THESIS

PHYTOALEXIN DEFICIENT4 (PAD4): A PLANT DEFENSE REGULATORY GENE WITH DISTINCT ALTERNATIVE SPLICING PATTERNS IN TOMATO (*SOLANUM LYCOPERSICUM*) AND SOYBEAN (*GLYCINE MAX*)

Submitted by

Rebecca Schmidt

Department of Agricultural Biology

In partial fulfillment of the requirements

For the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2023

Master's Committee:

Advisor: Vamsi Nalam

Cristiana Argueso
Anireddy Reddy

ABSTRACT

PHYTOALEXIN DEFICIENT4 (PAD4): A PLANT DEFENSE REGULATORY GENE WITH DISTINCT ALTERNATIVE SPLICING PATTERNS IN TOMATO (*SOLANUM LYCOPERSICUM*) AND SOYBEAN (*GLYCINE MAX*)

Alternative splicing is an important post-transcriptional regulatory mechanism that contributes to a plant's ability to perceive and respond to a variety of biotic and abiotic stressors. Alternative splicing has a documented role in plant immunity, as many *R* genes, which are important for plant defense against specialized pathogens, undergo alternative splicing in response to pathogen perception. Despite this, the role of alternative splicing in other components of plant defense responses is not well documented. As transcriptome data diversify to include more species and conditions, the extent of alternative splicing in plants has become apparent. *PHYTOALEXIN DEFICIENT4 (PAD4)*, plays an integral role in plant defense signaling to biotic stressors, and in regulating responses to abiotic stresses. *PAD4* undergoes alternative splicing in Soybean (*Glycine max*). Additionally, the expression pattern of *Glycine max* PAD4, *GmPAD4,* and its splice variant *GmPAD4-AS1* are further characterized in early growth stages. We hypothesize *PAD4* produces full-length and alternatively spliced transcripts in multiple species, and that *PAD4* gene structure may influence the occurrence of alternatively spliced transcripts. Here we characterize alternative splicing of *PAD4* in tomato (*Solanum lycopersicum*), identifying two splice variants. We also investigate the conservation of *PAD4*

intron-exon structure conservation across diverse species. *PAD4* expression patterns are characterized using available expression data.

ACKNOWLEDGEMENTS

I am extremely grateful to my advisor, Dr. Vamsi Nalam, without whom I would not have had the opportunity to explore plant pathology so deeply. I am indebted to him for his patience and critical input, which have helped to shape me into the scientist I am. I would like to express my deepest gratitude to my committee, Dr. Cris Argueso and Dr. Anireddy Reddy, for their insight throughout my graduate experience. I would especially like to thank Cris, for her unwavering guidance.

I am grateful for the expertise and companionship of previous and current Nalam Lab members. A special thanks to Jennifer Rushton, for being an incredible friend and scientist, and also, for helping me to move several large home appliances. In addition, I would like to thank the members of the Leach Lab, especially Emily Luna, for their valuable advice. I would like to acknowledge the help provided to me by Dr. Ana Cristina Fulladolsa, it has been a pleasure to collaborate with you.

I would like to deeply thank my family, friends, and, partner, for their constant support. I am confident that without all of you I would not be where I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: ALTERNATIVE SPLICING IN PHYTOALEXIN DEFICIENT4

Introduction

Plants are sessile organisms forced to be highly adaptable to survive in their environment. Therefore, they must be able to interpret and adjust to various environmental stressors. Plants possess numerous gene networks that readily respond to abiotic and biotic stressors. Coordinating responses and correctly distributing resources is essential to a plant's survival. There are numerous ways that this regulation occurs. Alternative splicing is a widespread phenomenon that is important for plants. It plays a role in growth and development, and response to abiotic and biotic stressors. Alternative splicing plays an important role in the regulation of complex gene networks. These networks allow plants to survive and thrive in ever-changing conditions. However, alternative splicing is understudied, and despite the knowledge that it occurs during these processes, the full breadth of its impact is not well understood.

Alternative Splicing

Alternative splicing is a post-transcriptional mechanism that allows organisms to increase transcriptome and proteome complexity, allowing plants to fine-tune responses to abiotic and biotic factors (Barbazuk et al., 2008). Rates of alternative splicing in plants can be highly variable and highly dependent on the species (Table 1). Some, such as Arabidopsis, have documented alternative splicing rates of up to 70% (Martín et al., 2021). In contrast, Stiff Brome is only documented to have alternative splice rates of 6.3% (Walters et al., 2013).

Despite the widespread prevalence of alternative splicing in plants, there is a lack of functional characterization of the impact of alternative splicing on plant proteins and their downstream functions. This is partly due to the difficulty of reliably predicting splice variants through sequence-driven technology (Brown et al., 2015). Emerging research indicates that alternative splicing is necessary for physiological and developmental processes and plant immunity (Staiger & Brown, 2013). Changes to alternative splicing prevalence within the transcriptome has been documented in Arabidopsis plants under high and low-temperature stress (Filichkin et al., 2010; Calixto et al., 2018). Transcriptome data indicates that alternative splicing is a part of regulatory processes that influence plant response to abiotic stressors (Mastrangelo et al., 2012). Alternative splicing events have been identified in several R genes, including *RESISTANCE TO PSEUDOMONAS SYRINGAE4 (RPS4)*, which is found to have an intron retention event that is necessary for induction of defenses in response to a virulence factor (Zhang & Gassmann, 2007).

A single multi-exon gene may have constitutive splicing. A single mature mRNA is produced after introns are removed, and exons are ligated back together in the order in which they appear in the gene (Figure 1). However, some multi-exon genes may also undergo alternative splicing, leading to variable mature mRNAs that result from variable use of splice sites (Reddy et al., 2013) (Figure 1). Alternative splicing is important for increasing transcriptome diversity but is also tightly paired with a transcript degradation pathway, non-sense mediated decay (NMD), which frequently targets transcripts with premature stop codons (Kalyna et al., 2012). While there are numerous ways to categorize splicing events, here we will discuss the events as they are most commonly categorized. The first, exon skipping (ES) results

in the removal of an exon and the formation of a truncated transcript. Second, alternative 3' or 5' splice sites can be used, which will also result in a truncated transcript. And third, intron retention (IR) can occur, typically resulting in longer transcripts.

Plant Defense

Plants possess a complex and adaptable immune system that encompasses several layers of immunity to defend against various pathogens and pests. Regulation of each level is essential to deter pathogens successfully and for a plant's long-term health and fitness. In order to conserve plant resources, induced immunity will only occur after the perception of a pathogen and the induction of downstream signals. The first layer of immunity aims to identify highly conserved pathogen patterns common to many pathogens, known as pathogen-associated molecular patterns (PAMPs) (Figure 2). Patterns that may be recognized include flg22, a pattern found in bacterial flagella, or chitin, a pattern common to fungi. The induction of Pattern-Triggered Immunity (PTI) through pattern recognition receptors (PRRs) can provide sufficient defense against non-specialized pathogens (Figure 2). However, some pathogens will evade PTI using effectors, molecules that influence host defense and physiology. In turn, plants have evolved to identify these pathogen or pest-produced effectors through nucleotide-binding leucine-rich proteins (NLRs), leading to effector triggered immunity (ETI). Identification of pathogen effector and induction of PTI or ETI frequently leads to the hypersensitive response (HR), a form of programmed cell death that allows a plant to stop pathogen spread. PTI and ETI play an essential role in immunity, offering rapid activation of defenses to general pathogen patterns and pathogen-specific effectors, respectively.

