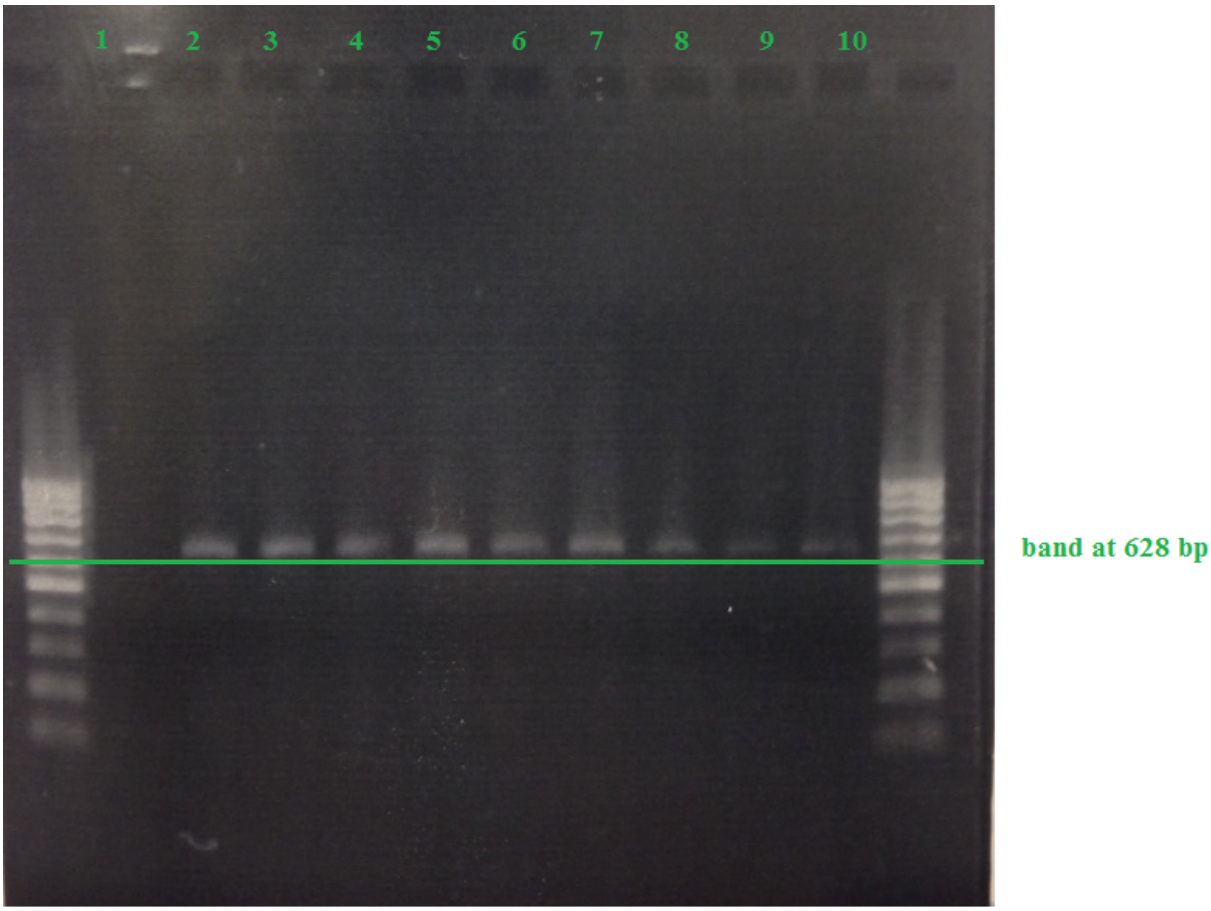

Figure 13. Semi-quantitative qRT-PCR gel showing the presence of candidate gene #1 in the transgenic plants from which RNA was extracted.

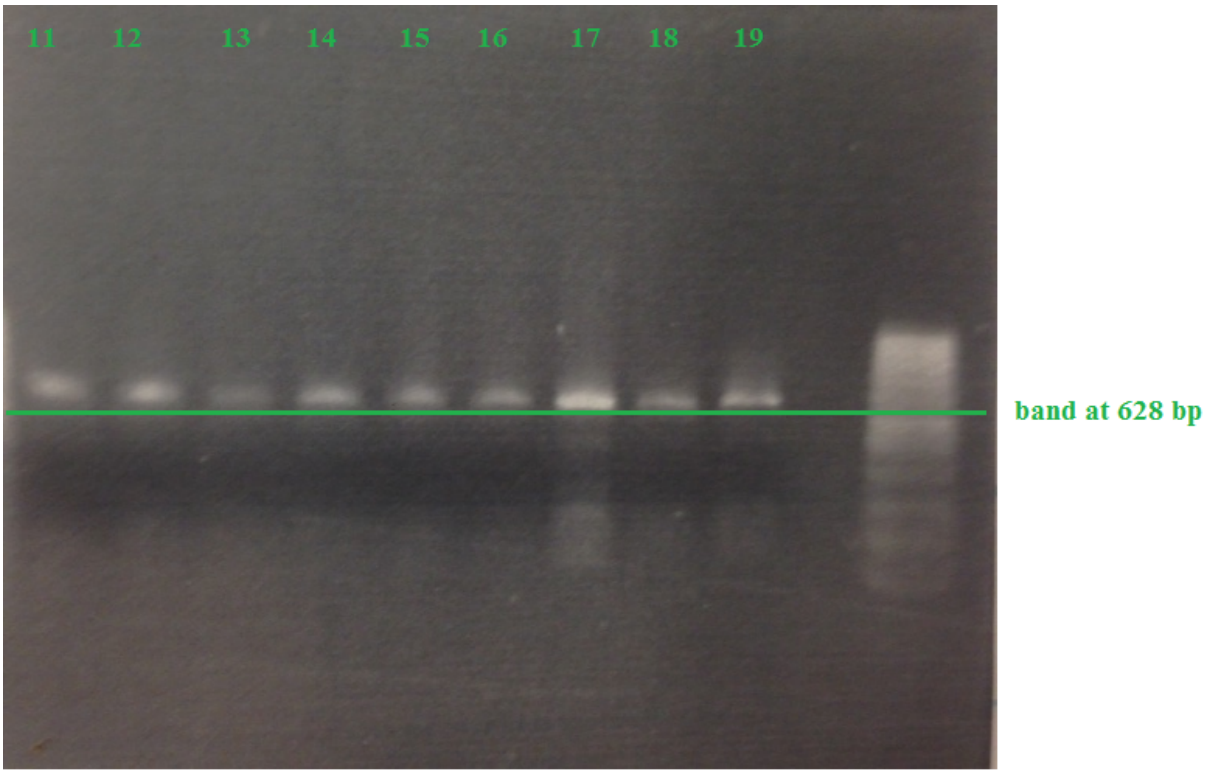

Figure 13 (cont.).

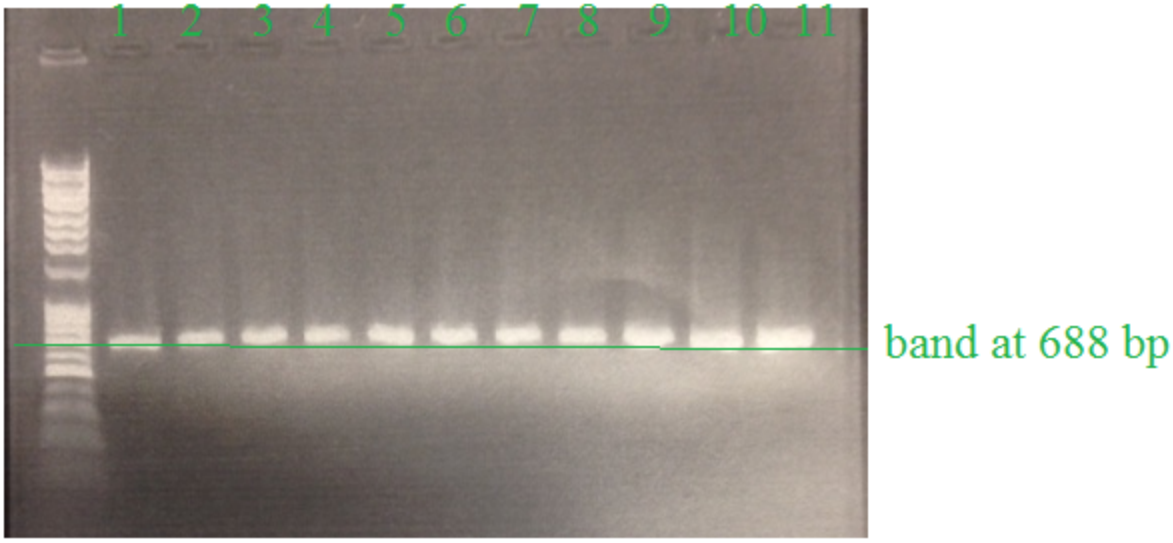

Figure 13 (cont.)

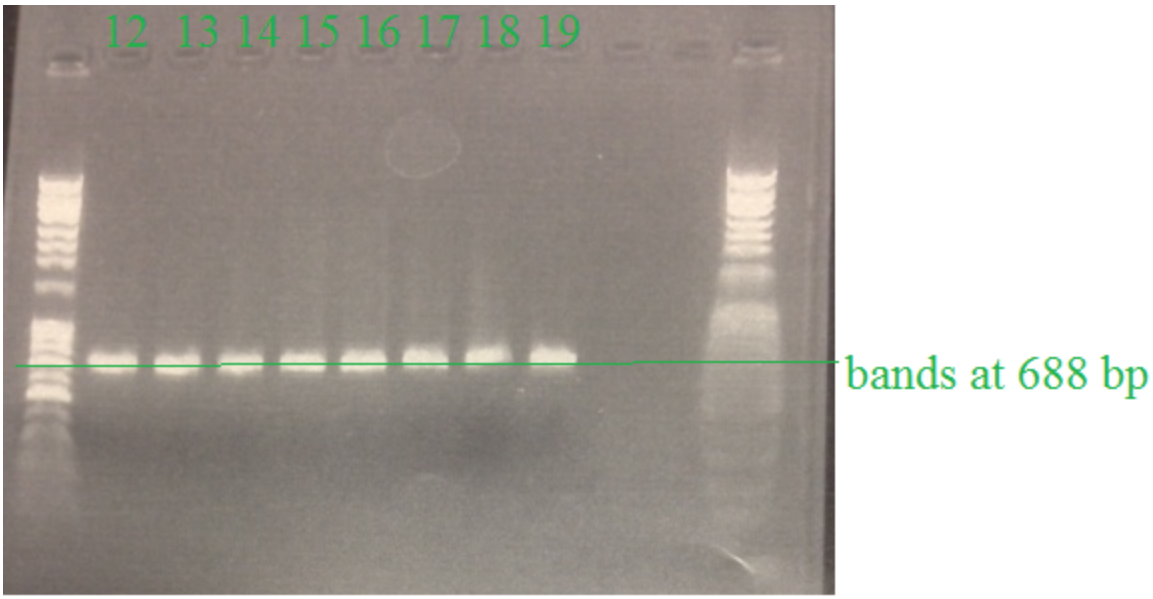

Figure 13 (cont.)

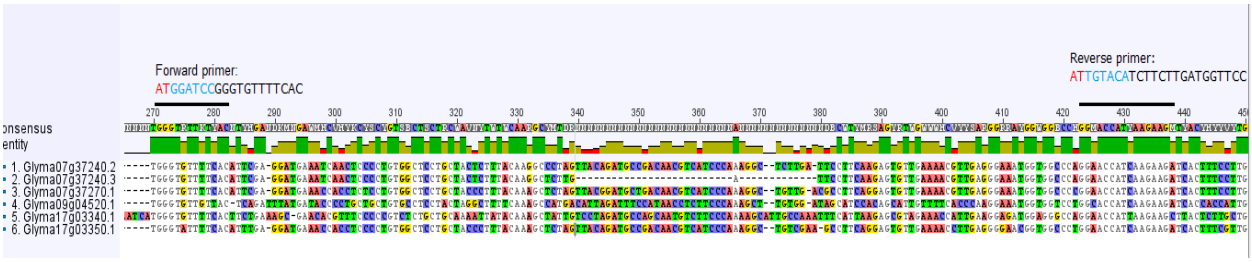

Figure 14. Alignment of the six candidate PR-10 genes to identified the most conserved region among all six. The primers designed are indicated above of the figure. The primers were designed using Primer3[@] software.

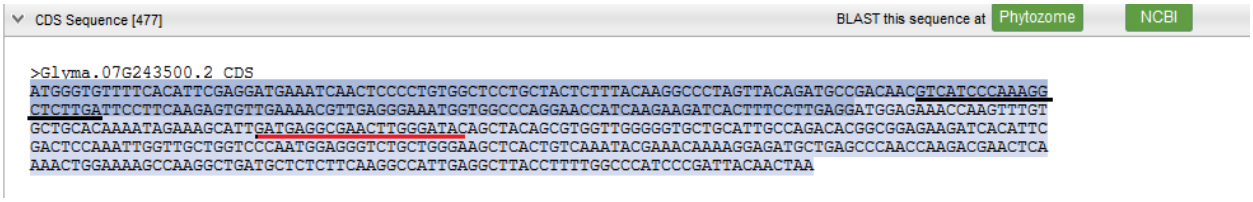

Figure 15. CDS sequence of candidate gene #1 (Glyma07g37240.2= Glyma.07G243500.2) obtained from Phytozome v10.2, and the region targeted for use in RNAi. The sequence matching the Forward primer is indicated in black, the sequence matching the Reverse primer is indicated in red.

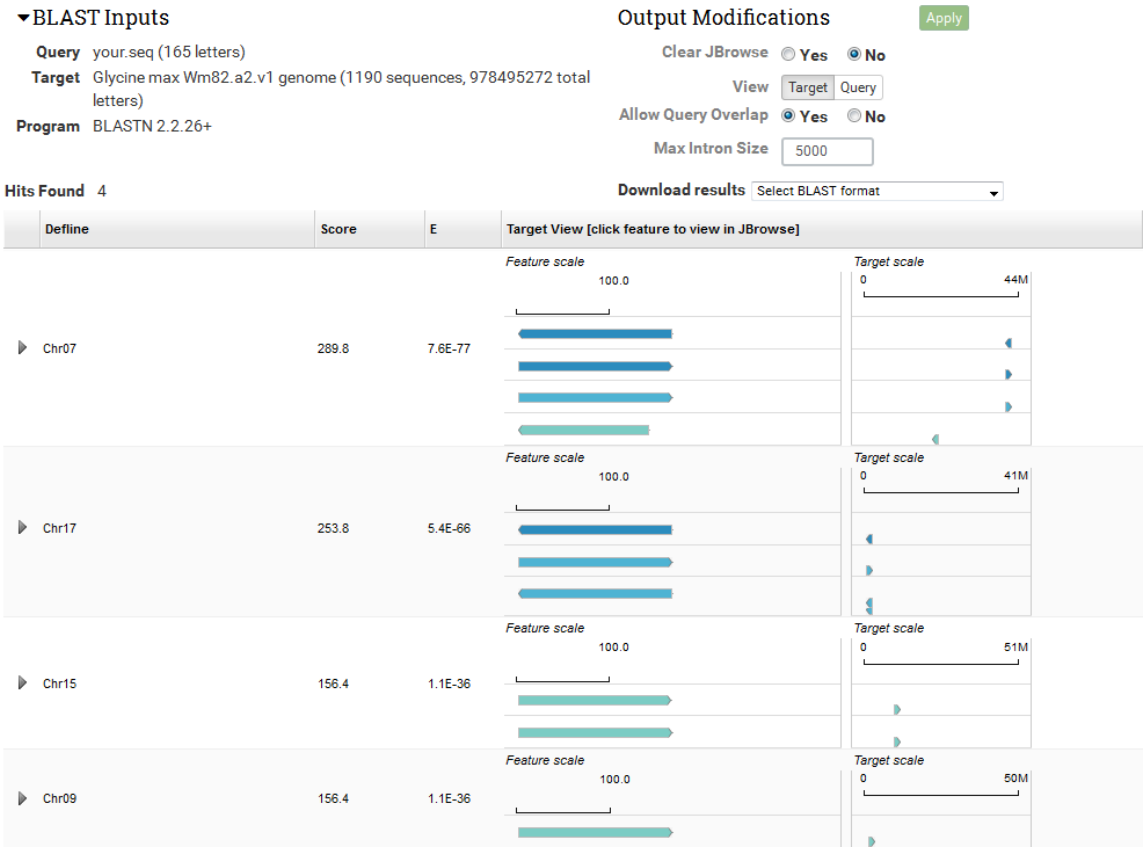

Figure 16. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for RNAi against multiple PR-10 genes.

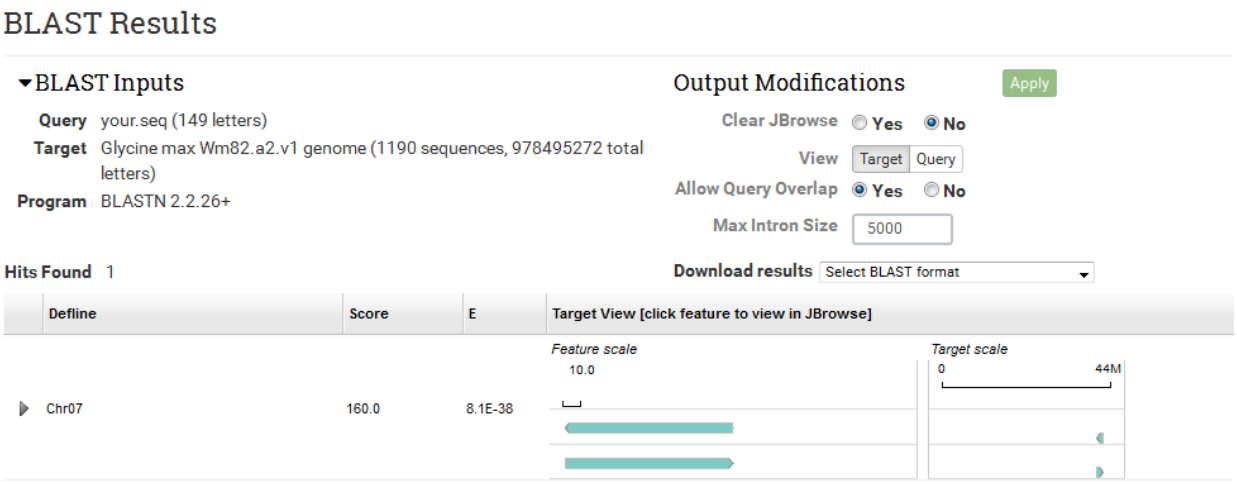

Figure 17. BLAST search matching fragment belonging to the CDS sequence for gene #1 used in RNAi

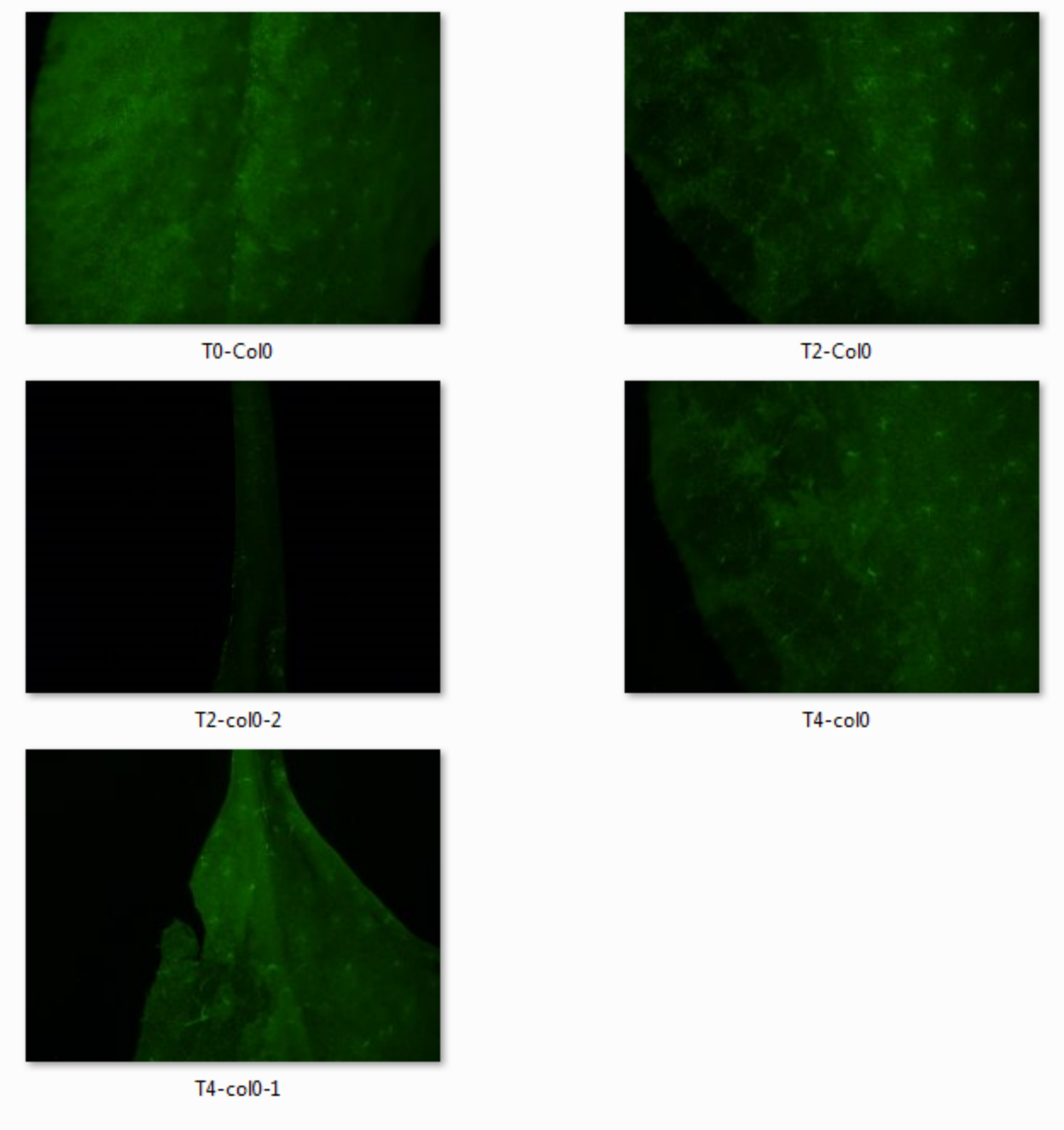

Figure 18. Fluorescent images of negative control, Columbia-0, before inoculation (T0), after 2 hours of inoculation (T2) and after 4 hours of inoculation (T4).