Induction of ETI will also lead to a more robust and longer-lasting defense response. A final layer of immunity, systemic acquired response (SAR), can be triggered after an initial challenge, leading to systemic defense responses throughout the plant. Induction of SAR can lead to local changes such as phytoalexin accumulation, cell wall modifications, salicylic acid (SA) accumulation, and systemic changes throughout the entire plant. Together these three components, PTI, ETI, and SAR, make up the innate immunity of a plant.

The complexity of plant defense is not surprising, as defenses will be tested by various pathogens, from necrotrophs seeking to encourage cell death to biotrophs requiring living cells to survive. Plant innate immunity is an adaptable and complex system that relies on large-scale transcriptional regulation. Underlying these large-scale transcriptional changes is an elaborate hormonal crosstalk system between SA, jasmonic acid (JA), and ethylene (ET) signaling (Glazebrook, 2005; Bari and Jones, 2009; Berens et al., 2017). The SA pathway primarily involves mounting responses to biotrophic and hemibiotrophic pathogens and is essential for activating SAR (Grant and Lamb, 2006; Bari and Jones, 2009). In contrast, JA and ET regulate plant responses to necrotrophic pathogens and herbivorous insects (Bari and Jones, 2009). The SA and JA/ET pathways are considered to be mutually antagonistic (Bari and Jones, 2009) although exceptions exist (Mur et al., 2006). This antagonism is essential for coordination of defense responses to pathogens of differing lifestyles. Ultimately, because the plant immune system must be adaptable enough to defend against biotrophs and necrotrophs, the hormonal pathways that facilitate defense against these diverse pathogens are tightly controlled by each other.

PHYTOALEXIN DEFICIENT4

PHYTOALEXIN DEFICIENT4 has a multifaceted influence on plant defenses and fitness. The gene is involved in regulating defenses against a variety of pathogens, including bacteria (Feys et al., 2005a), aphids (Pegadaraju et al., 2005; Louis et al., 2012; Louis and Shah, 2014), nematodes (Youssef et al., 2013; Wubben et al., 2008) and fungi (Rietz et al., 2011; Makandar et al., 2015). Arabidopsis PAD4 has been identified as a necessary regulatory component of PTI, ETI and SAR induction (Zhou et al., 1998; Tsuda et al., 2009; Rietz et al., 2011; Pruitt et al., 2021). *PAD4* was initially identified during screens for enhanced disease susceptibility when Arabidopsis mutants were exposed to the bacterial pathogen, *Pseudomonas syringae* (Glazebrook et al., 1996).

In addition to being a regulatory component for plant defense, *PAD4* influences overall plant fitness and response to abiotic stressors. *PAD4*, and its interacting partner, *ENHANCED DISEASE SUSCEPTIBILITY (EDS1)*, are conserved with orthologs found only in angiosperms, with no orthologs identified in Rhodophyta, Chlorophyta, or Bryophyta, suggesting that the PAD4/EDS1 hub evolved at a similar time to plant vasculature (Wagner et al., 2013, Baggs et al., 2020). Many identified *PAD4* orthologues are involved in plant defense against biotic pathogens; despite this, the signaling pathways they act through vary (Wiermer et al., 2005; Ke et al., 2014; Makandar et al., 2015). *AtPAD4* plays a role in the SA pathway that is integral for plant defense against biotrophic pathogens (Wiermer et al., 2005). In contrast, Rice *PAD4* (*OsPAD4*) mediates defenses against some biotrophic pathogens through the JA pathway (Ke et al., 2014). Additionally, some orthologues of *PAD4* have differences in localization, which may

imply differences in function. *OsPAD4* is a plasma membrane localizing protein, whereas *AtPAD4* is a nucleo-cytoplasmic protein (Ke et al., 2014, Feys et al., 2001; Rietz et al., 2011).

PAD4 is a lipase-like protein that contains an N-terminal α/β fold hydrolase domain, which contains a Ser-Asp-His (S-D-H) catalytic triad and a C-terminal EP (EDS1-PAD4) domain (Wiermer et al., 2005; Wagner et al., 2013). The EP domain is integral for the interaction between PAD4 and EDS1. Wagner et al. identify amino acids 1-299 as the lipase-like domain and 200-451 as the EP domain in Soybean, *Glycine max* PAD4. In *Arabidopsis,* PAD4 localizes to the cytoplasm, whereas the EDS1-PAD4 complex has been observed in the nucleus (Feys et al., 2001; Rietz et al., 2011). Despite the presence of a catalytic triad, PAD4 does not have a documented lipase function (Wagner et al., 2013). *SENESCENCE-ASSOCIATED GENE101 (SAG101)*, a second interacting partner of EDS1, shares similar protein topography to that of the EDS1 N-terminus, which is the assumed interaction site for the EDS1-PAD4 and EDS1-SAG101 heterodimers (Wagner et al., 2013)(Figure 3). Despite the structural similarities between PAD4 and SAG101, SAG101 lacks the catalytic triad within its lipase domain (Feys et al., 2005a; Wagner et al., 2013). Previous studies indicated that EDS1 could interact exclusively with PAD4 or SAG101 *in vitro* (Xing & Chen, 2006; Wagner et al., 2013). However, an *in planta* study indicates that there may be a tertiary complex, EDS1-PAD4-SAG101, that can be formed and that SAG101 and PAD4 may influence the localization of EDS1 to the nucleus and cytoplasm, respectively (Zhu et al., 2011).

PAD4, in conjunction with its interacting partners, *EDS1* and *SAG101*, form a defense regulatory hub. This hub is necessary for several defense pathways, including PTI, ETI, and SAR, primarily through the modulation of SA (Rietz et al., 2011). PAD4 and EDS1 are increased by SA,

which suggests that these proteins are essential for defense attenuation and amplification (Jirage et al., 1999; Zhou et al., 1998; Xiao et al., 2003). Due to the role of *PAD4* and *EDS1* in SA accumulation and signal attenuation, the role of *PAD4* in response to necrotrophic pathogens is unclear, as the induction of programmed cell death would further pathogen infection (Xing and Chen, 2006; Brodersen et al., 2006).

In Arabidopsis*, MAP Kinase 4 (MPK4)* has been identified as a negative repressor of the SA defense pathway and an activator of the JA/ET defense pathway against necrotrophs (Brodersen et al., 2006). MPK4 represses the SA pathway through interaction with PAD4 and EDS1, activating the JA/ET pathway due to the antagonistic nature of the SA and JA defense pathways (Brodersen et al., 2006).

PAD4 promotes hypersensitive response, a type of programmed cell death, in response to various pathogens, providing a mechanism to contain pathogens locally and conserve resources for the host plant (Rietz et al., 2011). Studies suggest that *PAD4* changes sphingolipid accumulation, leading to programmed cell death (Zeng et al., 2021). Resistance mediated by PAD4 to the sap-sucking green peach aphid (*Myzus persicae*) does not require *EDS1* and is not SA-dependent (Pegadaraju et al., 2005; Louis et al., 2012; Lei et al., 2014; Dongus et al., 2020). Additionally, *PAD4*-mediated resistance to GPA requires the residue S118 within the S-D-H catalytic triad in Arabidopsis (Louis et al., 2012; Dongus et al., 2020). *PAD4*-mediated defense acts through antibiosis, which negatively impacts aphid development and reproduction, and antixenosis, which deters aphids from settling or feeding. *Solanum lycopersicum* (Tomato) has a similar antibiosis and antixenosis effect on GPA (Singh and Shah, 2012). In addition, *SlPAD4* expression after GPA infestation was similar to that of AtPAD4 under the same conditions

(Singh and Shah, 2012). Together this evidence suggests that the antixenosis and antibiosis function of *PAD4* may be conserved across species.