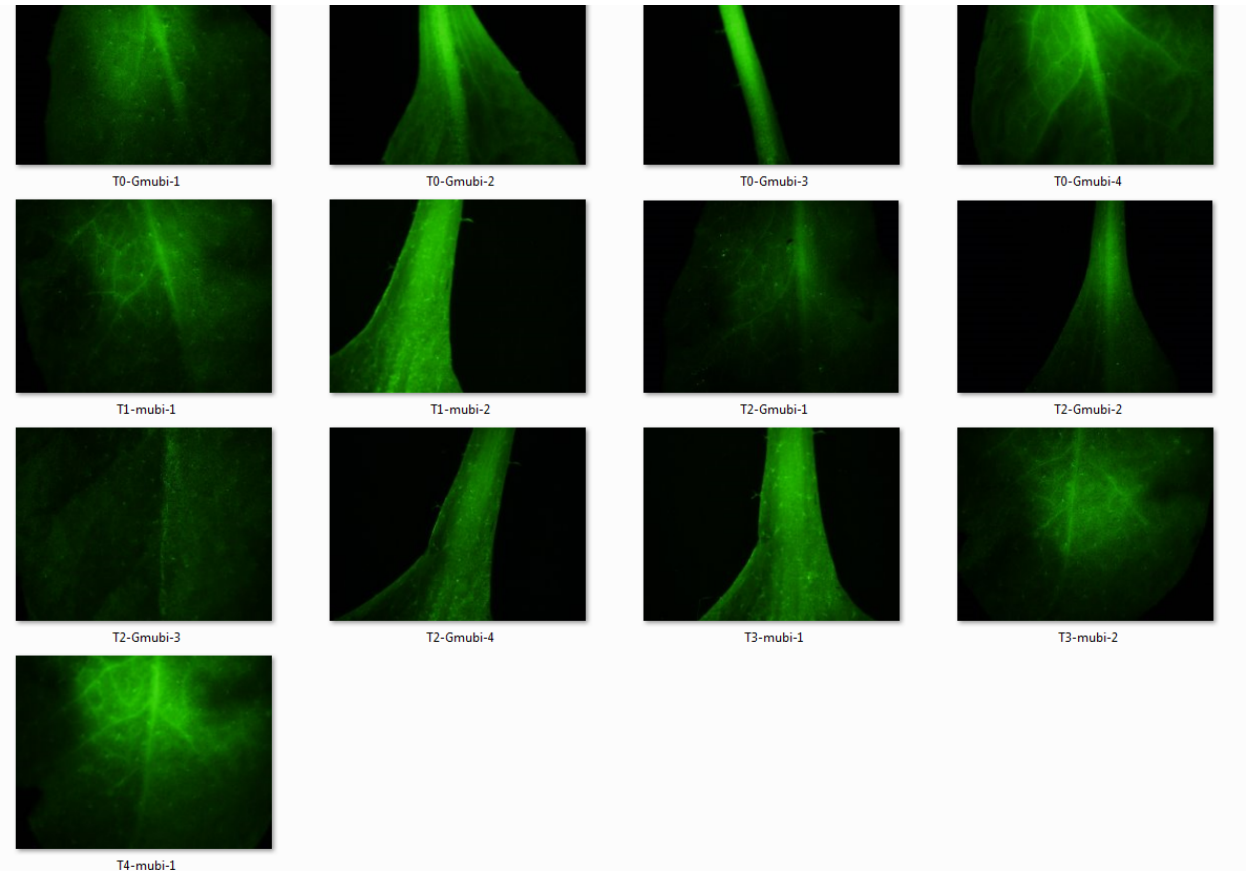

Figure 19. Florescent images of positive control, *Arabidopsis thaliana* containing the construct Gmubi::GFP/pCAMBIA1300. Pictures were taken before the inoculation (T0), after 1 hour (T1), 2 hours (T2), 3 hours (T3) and 4 hours (T4) of the inoculation.

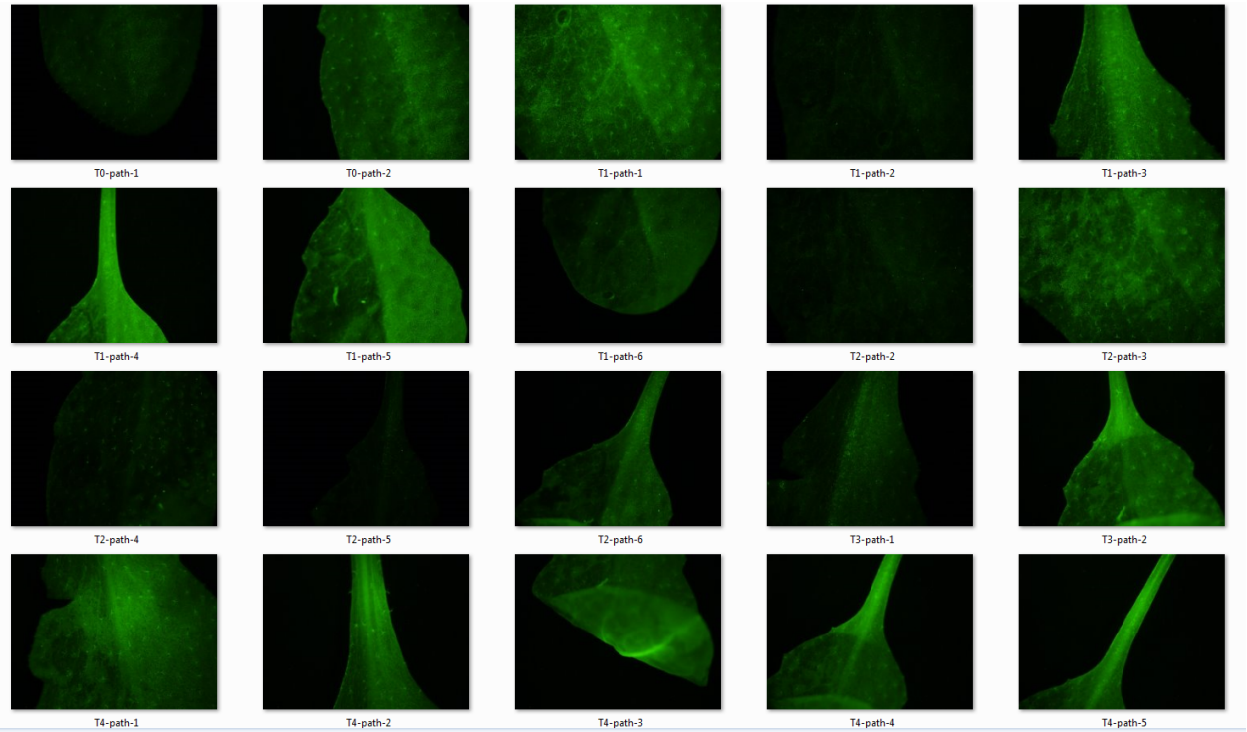

Figure 20. Fluorescent Iimages of *Arabidopsis thaliana* containing the construct GmPath21::GFP/pCAMBIA1300. Pictures were taken before the inoculation (T0), and after 1 hour (T1), 2 hours (T2), 3 hours (T3) and 4 hours (T4) of the inoculation.

REFERENCES

Albrecht B., Spangfort M. D., Kahlert H., Schlaak M., Becker W.-M. 1996. The major birch pollen allergen, *Bet v 1*, shows ribonuclease activity. Planta 199: 413-415.

Alcoba N.J., Bejarano N., Catacata J.R. 2005. Sistema Nacional Argentino de Vigilancia y Monitoreo de plagas [Online]. Available at http://www.sinavimo.gov.ar/plaga/sclerotinia-sclerotiorum (accessed 3 June 2015; verified 25 July 2015). SNAVMP. CABA

Arumuganathan, K., and E. D. Earle. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. 9**:** 208–219.

Berge, O., Monteil, C. L., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, D. C., & Morris, C. E. 2014. A User's Guide to a Data Base of the Diversity of *Pseudomonas syringae* and Its Application to Classifying Strains in This Phylogenetic Complex. *PLoS ONE*, *9*(9), e105547. doi:10.1371/journal.pone.0105547

Breiteneder H., Pettenburger K., Bito A., Valenta R., Kraft D., Rumpold H., Scheiner O., Breintebach M. 1989. The gene coding for the major birch pollen allergen *Bet v 1*, is highly homologous to a pea disease resistance response gene. EMBO J 8: 1935-1938

Calla B, Blahut-Beatty L, Koziol L, Zhang Y, Neece DJ, Carbajulca D, Garcia A, Simmonds DH, Clough SJ. 2014. Genomic evaluation of oxalate-degrading transgenic soybean in response to Sclerotinia sclerotiorum infection. Mol Plant Pathol. 15(6):563-75. doi: 10.1111/mpp.12115.

Choi, D. S., Hwang, I. S., & Hwang, B. K. 2012. Requirement of the Cytosolic Interaction between PATHOGENESIS-RELATED PROTEIN10 and LEUCINE-RICH REPEAT PROTEIN1 for Cell Death and Defense Signaling in Pepper. *The Plant Cell*, *24*(4), 1675–1690. doi:10.1105/tpc.112.095869.

Chung, W.-H., Jeong, N., Kim, J., Lee, W. K., Lee, Y.-G., Lee, S.-H., Jeong, S.-C. 2014. Population Structure and Domestication Revealed by High-Depth Resequencing of Korean Cultivated and Wild Soybean Genomes. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes*, *21*(2), 153–167. doi:10.1093/dnares/dst047.

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

Collmer, A., Badel, J. L., Charkowski, A. O., Deng, W.-L., Fouts, D. E., Ramos, A. R., Amos H. Rehm, Anderson D. M., Schneewind O., van Dijk K., and Alfano J. R. 2000. *Pseudomonas syringae* Hrp type III secretion system and effector proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(16), 8770–8777.

Desfeux, C., Clough, S. J., & Bent, A. F. 2000. Female Reproductive Tissues Are the Primary Target of *Agrobacterium*-Mediated Transformation by the Arabidopsis Floral-Dip Method. *Plant Physiology*, *123*(3), 895–904.

Dickinson M. 2000. Molecular Plant Pathology. John Wiley and Sons Ltd. Oxford, UK.

Dring N. C., Maliyakal E. J., Russell D and Amasino R. M. 1992. Characterization of a stress-induced, developmentally regulated gene family from soybean. Plant Molecular Biology. 18: 459-466.

Engl, C., Waite, C. J., McKenna, J. F., Bennett, M. H., Hamann, T., & Buck, M. 2014. Chp8, a Diguanylate Cyclase from *Pseudomonas syringae* pv. Tomato DC3000, Suppresses the Pathogen-Associated Molecular Pattern Flagellin, Increases Extracellular Polysaccharides, and Promotes Plant Immune Evasion. *mBio*, *5*(3), e01168–14. doi:10.1128/mBio.01168-14

Fernandes H., Michalska K., Sikorski M., Jaskolski M. 2012. Structural and functional aspects of PR-10 proteins. The FEBS Journal. 280:1169-1199. doi:10.1111/febs.12114.

Gyoung-Ah Lee, Gary W. Crawford, Li Liu, Yuka Sasaki, Xuexiang Chen. 2011. Archaeological Soybean (Glycine max) in East Asia: Does Size Matter?. PLoS ONE 6(11): e26720. doi: 10.1371/ journal.pone.0026720.

Hartman G., West E., Herman T. 2011. Crops that feed the World 2. Soybean—worldwide production, use, and constraints caused by pathogens and pests. Food Sec. 3:5–17. Doi: 10.1007/s12571-010-0108-x.

Hernandez-Garcia C. M., Bouchard, R. A., Rushton, P. J., Jones, M. L., Chen, X., Timko, M. P., & Finer, J. J. 2010. High level transgenic expression of soybean (*Glycine max*) GmERF and Gmubi gene promoters isolated by a novel promoter analysis pipeline. *BMC Plant Biology*, *10*, 237. doi:10.1186/1471-2229-10-237.

Hernandez-Garcia C. M., Finer J. J. 2014. Identification and validation of promoters and cis-acting

regulatory elements. Plant Sci. 2014 March; 217-218: 109–119. doi: 10.1016/j.plantsci.2013.12.007

Hinchee M.A.W., Connor-Ward D.V., Newell C.A., McDonnell R.E., Sato S.J., Gasser C.S., Fischhoff D.A., Re D.B., Fraley R.T., Horsch R.B. 1988. Production of transgenic soybean plants using *Agrobacterium*-mediated DNA transfer. Nat Biotech. 6:915–922. doi: 10.1038/nbt0888-915.

Kim, M. Y., Lee, S., Van, K., Kim, T.-H., Jeong, S.-C., Choi, I.-Y., … Lee, S.-H. 2010. Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. Proceedings of the National Academy of Sciences of the United States of America, 107(51), 22032–22037. doi:10.1073/pnas.1009526107

Li M., Liang X., Rollins J.A. 2012. *Sclerotinia sclerotiorum* γ-Glutamyl Transpeptidase (Ss-Ggt1) Is Required for Regulating Glutathione Accumulation and Development of Sclerotia and Compound Appressoria. APS. 25(3): 412-420. doi: 10.1094/MPMI-06-11-0159.

Libault M., Joshi T., Takahashi K., Hurley-Sommer A., Puricelli K., Blake S., Finger R.E., Taylor C.G., Xu D., Nguyen H.T., Stacey G. 2009. Large-Scale Analysis of Putative Soybean Regulatory Gene Expression Identifies a *Myb* Gene Involved in Soybean Nodule Development. *Plant Physiology*, *151*(3), 1207–1220. doi:10.1104/pp.109.144030

Neece, D. 2013. UIUC - Clough Lab. 2014. The Clough Lab website [Online]. Available at http://clough.cropsci.illinois.edu/pages/protocols.html (accessed 15 March 2014; verified 25 July 2015). UIUC, Urbana, IL.

Neece, D. 2014. UIUC - Clough Lab. 2014. The Clough Lab website [Online]. Available at http://clough.cropsci.illinois.edu/protocols/misc/pBIN-T_Cloning_Protocol.pdf (accessed 15 March 2014; verified 7 July 2015). UIUC, Urbana, IL.

Park, C.-J., Kim, K.-J., Shin, R., Park, J. M., Shin, Y.-C. and Paek, K.-H. 2004. Pathogenesis-related protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. The Plant Journal, 37: 186–198. doi: 10.1046/j.1365-313X.2003.01951.x.

Phipps P., Koenning S., Rideout S., Stromberg E., Bush E. 2010. Common diseases of soybean in the Mid-Atlantic Region [Online]. Available at: https://pubs.ext.vt.edu/3001/3001-1435/3001-1435.html Virginia State University-Virginia Cooperative Extension, Blacksburg, VA.

Radwan, O., Li M., Calla B., Li S., Hartmanand G., Clough S. J. 2012. Effect of *Fusarium virguliforme* phytotoxin on soybean gene expression suggests a role in multidimensional defense. Molecular Plant Pathology. 14(3):293-307. DOI: 10.1111/mpp.12006.

Resendis-Antonio, O., Hernández, M., Salazar, E., Contreras, S., Batallar, G. M., Mora, Y., & Encarnación, S. 2011. Systems biology of bacterial nitrogen fixation: High-throughput technology and its integrative description with constraint-based modeling. *BMC Systems Biology*, *5*, 120. doi:10.1186/1752-0509-5-120

Shurtleff W. and Aoyagi A. 2004. History of World Soybean Production and Trade - Part 1 from the Unpublished Manuscript, History of Soybeans and Soy foods, 1100 B.C. to the 1980s [Online]. Available at: http://www.soyinfocenter.com/HSS/production_and_trade1.php (accessed 5 June 2015; verified 25 July 2015). Soyfoods Center, Lafayette, Ca.

Spoeland S., Dong X. 2012. How do plants achieve immunity? Defense without specialized immune cell. Nature Review. Inmunology. Vol 12, p89-100. doi:10.1038/nri3141. Sussman M., Altmann T., Crosby B. 2013. The Multinational Coordinated Arabidopsis thaliana Functional Genomics Project [Online]. Available at: http://www.nsf.gov/pubs/2002/bio0202/model.htm (accessed 4 June 2015; verified 25 July 2015). The National Science Foundation. Arlington, VA.

Tariq M. Rana. 2007. Illuminating the silence: understanding the structure and function of small RNAs. Nature Reviews Molecular Cell Biology 8, 23-36. doi:10.1038/nrm2085.

Tao Y., Xie Z., Chen W., Glazebrook J., Chang H.S., Han B., Zhu T., Zou G., Katagiri F. 2003. Quantitative Nature of Arabidopsis Responses during Compatible and Incompatible Interactions with the Bacterial Pathogen *Pseudomonas syringae.* The Plant Cell. 15:317-330.

Upadhyay P, Rai A, Kumar R, Singh M, Sinha B. 2014. Differential Expression of Pathogenesis Related Protein Genes in Tomato during Inoculation with *A. Solani*. J Plant Pathol Microb 5:217. doi: 10.4172/2157-7471.1000217.

van Loon L.C., Rep M., Pieterse C.M.J. 2006. Significance of Inducible Defense-related Proteins in Infected Plants. Annual Review of Phytopathology. 44: 135-162. Doi: 10.1146/annurev.phyto.44.070505.143425.

Wanderley-Nogueira, A. C., Belarmino, L. C., Soares-Cavalcanti, N. da M., Bezerra-Neto, J. P., Kido, E. A., Pandolfi, V., Benko-Iseppon, A. M. 2012. An overall evaluation of the Resistance (*R*) and Pathogenesis-

Related (*PR*) superfamilies in soybean, as compared with *Medicago* and *Arabidopsis*. *Genetics and Molecular Biology*, *35*(1 Suppl), 260–271. doi:10.1590/S1415-47572012000200007.

Yu, I-c., K. A. Fengler, S. J. Clough, and A. F. Bent. 2000. Identification of Arabidopsis mutants exhibiting an altered hypersensitive response in gene-for-gene disease resistance. Mol. Plant Microbe Interact. 13:277-286.

Zhou X., Lu S., Xu Y., Wang J., Chen X. 2002. A cotton cDNA (*GaPR-10*) encoding a pathogenesis-related 10 protein with in vitro ribonuclease activity. Plant Science. 162(4):629–636. Doi: 10.1016/S0168-9452(02)00002-X.

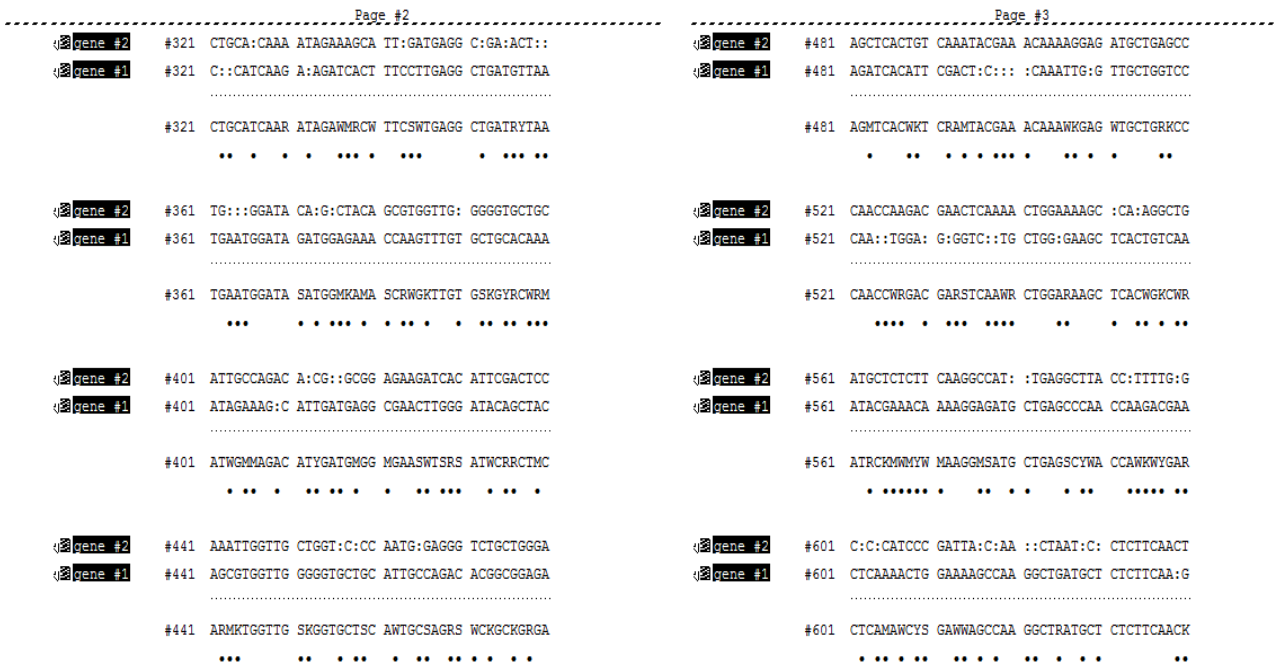

Figure A1. Assembly of candidate gene #1 and #6 (gene #1 is in the top lines, gene #6 is in the bottom lines)

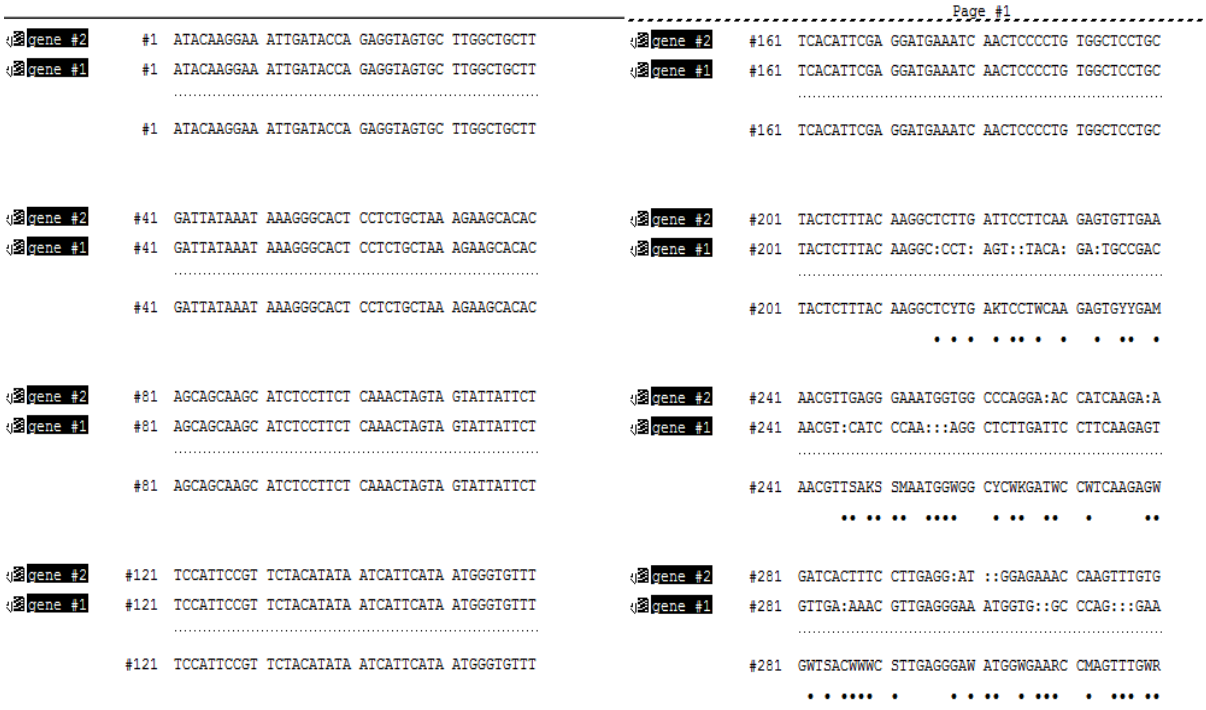

Figure A1 (cont).