PAD4 is an essential component of plant response to abiotic stress, including excess excitation energy, which produces reactive oxygen species, and root hypoxia (Mühlenbock et al., 2008). *PAD4* and *EDS1* positively control the induction of foliar ET and H_2O_2 in response to redox status during light stress (Mühlenbock et al., 2008). Additionally, *PAD4* and *EDS1* regulate water use efficiency, photosynthesis efficiency, and seed yield (Wituszyńska et al., 2013; Szechyńska-Hebda et al., 2016). *PAD4*, *EDS1*, and *LESION STIMULATING DISEASE1* (*LSD1*) are suspected of acting as a regulatory hub in response to abiotic factors through regulating H_2O_2, SA, and ET concentrations in cells (Wituszyńska et al., 2013, Ng et al., 2011, Mühlenbock et al., 2008). Defense against a pathogen is not only about host identification and defending against a pathogen but also maintaining the fitness of a host plant after the infection. In Arabidopsis, it is known that *PAD4* is influential in the survival of plants under drought stress and has an overall regulatory effect on plant vegetative growth (Szechyńska-Hebda et al., 2016). Additionally, *PAD4* modulates plant acclimation and survival in freezing conditions, with AtPAD4 knockout plants displaying increased survival with and without cold acclimation (Chen et al., 2015). In *Populus tremula x tremuloides*, *PAD4* influences plant survival by regulating vegetative mass production and influences cell wall structure and wood properties, likely by controlling the cell death ratio to cell division (Szechyńska-Hebda et al., 2016).

PAD4 contributes to the defense against pathogens in several ways, individually and with *EDS1* and *SAG101*. *PAD4* is essential for the temporal regulation of premature cell death through SAGs (Louis et al., 2010). Coordinated temporal responses of programmed cell death to

a pathogen impact immune function and are essential for maintaining seed yield as the host plant matures (Louis et al., 2010). Arabidopsis *pad4-1* plants were found to have reduced seed yield after GPA infestation, indicating that *PAD4* may help to conserve resources in response to GPA and other pathogens (Pegadaraju et al., 2005, Wituszyńska et al., 2013). In addition, *PAD4* involvement in coordinating responses to abiotic stress contributes to its overall role in maintaining plant fitness.

Preliminary research indicates that the *PAD4* gene undergoes alternative splicing in Soybean (Selig, 2017). In Soybean, GmPAD4 transcription results in a full-length transcript of 1,902 bp and a splice variant (GmPAD4-AS1) with a length of 1,178 bp. Analysis of these sequences shows that exon 3 is spliced out in GmPAD4-AS1. In addition, evaluation across eight *Glycine max* cultivars indicates that GmPAD4 and GmPAD4-AS1 transcripts are similar between cultivars (Selig, 2017).

Objectives

The first objective of this work is to evaluate the conservation of *PAD4* structure and expression across diverse plant tissues and species. I hypothesize that *PAD4* structure and expression patterns are conserved across tissues and species. The second objective is to identify *PAD4* splice variants in the dicot plant *Solanum lycopersicum.* I hypothesize that due to the intron-exon structure that splice variants can be identified and cloned in *Solanum lycopersicum.* The third objective is to characterize *Glycine max PAD4* overexpression lines for future use and evaluate the generation of transgenic GmPAD4 Arabidopsis lines for future use.

Methods

Plant Growth Conditions

The soybean plants used for soybean colonies and aphid assays were grown in PRO-MIX BS soil. Soybeans that required sampling of early-stage soybeans and root tissue were grown in a mixture of 1:1 vermiculite and PRO-MIX BS soil to allow for better cleaning of tissues before extraction. Soybean plants were cultivated in a growth tent at 23°C with a light regime of 16:8 L:D. Tomato plants were grown at 20°C on racks with growth lights set to 16:8 L:D. Arabidopsis plants were grown in a growth chamber or tent at 23°C and 16:8 L:D cycles. All plants were fertilized weekly with Plant Food (Miracle-Gro). Plants were watered when soil surface appeared slightly dry. Watering was performed by sub-irrigating and excess water was allowed to drain after each watering.

Soybean Aphid Colony

Soybean aphids were collected from Pinney Purdue Agricultural Center (PPAC), Watanah, Indiana. The colony was maintained within a 2x2x4 foot growth tent (CoolGrows). The colony was maintained at 23°C with a light regime of 16:8 L:D within grow tent on *Glycine max* cv. Williams 82 plants. Soybean plants were replenished weekly to ensure colony health.

Cloning

Cloning was performed using a pCR™8/GW/TOPO (Invitrogen) Cloning kit. PCR reaction was performed, and the products were separated on a 0.7% agarose gel at low voltage. Immediately after imaging, bands were excised from the gel using a razor blade and placed into individual tubes. Care was taken to excise the whole individual band but retain little excess

agarose gel. After excision, individual bands were cloned directly from excised gel bands using the procedure described in pCR™8/GW/TOPO (Invitrogen) Cloning kit manual. Following TOPO cloning reaction, the entire volume was added to a transformation reaction for TOP10 *E. coli* cells and transformed into cells via heat shock (transformation as described in the manual). After transformation, colonies were plated on LB with Spectinomycin at 100 µg/mL and incubated overnight at 37°C. TOP10 *E. coli* colonies are then screened using M13 primers. Positive colonies were cultured overnight at 37°C in LB with Spectinomycin at 100 µg/mL. Plasmid extraction was performed with Zyppy® Plasmid Miniprep Kit (Zymo Research). Cloned insertions were Sanger sequenced using M13 primers (Table 2)(GENEWIZ). These sequences were then aligned with the known sequence of PAD4 for the species to characterize PAD4 splicing within each species. Sequenced plasmids were then translated to proteins using the Benchling translation function and aligned using EMBL-EBI Clustal Omega Multiple Sequence Alignment with STRAP Alignment Annotations (Gille et al., 2014).

Semi-quantitative (RT-PCR) Analysis

Plant tissue was harvested and flash frozen in liquid nitrogen and stored at -80°C for downstream use. RNA was extracted from tissue samples using the Direct-zol RNA Miniprep Plus Kit (Zymo Research). cDNA was synthesized using the Verso cDNA Synthesis Kit (Thermo Scientific™). cDNA synthesis was a normalization step, and all samples were normalized to 1 µg of RNA template based on Nanodrop values (Thermo Scientific™). Following cDNA synthesis, PCR was performed using PAD4.2F and PAD4.2R primers, designed for Soybean PAD4, predicting product sizes of 200 bp for GmPAD4-AS and 900 bp for GmPAD4 and the *Glycine max* housekeeping gene primers GmFboxF and GmFboxR. The PCR product was then run on a 1.5%

agarose gel and visualized. Digital images of the PCR products on agarose gels were quantified using ImageJ for the products at the 200 bp (GmPAD4-AS) and 900 bp (GmPAD4) sizes. Plots were made using ggplot2 in R.

Soybean Aphid Fecundity

A fecundity assay was performed by placing ten adult aphids (3-4 per leaflet) on the first trifoliate leaf after full emergence. Petroleum jelly was applied to the petiole of the first trifoliate to ensure that all aphids were confined to the first trifoliate. Adult and nymphs were counted each day for five days; only aphids on the first trifoliate were counted. Nymphs per adult aphid (NpA) was calculated by dividing the number of nymphs by the corresponding adult aphid count.

Statistical Analysis

Statistical analysis was carried out in R with p-values of 0.05 as the cutoff for significance. Datasets were checked for normality using the Shapiro-Wilk test and for equality of variance using Levene's test. For parametric datasets, differences in means were analyzed using a one-way ANOVA. For datasets that were non-parametric, the Kruskal-Wallis one-way analysis of variance was used.

Generation of GmPAD4 Arabidopsis Transformants

Arabidopsis plants with background Col-0 (wild ecotype) and *pad4-1* (PAD4 knockout ecotype) have been transformed using the floral dip method (Clough & Bent, 1998) to introduce a plant vector (pEarleyGate100, Earley et al., 2006) (Figure 4). pTWIST ENTR vectors with 35S:GmPAD4 and 35S:GmPAD4-AS1 were ordered from pTWIST Biosciences and transformed

into Top10 *E. coli* using heat shock. Transformed *E. coli* were then plated on LB and Kanamycin (50 µg/mL) plates and incubated at 37°C overnight.

Successful transformation was confirmed by colony PCR using the primer set attB1-PAD4 F and PAD4seq1R which spanned the insertion site and the start of PAD4 constructs and had an anticipated product size of 179bp (Figure 5). Positive colonies were then grown in culture, and plasmids were miniprepped using the Zyppy Plasmid Miniprep Kit (Zymo Research). GmPAD4 inserts were then Gateway cloned from the entry vector, pENTR into the pEarleyGate100 destination vector through a Gateway™ LR Clonase™ (Invitrogen) reaction. Following the reaction, TOP10 *E.coli* cells were transformed with the GmPAD4-pEarleyGate100 plasmids and allowed to incubate overnight on plates. The following day colonies were screened in the same method as above (Figure 6). After confirmation of GmPAD4 construction insertion, plasmids were prepped and sent for Sanger sequencing. Due to the length of the GmPAD4 inserts, a collection of 9 forward primers were used for sequencing (Figure 7).

Sequencing results were then aligned against the known insert sequences (Figure 8). Note that regions boxed in red were modified by Twist Biosciences during construct synthesis and are not anticipated to make changes to the protein product.

Following insert verification, pEarleyGate100-GmPAD4 and pEG100-GmPAD4-AS1 vectors (Figure 4, A and B) were transformed into *Agrobacterium tumefaciens* (GV3101) via the freeze-thaw method (Hofgen & Willmitzer, 1988). Recovered cells were then plated on LB plates with Kanamycin, Rifampicin, and Gentamycin and incubated for two nights at 28°C. Colonies were then confirmed via colony PCR (Figure 9).

GV3101 Agrobacterium lines that were confirmed to have GmPAD4-pEarleyGate100 vectors in them were then cultured for use in Arabidopsis transformation. Plant transformation was achieved through the floral dip method on the accessions *pad4-1* and Col-0. Transformed Arabidopsis plants were grown to seed (T_0). T_0 seeds were then selected using BASTA (diluted 1:1000). BASTA selection was done in soil by spraying a 1:1000 BASTA solution on Arabidopsis seedlings after vernalization at 10, 12, 15, and 17 days after planting. Seedlings that survived BASTA selection then had lead tissue harvested for DNA extraction via a rapid DNA extraction protocol as described by Edwards et al., 1991. PAD4.2 primers were used to confirm PAD4 insertion with amplicon sizes of 185bp for35S:GmPAD4-AS1 transformants, and 905 bp for 35S:GmPAD4 transformants. Plants that survivied the selection were sampled for PCR confirmation of GmPAD4 insertions (Figure 10).

35S:GmPAD4 and 35S:GmPAD-AS1 transformants were intended to perform the No-choice Bioassay (as described in Nalam et al., 2018). For the assay, plants are grown until 24-26 days old, and then 20 adult apterous (wingless) aphids are placed on each Arabidopsis plant. After 48 hours, adults and nymphs are counted for each plant. 35S:GmPAD4-mCherry and 35S:GmPAD4-AS1-mCherry (Figure 4, C and D) transformants were intended to be used for a confocal microscope localization study. For confocal microscopy, plants would be grown until 24-26 days old and used before and after green peach aphid infestation.

Phylogenetic Analysis

Phylogenetic analysis of protein sequences was carried out using MEGA-X (Kumar et al., 2018). Phylogenetic trees were predicted using the maximum likelihood model. Protein sequences for analysis were acquired from NCBI and Phytozome based on sequence similarity

to Arabidopsis PAD4 and annotation. Intron-exon structure was acquired from NCBI and Phytozome and were predicted from nucleotide sequences.

Results

Alternate Splicing of PAD4 is Conserved Across Species

Comparative analysis of PAD4 protein sequence and intron-exon structure

Given that *PAD4* plays a dual role in plant response to abiotic and biotic stressors, and the fact that it undergoes alternative splicing, we conducted an analysis of *PAD4* protein sequence, and intron-exon structure across the published genomes of 84 plant species. The underlying mechanisms behind *PAD4's* dual role is not well understood, therefore investigation of PAD4 intron-exon structure may shed light on conserved *PAD4* intron exon structure, which in turn may suggest a conservation of alternative splicing events. Alternative splicing occurs in *Glycine max PAD4*, which contains four exons, but has not previously been documented in any other species. *PAD4* protein sequences were identified through similarity and annotation to Arabidopsis (AT3G52430) protein sequence with Phytozome and NCBI Blast. Protein sequences were then aligned using MEGA-X with the Maximum Likelihood Model. In total *PAD4* sequences were identified for 84 species, from 24 orders (Figure 11).

These results indicate that *PAD4* has diverse intron-exon structures across species (Figure 11). 61% of the species with available intron-exon structure data contain more than two exons. Predicted exon numbers range from two (such as in the Brassicaceae family) to five (such as the eastern cottonwood). Alternative splicing within *PAD4* has not been observed in

the most common model plant, Arabidopsis, which may be connected to its intron-exon structure, as PAD4 is predicted to contain only two introns.

For plants with predicted intron-exon structure, an order-level evaluation was performed to determine the conservation of *PAD4* structure within each order. Within the Brassicales order, which contains mustards and allies, the primary structure for *PAD4* includes two exons, with 61% of the surveyed species in the order being predicted to have two exons. The Fabales order, a group containing many important crops such as Soybean, Medicago, and Pecan, has a *PAD4* structure that indicates four exons (Table 3). The Poales order, which contains grasses, bromeliads, and sedges, had more variation than previous orders, but most plants were predicted to have four exons. The Malpighiale order which contains plants such as poplars, willows, and castor bean, with PAD4 exon numbers between three and five. Finally, the Rosales order, which contains plants such as Strawberry, Apple, Pear, and China Rose, is predicted to have a *PAD4* exon numbers ranging from of four to five. The analysis of *PAD4* intron-exon structure across diverse species indicates that the structure is relatively conserved at the order level, in addition to PAD4 orthologues in diverse species (Table 3).