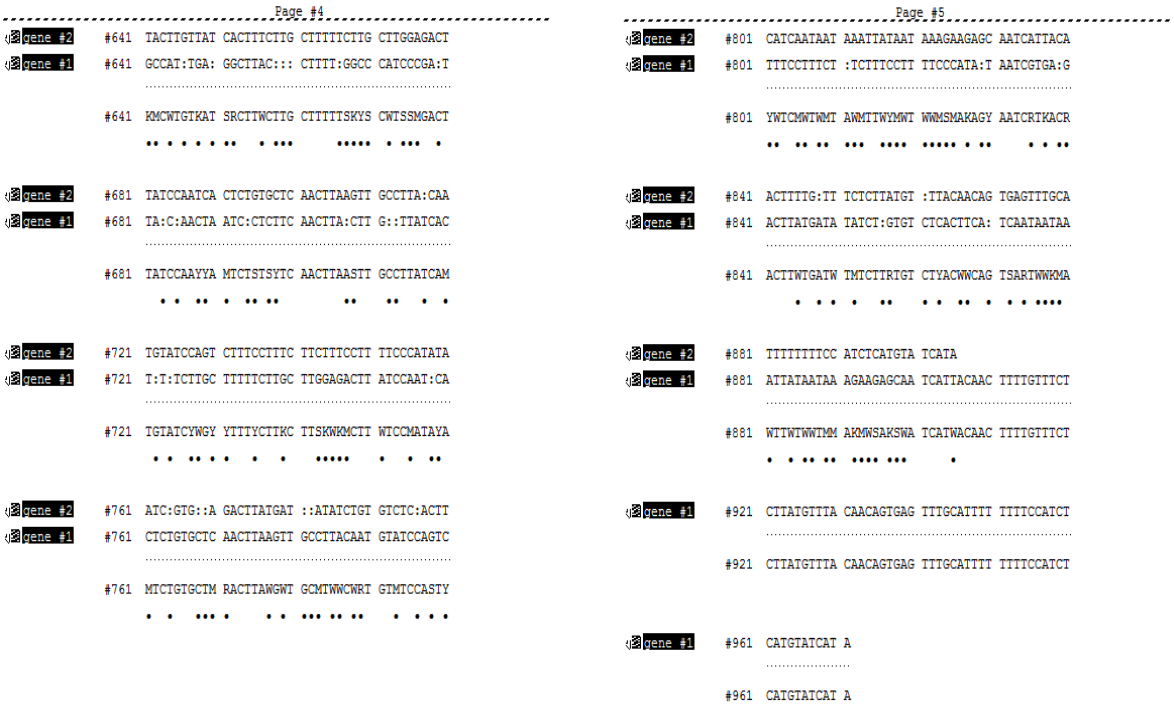

Figure A1 (cont).

APPENDIX B

Figure B1. Columbia-0, 24 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

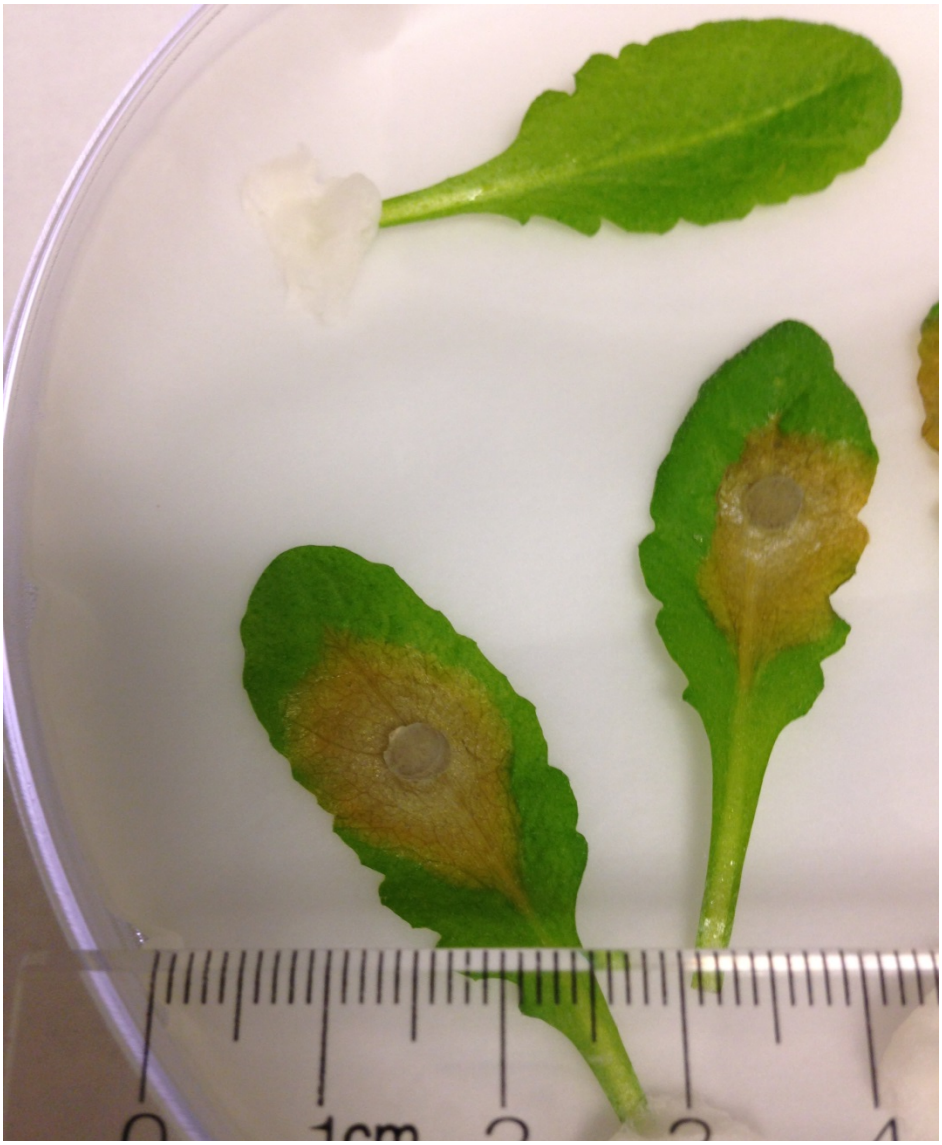

Figure B2. Columbia-0, 36 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

Figure B3. Columbia-0, 48 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

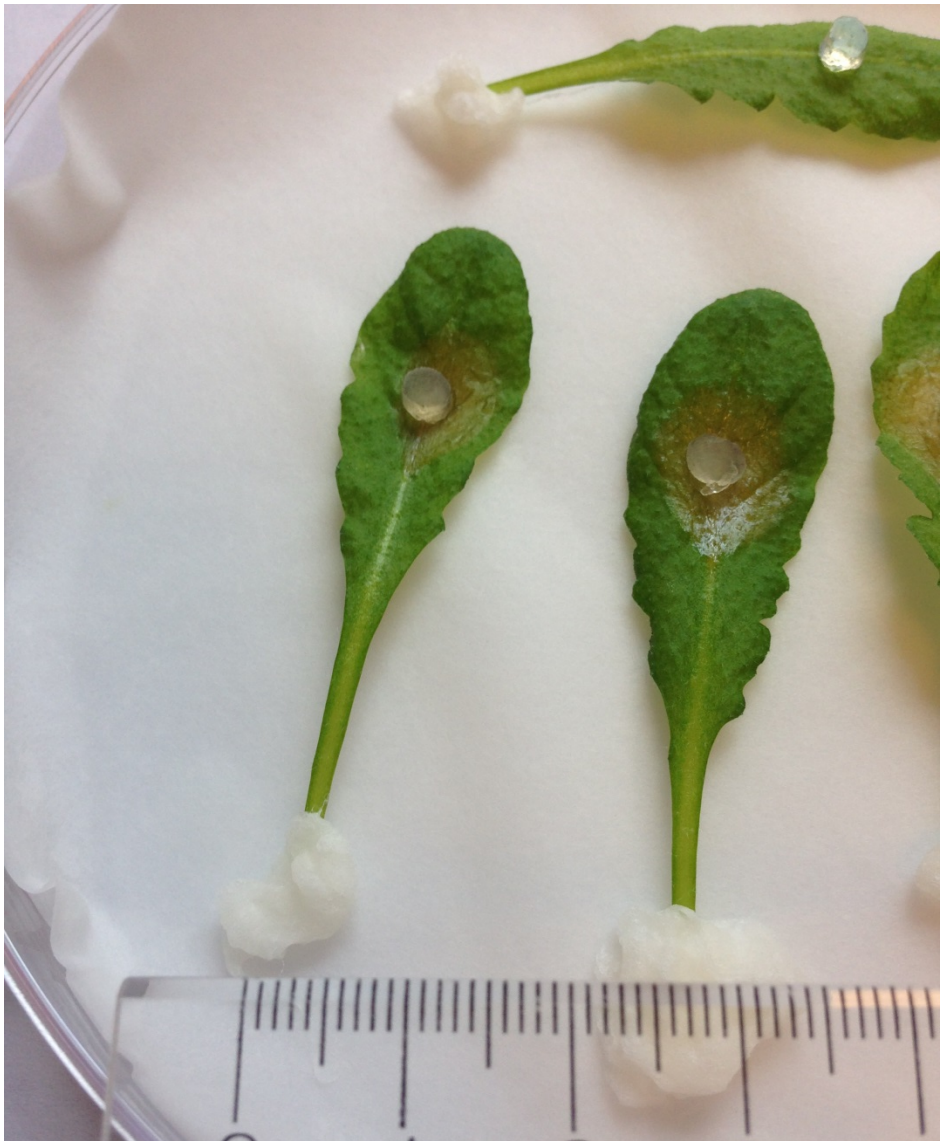

Figure B4. Transgenic line #1, 24 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