Analysis of PAD4 expression across species and tissue

Conservation of expression patterns within tissues across multiple species can implicate conservation of gene function, for that reason, we investigated the expression patterns of *PAD4* within the species Arabidopsis, Poplar, Tomato, Potato, Soybean, and Rice. Due to *PAD4's* involvement in plant response to abiotic stressors we also analyzed *PAD4* expression changes in leaf tissue after exposure to a variety of abiotic stressors. *PAD4* orthologues were identified by

sequence comparison to Arabidopsis *PAD4* (AT3G52430) and compared to orthologues identified by the Bio-Analytic Resource of Plant Biology (BAR) Navigator viewer (http://bar.utoronto.ca/eplant/)(Table 3). *PAD4* expression data was collected for root, leaf, flower, and fruit. In addition, several species had expression data for *PAD4* under varying abiotic stresses. Tissue-specific *PAD4* expression data is available for Arabidopsis, Poplar, Tomato, Potato, Soybean, and Rice. *PAD4* expression data for plants undergoing abiotic stressors were available for the species Arabidopsis, Potato, and Rice. Table 3 contains a complete list of data acquisition methods, normalization methods, and citations for all data collected from BAR.

In Arabidopsis, Poplar, and Soybean, a similar pattern of PAD4 expression was observed, with highest expression observed in the root tissues, followed by the leaf tissue, and lowest expression in flowers (Figure 12). However, in Tomato and Potato, distinct differences were observed in PAD4 expression patterns across the different plant tissues. High expression of *PAD4* was observed in tomato roots; and also in fruits (Figure 12). Potato also shows a unique profile, with the highest gene expression observed in leaves, followed by fruits (Figure 12). Both Solanaceae family plants show a higher relative expression of *PAD4* within fruit tissues than Soybean, in which *PAD4* expression in fruit tissues is lower than in the other tissues, root, leaf, or flower.

Abiotic stress-induced changes in *PAD4* expression were analyzed in Arabidopsis, Potato, and Rice (Figure 13). The most comprehensive data is available for Arabidopsis, with sampling for each condition performed in a time course at 1, 3, 6, 12, and 24 hours post stress. Significant fold changes in *PAD4* expression were most common at 24 hours, with fold changes

(FC) ≥ 2 recorded for cold stress (FC 2.1), genotoxic stress (FC 2.0), and oxidative stress. During UV-B stress, *PAD4* expression levels compared to the control plant were larger than FC 2 for all time points, with the highest FC of 9.7 observed at 6 h post-stress and lowest at 12 h (FC 2.7) post-stress. In Potatoes, a slight reduction (less than -0.5 FC) of *PAD4* expression under heat and salt stress was observed. Rice *PAD4* expression shows significant FC in response to cold stress (FC 2.3), drought (FC 1.1), and salt (FC 0.9). Taken together this data suggests that PAD4 does have tissue-specific expression. In addition, this data indicates that *PAD4* expression does undergo significant changes after specific abiotic stressors.

Glycine max (GmPAD4) shows an expression of truncated transcript in multiple growth stages and tissues

Soybean PAD4 undergoes alternative splicing to produce one splice (*GmPAD4-AS1*) variant in which exon 3 is skipped in addition to the full-length transcript (*GmPAD4*). To elucidate if the splice variant has a role in plant development, we analyzed *GmPAD4-AS1* and *GmPAD4* expression in several tissues and growth stages. In Soybean*, PAD4* expression shows differences in expression and splicing ratios at the tissue level and in early growth stages, VE, VC, and V1 (Figure 14). At the VE stage, when the cotyledons have emerged from the soil, it shows the highest total *PAD4* expression (Figure 14B, C). This is followed by the V1 stage, when the unifoliate leaves have emerged, with slightly lower *PAD4* expression (Figure 14B, C). The final stage analyzed, VC, when the first unifoliate leaf has fully emerged, has the lowest *PAD4* expression (Figure 14B, C). However, the total *GmPAD4* expression level is not a good predictor of the level of *GmPAD4-AS1*. The V1 growth stage has the highest percentage of alternative splicing, with 46% of total *PAD4* expression attributed to the splice variant (*GmPAD4-AS1*),

followed by the VE stage, with 17% of *PAD4* expression being attributed to *GmPAD4-AS1*. The VC stage has the lowest percentage of *PAD4* splicing, with a value of 9%.

Total *PAD4* expression was higher in radicle (VE, VC) followed by in the stem (VC, V1), and then the unifoliate (VC, V1) (Figure 14). Differences in *GmPAD4* to *GmPAD4-AS1* ratios varied among tissues. Samples from the stem and unifoliate displayed a higher ratio of *GmPAD4-AS1* to *GmPAD4* than other sampled tissues (Figure 15). These data suggest that *PAD4* does in fact have different expression levels, and ratios of *GmPAD4* and *GmPAD4-AS1* ratios across early tissues and growth stages.

Solanum lycopersicum (SlPAD4) PAD4 produces truncated transcripts

In Soybean, the genomic structure of *PAD4* contains four exons and produces two distinct transcripts, one of which is the result of an exon skipping event. Similar to Soybean *PAD4,* Tomato *PAD4* is also predicted to contain four exons (Figure 16A, B), and for this reason, we used RT-PCR to investigate *PAD4* transcripts expressed in Tomato. To determine the occurrence of alternative splicing in SlPAD4, sequence-specific primers flanking*SlPAD4* (Table 2) were used to amplify transcripts within leaf tissue. A total of 3 distinct sequences were identified; (1) A full-length *SlPAD4* transcript (Solyc02g032850), (2) one lacking part of exon 3 (*SlPAD4-AS1*), and (3) one lacking the entirety of exon 3 (*SlPAD4-AS2*) (Figure 16B). *SlPAD4* results in a transcript that is 1,737 bp long. Based on sequence alignments it is possible that *SlPAD4-AS1* likely results from an alternate acceptor or alternative 3' splicing event, including only 118 bp of the 3' end of exon 3 in the mature transcript and not the complete exon 3. *SlPAD4-AS2* results from an exon skipping event, as evidenced by the mature transcript lacking all 627 bp of exon 3 in the mature transcript. Translated nucleotide sequences reveal that both

SlPAD4-AS1 and *SlPAD4-AS2* result in proteins and that no premature codons are introduced due to the splicing modifications (Figure 17). *SlPAD4* produces a protein that is 578 AA long, whereas *SlPAD4-AS1* is predicted to produce a protein that is 407 AA, and *SlPAD4-AS2* is predicted to produce a protein 369 AA long (Figure 16B, C, D; Figure 17). The conserved catalytic S-D-H residues were identified in SlPAD4 as S129, D188, and H275 (Figure 17, Figure 19). Further analysis of the predicted amino acid sequence reveals that both SlPAD4-AS1 and SlPAD4-AS2 lack part of the predicted lipase region (Figure 17; Figure 18) and the S-D-H catalytic residues (Figure 17). Our data suggests that *SlPAD4* undergoes two separate alternative splicing events and that in both events, the protein produced would lack part of the lipase region and the conserved catalytic residues.

Characterization of Transgenic 35S:GmPAD4 *Glycine max* Lines

Previous preliminary data indicated several lines of transformed Soybean (cv. Jack) in the T2 Generation had GmPAD4 insertions after biolistic transformation. In addition, these lines showed overexpression of GmPAD4. Subsequent soybean aphid fecundity assays were performed on T2 plants, which showed that total aphid count was significantly lower on transformed lines compared to the wild-type control. However, upon analysis of T3 plants from these lines, insertions were confirmed via PCR with PAD4.2 F and PAD4.2 R primers (Figure 20)(Table 2), but overexpression could not be validated via qPCR due to high Cq values (>29), regardless of steps taken for qPCR optimization. Furthermore, soybean aphid fecundity assays were carried out in the documented manner as performed on the T2 generation. No significant difference ($p>0.05$) was found at any timepoint when means from the E6-2 transformation line were compared through one-way ANOVA (for parametric datasets) or one-way Kruskal-Wallis

(for non-parametric datasets) to soybean cultivars Jack and Williams 82 (Figure 21). Additionally, in the T3 generation, the mean count of nymphs per adult aphid was also found to be non-significant ($p>0.05$) (Figure 21).

Phenotype data for the T2 Generation of overexpression soybean lines are not available. However, significant phenotypic differences were recorded between T3 overexpression and control lines (Williams 82 and Jack). These differences include decreased leaf size, plant height, and overall biomass (dwarfing) (Figure 22). In addition to these phenotypic changes, overexpression lines also displayed significantly decreased germination rates.

Generation of GmPAD4 Arabidopsis transformants

Generation of stable 35S:GmPAD4 and 35S:GmPAD4-AS1 Arabidopsis transformants was initially carried out for downstream use in pest/pathogen infection assays to investigate the role of GmPAD4 and GmPAD4-AS1 in plant response to biotic stressors. Additionally, the generation of 35S:GmPAD4:mCherry and 35S:GmPAD4-AS1:mCherry Arabidopsis transformants was initially planned to elucidate if there are differences in localization between GmPAD4 and GmPAD4-AS1. Confirmation of GmPAD4 constructs were carried out at each step (Figure 5, 6, 9). Creation of GmPAD4, GmPAD4-AS1, GmPAD4:mCherry and GmPAD4-AS1 mCherry Arabidopsis lines was achieved (Figure 10). After BASTA selection transformed plants were checked via PCR to confirm GmPAD4 insertions (Figure 10). However, these lines indicate that there are multiple constructs inserted and therefore require further characterization to ensure that the transformed lines have a single construct (Figure 10).

Discussion

Alternative splicing is a common but understudied phenomenon in plants. While alternative splicing events are documented in *R* genes, indicating a role in plant defense, the extent of alternative splicing in other plant defense regulatory components is understudied. To elucidate the role of alternative splicing in plant defense regulation, we focused on the highly conserved plant defense regulatory gene, *PAD4* (Baggs et al., 2020). *PAD4* appears to be uniquely positioned as a component of a regulatory hub that modulates plant responses to both biotic and abiotic stimuli.

PAD4 and its interacting partner(s) are highly conserved; despite this, the intron-exon structure of *PAD4* is highly variable between species (Figure 11). Comparative phylogenic analysis of *PAD4* protein sequences supports that *PAD4* is highly conserved across diverse species and orders (Figure 11). In addition, analysis at the order level reveals that *PAD4* intron-exon structure is highly conserved within several orders. Within the Fabales order, PAD4 is predicted to encode four exons consistently. In contrast, *PAD4* is predicted to encode 4-5 exons within the Rosales order. In species within the Brasssicales family, the typical number of *PAD4* exons is two, including Arabidopsis. Typically, outside the Brassicales order, *PAD4* is predicted to encode a transcript containing between three and five exons. Alternative splice events have not been documented in *AtPAD4*, making studying these events difficult in the model plant Arabidopsis.

Investigation of *PAD4* expression through BAR data indicates that *PAD4* expression is tissue-dependent (Figure 12). Species surveyed include Arabidopsis, Poplar, Tomato, Soybean, and Potato. All species, except for Potato, showed the highest expression of *PAD4* within root

tissues (Figure 12). *PAD4* has been cited as a regulatory component for the formation of lysigenous aerenchyma in Arabidopsis root tissue under hypoxic conditions (Mühlenbock et al., 2007). In addition, BAR expression data for Potato, Rice, and Arabidopsis under various abiotic stresses indicated that *PAD4* expression experiences significant fold changes under specific abiotic stressors such as genotoxic, oxidative, UV-B, and cold stress (Figure 13). These results are consistent with previous literature, which indicates that *PAD4* plays a role in plant response to excess light stress in Arabidopsis through redox signals transduction and promoting ET and ROS signaling (Mateo et al., 2004; Mühlenbock et al., 2008). *PAD4* expression changes have been documented under numerous abiotic stressors. Still, UV-B, heat stress, and light stress all lead to changes in redox status, suggesting that PAD4 involvement in these pathways is indeed involved with the potentiation of redox signals (Mateo et al., 2004; Mühlenbock et al., 2008; Bernacki et al., 2019). Experiments in Arabidopsis indicate that *pad4-1* knockout plants display increased tolerance to cold stress and experience a reduction in cell death (Chen et al., 2015). *PAD4* likely facilitates programmed cell death in response to cold stress through positive regulation of ROS and SA (Chen et al., 2015).

This work identifies two *PAD4* splice variants in Tomato, *SlPAD4-AS1,* and *SlPAD4-AS2* (Figure 16). Both transcripts are predicted to produce truncated proteins with no premature stop codons (Figure 17). *SlPAD4-AS1* appears to be the result of an alternate acceptor splicing event, while *SlPAD4-AS2* results from a rare exon skipping splicing event (Figure 17). In addition to *SlPAD4*, which is predicted to have four exons, Soybean *PAD4* is also predicted to have four exons (Selig, 2017). In Soybean, GmPAD4 undergoes a rare alternate splicing event, exon skipping, to produce a truncated transcript *GmPAD4-AS1* similar to *SlPAD4-AS2* (Selig, 2017).

Comparative analysis of AtPAD4, SlPAD4, and GmPAD4 amino acid sequences indicate that all three contain the conserved N-terminal lipase and C-terminal EP domains (Figure 18; 19) (Wiermer et al., 2005; Wagner et al., 2013). Within the *AtPAD4* lipase domain, an S-D-H catalytic triad was identified as S118, D178, and H229 (Wagner et al., 2013). Analysis of AtPAD4, SlPAD4, and GmPAD4 amino acid sequences allowed for the identification of the conserved S-D-H catalytic triad motif in both SlPAD4 and GmPAD4. In SlPAD4, the catalytic triad was identified as S129, D188, and H274 (Figure 19). Previous data from Selig 2017 indicated that the catalytic residues within GmPAD4 are S137, D199, and H224. However, alignments of AtPAD4, SlPAD4, and GmPAD4 suggest that the conserved H residue is H313 rather than H229, as all three amino acid sequences are highly similar around the conserved residues. In addition, both AtPAD4 and SlPAD4 lack the region that H299 resides in. All three sequences contain highly similar sequences around H313 (Figure 19). High conservation of the lipase region and the catalytic triad within suggests that this region may serve a function despite a lack of documented lipase activity. Previous work indicates thatAtPAD4 S118 is essential for defense against green peach aphids (Louis et al., 2012; Dongus et al., 2020). In addition, *GmPAD4* and *SlPAD4* have documented effects on aphid antibiosis and antixenosis in plant defense against soybean aphids and green peach aphids, respectively (Selig, 2017, Singh and Shah, 2012; Dongus et al., 2020). Taken altogether, the similarity between amino acid sequences, conservation of the catalytic triad, and previous literature that indicates *PAD4* involvement in aphid defense, I suggest that these results indicate a conserved role of *PAD4* in plant defense against aphids across species.