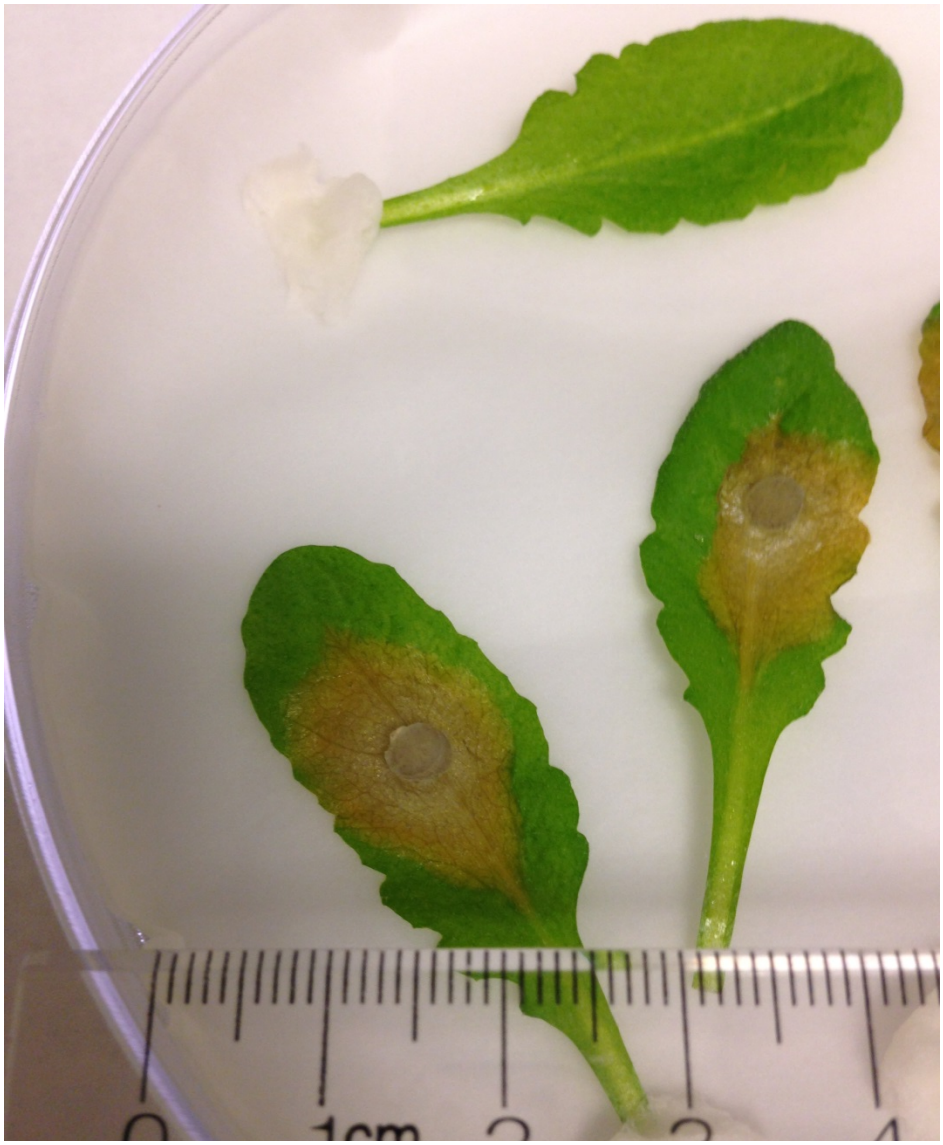

Figure B5. Transgenic line #1, 36 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

Figure B6. Transgenic line #1, 48 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

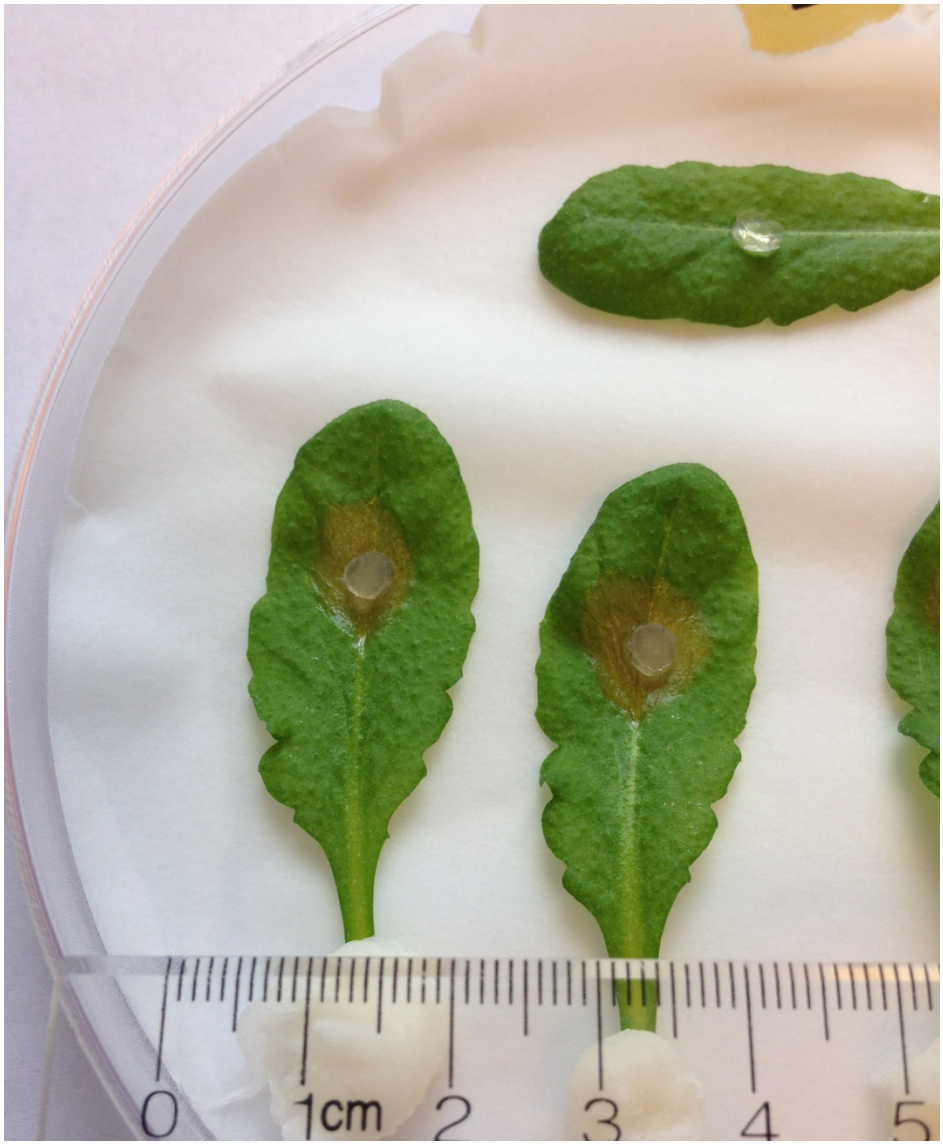

Figure B7. Transgenic line #4, 24 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

Figure B8. Transgenic line #4, 36 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

Figure B9. Transgenic line #4, 48 hrs after being inoculated with Sclerotinia sclerotiorum initiating from a plug of potato dextrose agar medium. Replicate 1-A

Figure B10. Composite image of Sclerotinia assay for Col-0, candidate gene line #1 and candidate gene line #4 at 24 hours.

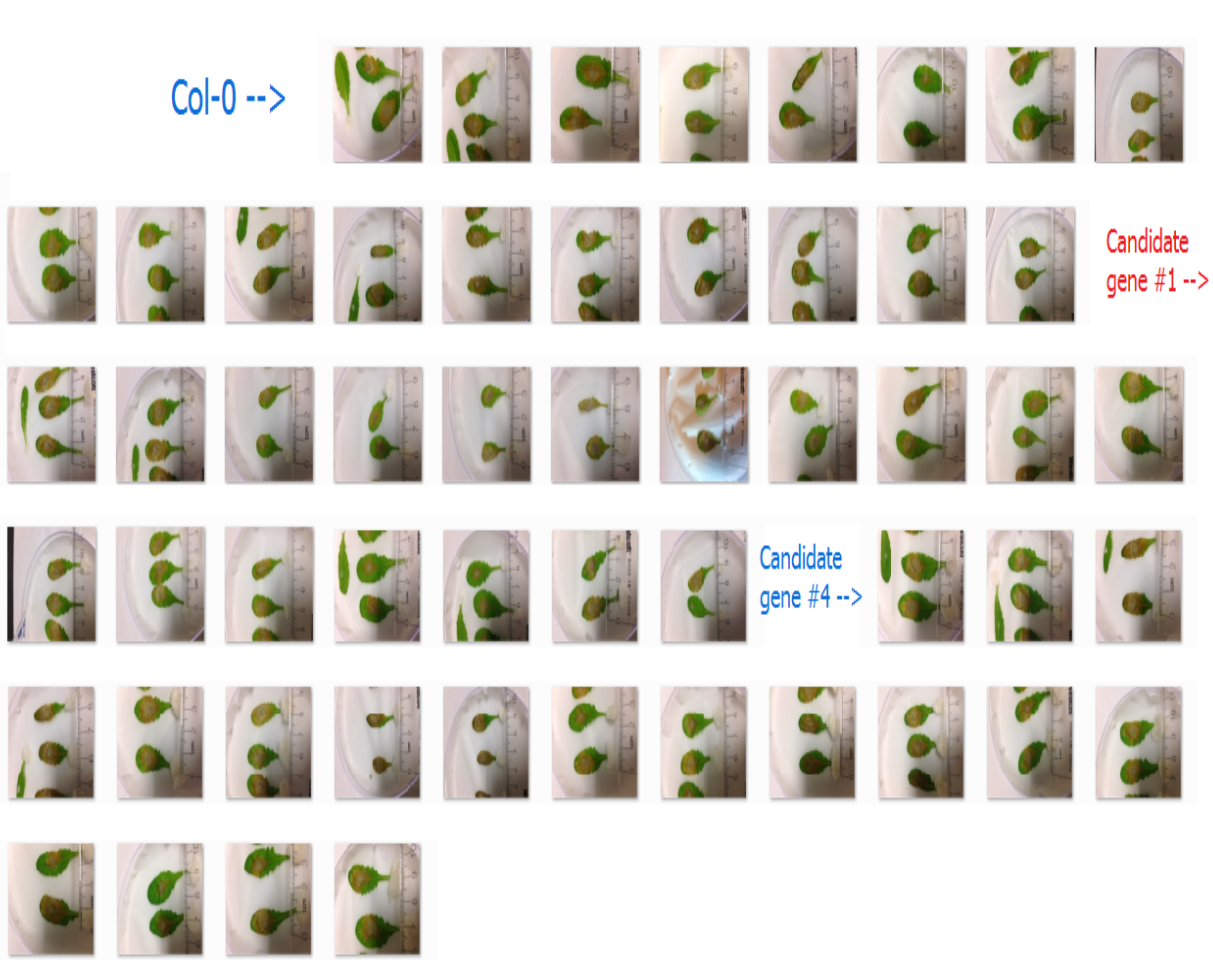

Figure B11. Composite image of Sclerotinia assay for Col-0, candidate gene line #1 and candidate gene line #4 at 36 hours.

Figure B12. Composite image of Sclerotinia assay for Col-0, candidate gene line #1 and candidate gene line #4 at 48 hours.