Analysis of *GmPAD4* and *GmPAD4-AS1* transcripts suggest that total *PAD4* expression is not a predictor of *GmPAD4-AS1* prevalence. In growth stage analysis, total *PAD4* was highest in the VE stage after seedling emergence. In contrast, the V1 stage, after trifoliate emergence, had the highest percentage of *GmPAD4-AS1,* with 46% of *PAD4* expression accounted for by the splice variant (Figure 14). These findings are not surprising as alternate splicing can be essential to plant development (Staiger & Brown, 2013). Further studies would be needed to elucidate if *GmPAD4* and *GmPAD4-AS1* impact soybean development. Tissue analysis at each growth stage shows that *GmPAD4* and *GmPAD4-AS1* expression varies throughout tissues at differing growth stages. These results reflect the findings of generalized growth stage analysis, which show that *GmPAD4* expression is highest in the V1 stage, in addition to *GmPAD4-AS1* having the highest percentage of expression (Figure 14, Figure 15).

Previous unpublished preliminary data indicated several lines of 35S:GmPAD4 overexpression lines created in the soybean cultivar, Jack. T2 soybean aphid assays indicated that transformed lines had reduced total aphid fecundity. However, soybean aphid assays in the T3 Generation did not show reduced total aphid fecundity. Validation of the 35S:GmPAD4 insertion was performed via DNA extraction, but overexpression could not be validated via qPCR. In addition, highly reduced germination of overexpression lines and phenotypic differences may indicate pleiotropies within the overexpression line. After characterization of these lines with soybean aphid assays and qPCR we concluded that these lines might be experiencing post-translational silencing (Flavell, 1994). Post-transcriptional silencing of transgenes is not uncommon in plants and can lead to the suppression of an endogenous homologous gene in addition to the transgene (Flavell, 1994).

Using the floral dip method; we attempted to generate transgenic Arabidopsis plants over-expressing GmPAD4 alone and with a 3' mCherry fusion. However, T1 Arabidopsis plants selected with BASTA showed multiple insertions with different constructs, and the identity of T1 plants could not be verified. Mitigating concerns of multiple construct insertions in the future can be done in various ways, (1) comprehensive cleaning during the floral dip procedure and complete separation of inoculated plants not just for the first three days but for the entirety of the growth period to ensure no-cross contamination between constructs, (2) use of primers to confirm plasmid insert characteristics (such as length to distinguish GmPAD4 and GmPAD4-AS1, such as the PAD4.2 primer set), not just the presence of an insert, (3) both complete plasmid sequencing and restriction enzyme digestions to confirm the integrity and identity of the vector. These lines will need to be carried into further generations to ensure a single insertion and check for pleiotropies from the transformation. In addition, observation of segregation ratios will allow the confirmation of a single insertion being passed down independently. After segregation ratios are confirmed in the T3 or T4 generation, the homozygous lines can be used for further experiments when all plants survive selection. These transformed Arabidopsis lines can be used for green peach aphid assays and localization studies to elucidate further the role of *GmPAD4-AS1* in defense against green peach aphids.

Here our findings indicate that the plant defense regulatory gene *PAD4* is both highly conserved and has diverse intron-exon structures, with 61% of species surveyed having predicted intron-exon structures that contain three to five exons. Identified *PAD4* splice variants in Tomato indicate two distinct types of alternative splice events, one of which is an uncommon exon-skipping event. Similarly, GmPAD4-AS1 shows a similar exon-skipping event.

AtPAD4, SlPAD4, and GmPAD4 all contain a conserved S-D-H catalytic triad, however the splice variants *SlPAD4-AS1*, *SlPAD4-AS2*, and *GmPAD4-AS1* all result in proteins that lack these conserved residues. PAD4 has a documented role in modulating the SA pathway and the JA/ET pathway in response to biotic stressors. In addition, *PAD4* regulates responses to abiotic stressors, likely through the potentiation of ethylene and redox signals. Because all splice variants lead to the removal of residues that are important to *PAD4*-mediated defense against aphids, we hypothesize that alternative splicing may help to balance the multifaceted regulatory hub formed by *PAD4*, *EDS1,* and *SAG101* and allow for *PAD4* to play a dual role in response to both abiotic and biotic stressors. These findings are especially interesting as the allocation of plant resources is dynamic and highly regulated. Often increases in plant yield come with a sacrifice in plant immunity efficiency (Karasov et al., 2017). Overcoming this trade-off in crops is especially important as crops ideally would produce a high yield and successfully defend against pathogens. Field trials with Arabidopsis and Poplar indicate that plants with mutations or silencing of *PAD4* did not show increased susceptibility to pathogens (Bernacki et al., 2019).

Conclusion

PAD4 and its interacting partners are uniquely positioned at the intersection of hormone regulatory and signaling pathways that allow it to modulate plant responses to biotic and abiotic stressors. This research provides evidence that alternative splicing within *PAD4* is conserved in the species *Solanum lycopersicum* and suggests that these alternative splicing events functionally impact the biologically relevant lipase domain of PAD4. In addition, studies of *PAD4* expression data indicate that *PAD4* expression levels are tissue-specific and responsive

to several abiotic stimuli. Investigation of soybean *PAD4* and its splice variant reveals that *PAD4* alternative splicing may also be modulated throughout growth stages and within tissues. A deeper understanding of the role of alternative splicing in *PAD4* may help elucidate how this hub influences many important plant defense and fitness characteristics. Changing climate conditions intensify crop stress by increasing abiotic stressors such as higher temperatures and increased drought. Additionally, changing conditions result in different growth conditions for pathogens and can allow pathogens to overwhelm host defenses. Studying *PAD4* and the hub that it helps to regulate may offer a path to modulating crop responses to these unfavorable conditions and allow crop loss to be mitigated.

Tables and Figures

Table 1: Prevalence of AS in diverse species

Common Name	Scientific Name	Total AS[a]	Percent IR	Percent ES	Percent AltD	Percent AltA	Reference
Arabidopsis	*Arabidopsis thaliana*	70%	40%	2.7%	7.5%	15.4%	Martín et al. (2021), Marquez et al. (2012)
Tomato	*Solanum lycopersicum*	65%	18.9%	6%	7.3%	12.9%	Clark et al. (2019)
Rice	*Oryza sativa*	33%	47.5%	25.4%	7.9%	14.6%	Zhang et al. (2010)
Cucumber	*Cucumis sativus*	54.4%	55.7%	4.7%	12.1%	22.3%	Guo et al. (2010)
Stiff Brome	*Brachypodium distachyon*	6.3%	55.5%	5%	8.8%	16.7%	Walters et al. (2013)
Cotton	*Gossypium raimondii*	32%	40%	10%	14%	25%	Li et al. (2014)
Soybean	*Glycine max*	63%	26.4%	8.8%	11.2%	14.8%	Shen et al. (2014)
Grape	*Vitis vinifera*	44.6%	43.3%	5.9%	17.3%	23%	Potenza et al. (2015)

Corn	*Zea mays*	40%	58%	39%	26%	29%	Thatcher et al. (2014)
Barley	*Hordeum vulgare*	51%	54%	4%	7%	17%	Panahi et al. (2015)
Poplar	*Populus alba*	13.5% [b]	43.8%	8.6%	23.7%	23.7%	Wang et al. (2020)
Cassava	*Manihot esculenta*	31.6% [c]	22%	18%	-	-	Li et al. (2019)