Table B1. Detailed information about the lesion area measurement for Columbia-0 infected with Sclerotinia sclerotiorum. (expressed in cm^2)

		24 hs		36 hs		48 hs	
col-0	1-a	0.39	0.29	1.42	1.02	2.04	1.46
	1-b	0.72	0.56	1.55	1.34	1.9	1.82
	2-a	0.61	0.41	1.29	1.43	1.76	2.53
	2-b	0.6	0.54	1.1	1.32	1.79	1.84
	3-a	0.7	0.45	1.27	0.9	2.11	1.26
	3-b	0.31	0.14	1.05	0.6	2.02	1.86
	4-a	0.4	0.49	1.21	1.1	2.68	2.59
	4-b	0.55	0.77	1.27	1.55	2.12	2.34
	5-a	0.56	0.42	1.64	1.24	2.35	2.19
	5-b	0.38	0.46	1.47	1.26	2.25	2.06
	6-a	0.54	0.45	1.25	1.34	1.88	1.66
	6-b	0.46	0.37	1.98	1.05	2.35	1.29
	7-a	0.46	0.53	1.21	1.25	1.52	1.68
	7-b	0.49	0.57	1.18	1.18	1.64	1.7
	8-a	0.6	0.62	1.26	1.67	1.94	2.91
	8-b	0.64	0.65	1.45	1.07	2.75	2.08
	9-a	0.68	0.47	1.79	1.24	2.23	2.14
	9-b	0.6	0.51	1.47	1.24	1.98	1.7

Table B2. Detailed information about the lesion area measurement for the transgenic line #1 infected with Sclerotinia sclerotiorum. (The areas are expressed in cm^2)

		24 hs		36 hs		48 hs	
#1	1-a	0.38	0.63	1.31	1.54	1.98	2.49
	1-b	0.61	0.57	1.74	1.47	2.73	2.9
	2-a	0.25	0.36	0.446	0.629	0.913	1.45
	2-b	0.27	0.19	0.554	0.575	1.32	1.32
	3-a	0.585	0.341	0.78	0.514	1.26	1.15
	3-b	0.268	0.799	0.684	0.567	1.09	0.925
	4-a	0.518	0.203	1.52	0.464	1.81	0.825
	4-b	0.316	0.232	0.854	0.747	1.37	1.6
	5-a	0.483	0.767	1.04	1.47	1.83	2.07
	5-b	0.396	0.269	0.924	1.26	1.57	2.32
	6-a	0.541	0.69	1.77	1.88	2.34	2.81
	6-b	0.668	0.73	1.61	1.8	2.48	2.37
	7-a	0.733	0.657	1.25	1.52	1.93	2.1
	7-b	0.58	0.571	1.16	1.24	1.96	1.82
	8-a	0.247	0.245	0.946	0.814	1.72	2.2
	8-b	0.462	0.292	1.1	0.601	1.84	1.5
	9-a	0.579	0.433	1.7	0.907	1.64	1.26
	9-b	0.083	0.478	0.082	1.24	0.283	1.3

Table B3. Detailed information about the lesion area measurement for the transgenic line #4 infected with Sclerotinia sclerotiorum (The areas are expressed in cm^2)

		24 hs		36 hs		48 hs	
#4	1-a	0.494	0.644	1.41	1.75	2.41	2.65
	1-b	0.622	0.621	1.75	1.3	2.35	2.23
	2-a	0.779	0.843	1.27	1.25	2.1	2.03
	2-b	0.822	0.729	1.16	0.829	1.82	1.49
	3-a	0.384	0.734	1.21	1.85	2.22	2.74
	3-b	0.401	0.624	1.25	1.73	2.2	2.38
	4-a	0.608	0.714	1.12	1.05	1.14	1.51
	4-b	0.818	0.809	1.06	1.2	1.28	1.63
	5-a	0.573	0.481	1.4	0.711	2.41	1.96
	5-b	0.683	0.389	1.27	0.735	2.5	2.07
	6-a	0.519	0.763	1.5	1.54	2.55	2.38
	6-b	0.702	0.639	1.71	1.69	2.81	2.79
	7-a	0.55	0.722	1.54	1.74	2.68	2.66
	7-b	0.693	0.546	1.61	1.45	2.89	2.69
	8-a	1.03	0.324	1.75	0.781	2.57	1.55
	8-b	0.327	0.269	0.904	0.466	2.23	1.5
	9-a	0.807	0.557	1.51	0.861	2.63	2.37
	9-b	0.58	0.512	1.45	1.24	3.72	3.14

APPENDIX C

Figure C1. Plant transformed with candidate gene #1 showing the stem with three rosettes.

Figure C2. On the rosette, a new stem grew, forming a new individual plant.

Figure C3. Picture of the transgenic line #1 after 22 days of putting the stem in contact with the soil. On the right, a wild type *Arabidopsis* for comparison.

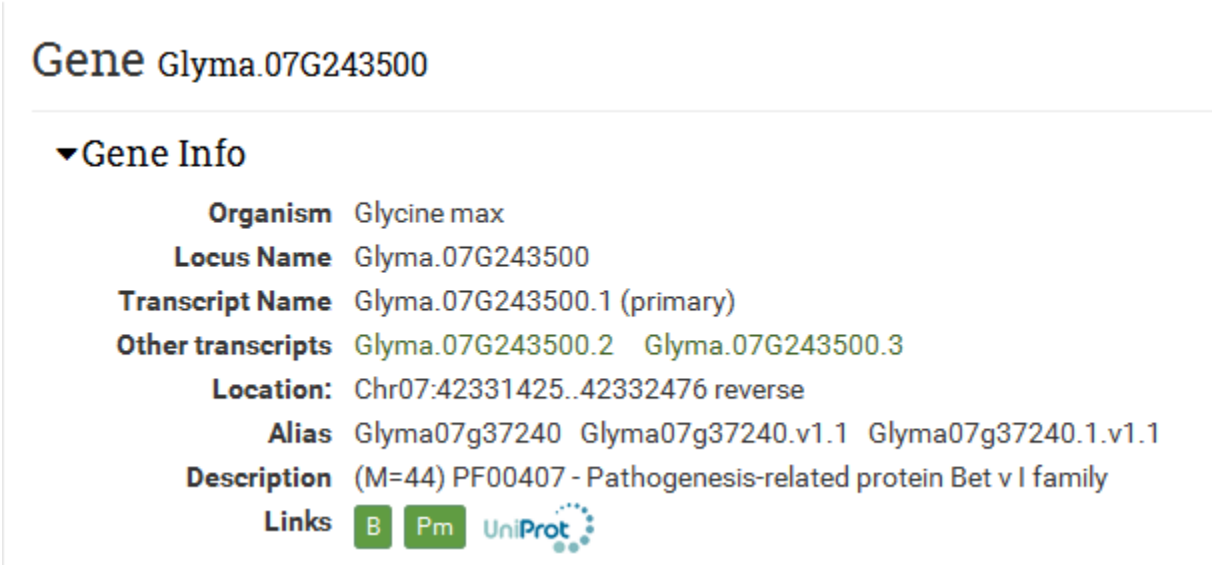

Figure D1. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10: Chr07:42332153..42332333 reverse

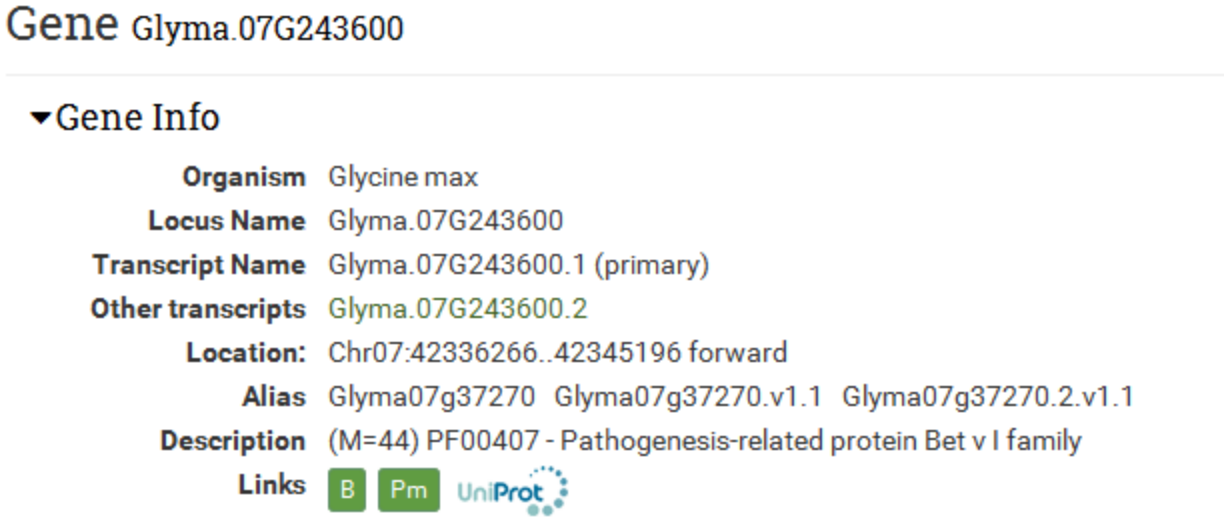

Figure D2. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr07:42336266..42345196 forward

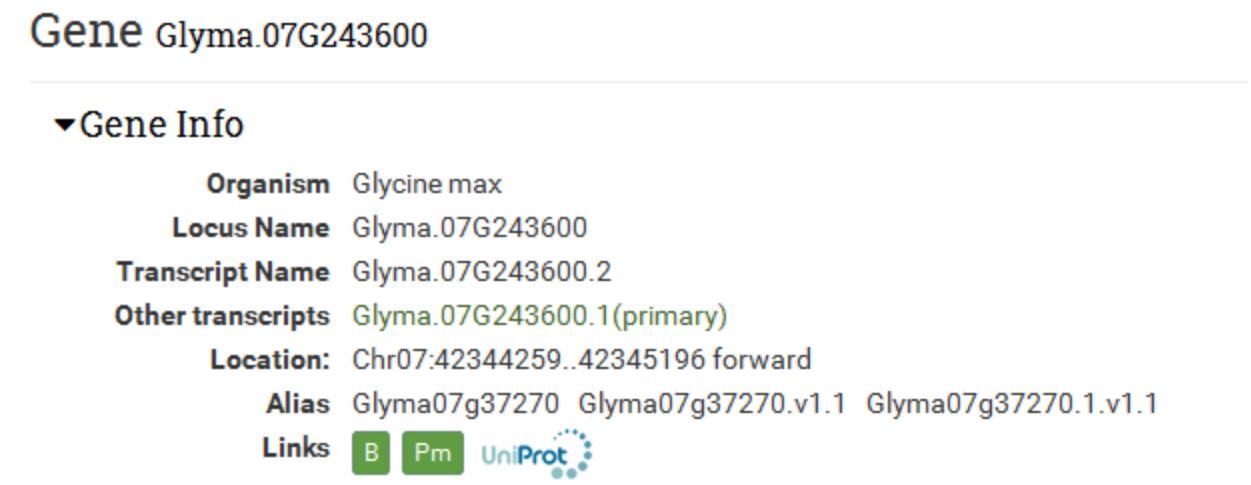

Figure D3. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr07:42344335..42344515 forward

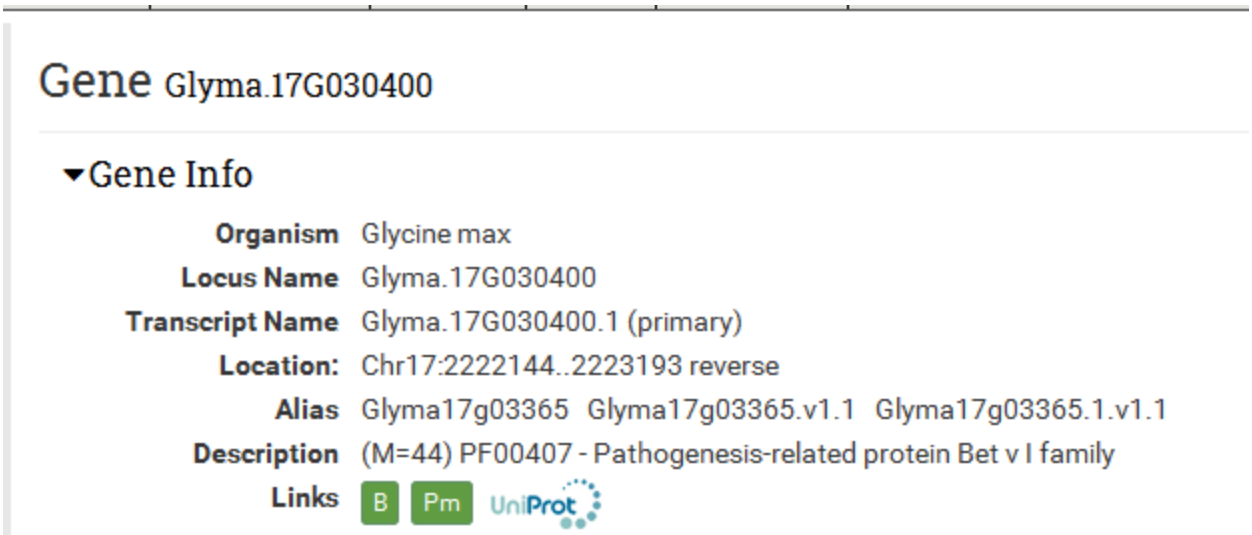

Figure D4. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr17:2222144..2223193 reverse

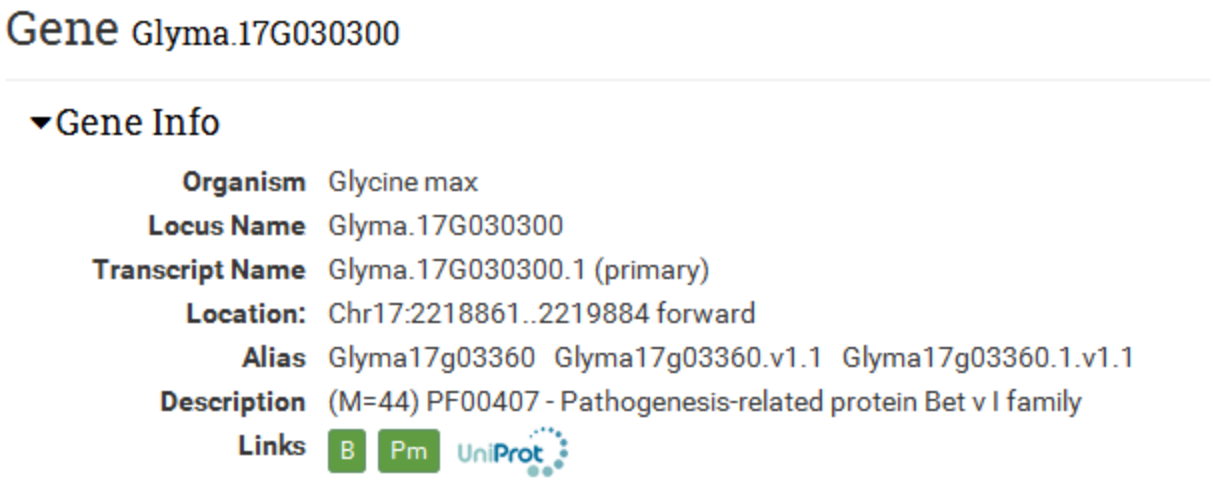

Figure D5. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr17:2218861..2219884 forward

Figure D6. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr17:2215796..2216870 reverse

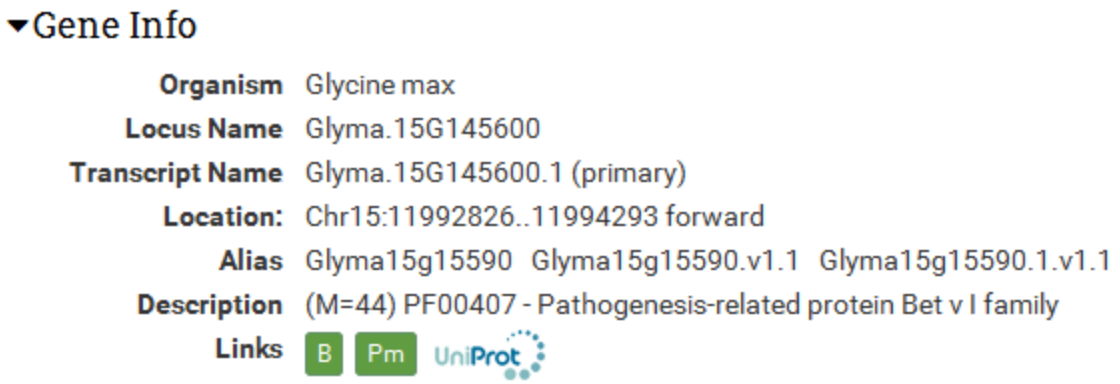

Figure D7. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr15:11992826..11994293 forward

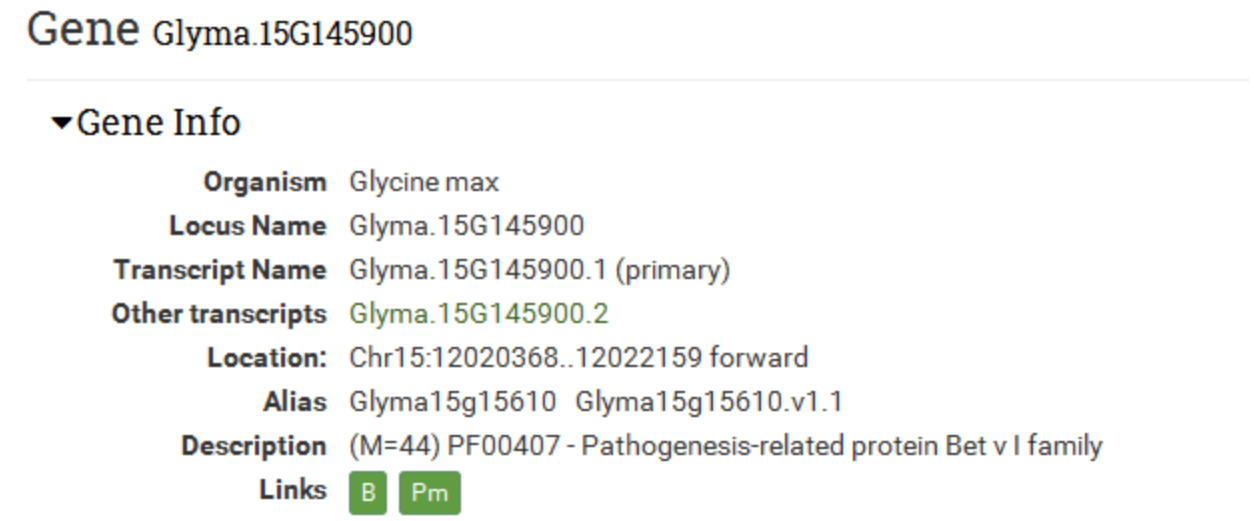

Figure D8. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr15:12020368..12022159 forward

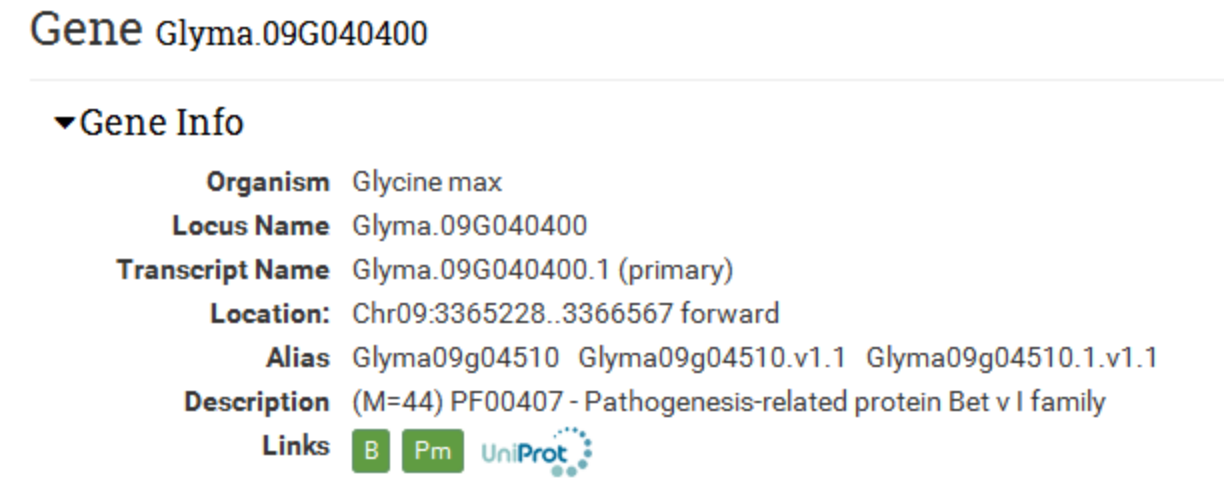

Figure D9. BLAST results in Glycine Max Wm82.a2.v1 after matching with the consensus sequence for PR-10 location at Chr09:3365228..3366567 forward

Figure D10. BLAST results in Glycine Max Wm82.a2.v1 after matching fragment belonging to the CDS sequence for gene #1, location at Chr07:42331425..42332476 reverse.

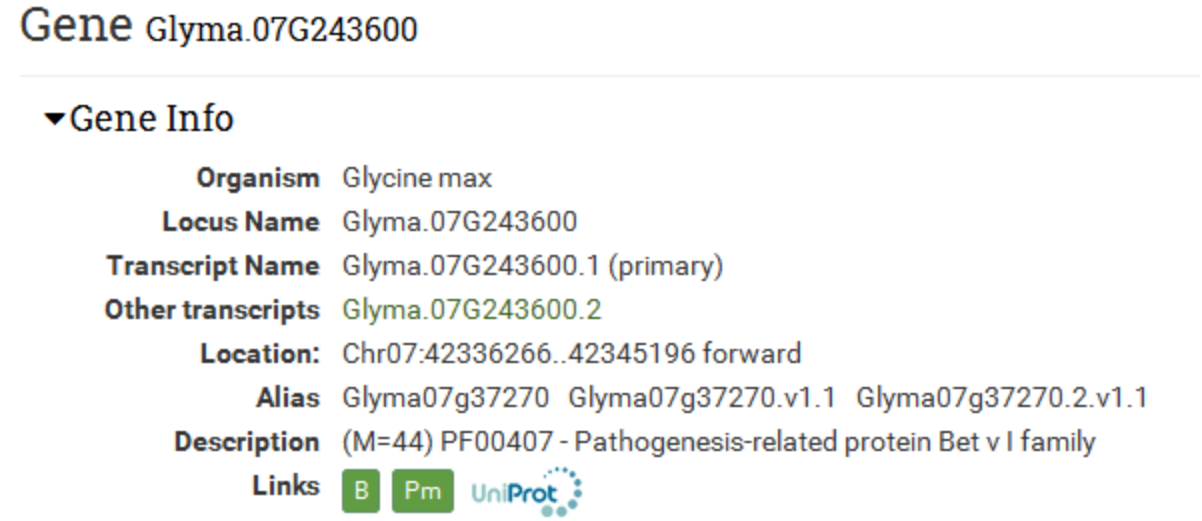

Figure D11. BLAST results in Glycine Max Wm82.a2.v1 after matching fragment belonging to the CDS sequence for gene #1, location at Chr07:42336266..42345196 forward