[a] Percentage of AS in multi-exon genes

[b] Percentage of AS in expressed genes

[c] Percentage of AS in protein coding genes

Table 2: List of Primers

Name	Species	Sequence (5' -> 3')	Amplicon	Application
SlPAD4F	*Solanum lycopersicum*	ATGGAATCGGAAGCTTCATCGTTCG	Various	Cloning
SlPAD4R	*Solanum lycopersicum*	TCAAGGAAACTGAGGTTGGAGCAGCTG	Various	Cloning
PAD4.2F	*Glycine max*	CCTCTGTTTTCGTCTCGGC	Various	Various
PAD4.2R	*Glycine max*	AGCATTCAGGGTTGGTGAAG	Various	Various
SlGAPDH_F	*Solanum lycopersicum*	CTGCTCTCTCAGTAGCCAACA	157 bp	Housekeeping Primer
SlGAPDH_R	*Solanum lycopersicum*	CTTCCTCCAATAGCAGAGGTT	157 bp	Housekeeping Primer
M13-F (20)	-	GTAAAACGACGGCCAGT	Various	Cloning
M13-R (24)	-	AACAGCTATGACCATG	Various	Cloning
GmFBoxF	*Glycine max*	AGATAGGGAAATTGTGCAGGT		Housekeeping Primer
GmFboxR	*Glycine max*	CTAATGGCAATTGCAGCTCTC		Housekeeping Primer
attB1-PAD4 F	-	GGGGACAAGTTTGTACAAAAAAGCAGGCTAAGGCGCATGGCTTCCAACGAAACTTCA	Various	Colony PCR, Sanger Sequencing

PAD4Seq1R	-	CCGGAGAAAGCCACATACAC	Various	Colony PCR
PAD4Seq1F	-	CCAACGAAACTTCACCGTTTG	NA	Sanger Sequencing
PAD4Seq2F	-	CCTGCCTCTGTTTTCGTCTC	NA	Sanger Sequencing
PAD4Seq3F	-	AATGCCAAGGTTGCTCTTTG	NA	Sanger Sequencing
PAD4Seq4F	-	GAAGGAGCAGTGTGTGTGGATAG	NA	Sanger Sequencing
PAD4Seq5F	-	CAGAGCAACTCCATGCAAAAG	NA	Sanger Sequencing
PAD4Seq6F	-	CAGGGACTCTTCCAGTTCCA	NA	Sanger Sequencing
PAD4Seq7F	-	GAAGGGGATGCATAGAACCA	NA	Sanger Sequencing
PAD4Seq8F	-	GAAGCAAGGGACTGGTTGAA	NA	Sanger Sequencing

Table 3: Exon number is conserved across orders

Order	Most Common Exon Number	% of Analyzed Sequences with Most Common Exon Number	Total Number of Sequences Analyzed[a]
Arecales	4	100%	2
Brassicales	2	94%	49
Fabales	4	100%	4
Fagales	4	100%	2
Malpighiales	4	80%	5
Poales	4	67%	6
Rosales	4	80%	5

Table 4: BAR data for PAD4 expression analysis

Species	Gene ID	Data Type	Normalization	Source
Arabidopsis (Tissue)	AT3G52430	Affymetrix ATH1	GCOS	Schmid et al., 2005, Nature Genetics 37:501 and Nakabayashi et al., 2005, The Plant Journal, Vol 41:697
Arabidopsis (Stress)	AT3G52430	Affymetrix ATH1		Kilian et al., 2007, The Plant Journal, 50:347
Poplar	POTRI.007G100600	Affymetrix expression data	GCOS	Campbell Laboratory
Soybean	GLYMA.08G002100	mRNA-Seq		An integrated Transcriptome Atlas of the Crop Model Glycine Max, and its use in Comparative Analysis in Plants: Libault, M. et al (2010). The Plant Journal 63, 86-99
Tomato	SOLYC02G032850	Illumina-derived	RPKM	The tomato genome sequence providing insights into fresh fruit evolution: Tomato Genome Consortium. Nature 2012 485 (7400): 635-641
Potato	PGSC0003DMG400019873			The Transcriptome of the Reference Potato Genome Solanum tuberosum Group Phureja Clone DM1-3 516R44. Massa et al. (2011) PLoS ONE https://doi.org/10.1371/journal.pone.0026801 and Genome sequence and analysis of the tuber

				crop potato. The Potato Genome Sequencing Consortium (2011) Nature 475: 189–195
Rice	LOC_OS11G09010	Affymetrix ATH1	GCOS	Schmid et al., 2005, Nature Genetics 37:501 and Nakabayashi et al., 2005, The Plant Journal, Vol 41:697

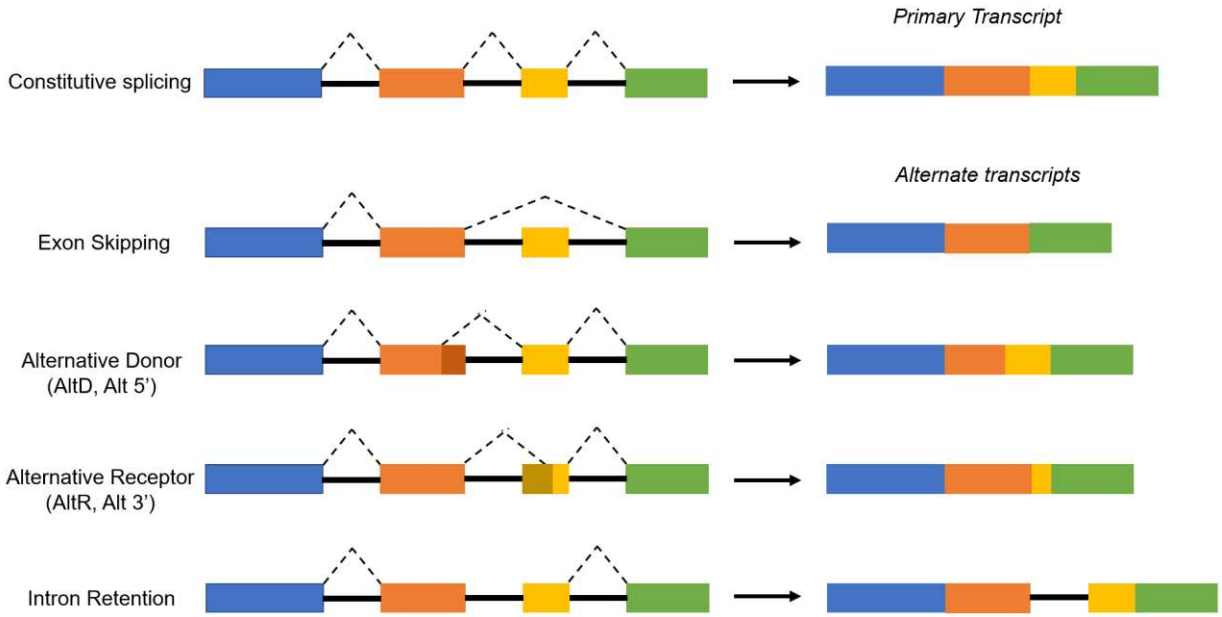


Figure 1: Alternative splicing event types The most common AS events include Exon Skipping (ES), Alternative Donor (AltD), Alternative Receptor (AltR) and Intron Retention (IR). A constitutive splicing event is depicted on the top, with alternate splicing events depicted below. An unspliced transcript is shown on the left, with colored boxes denoting exons and black lines denoting introns. Splicing events are indicated with dashed lines. The spliced transcript is shown on right with the splicing changes indicated to the left.

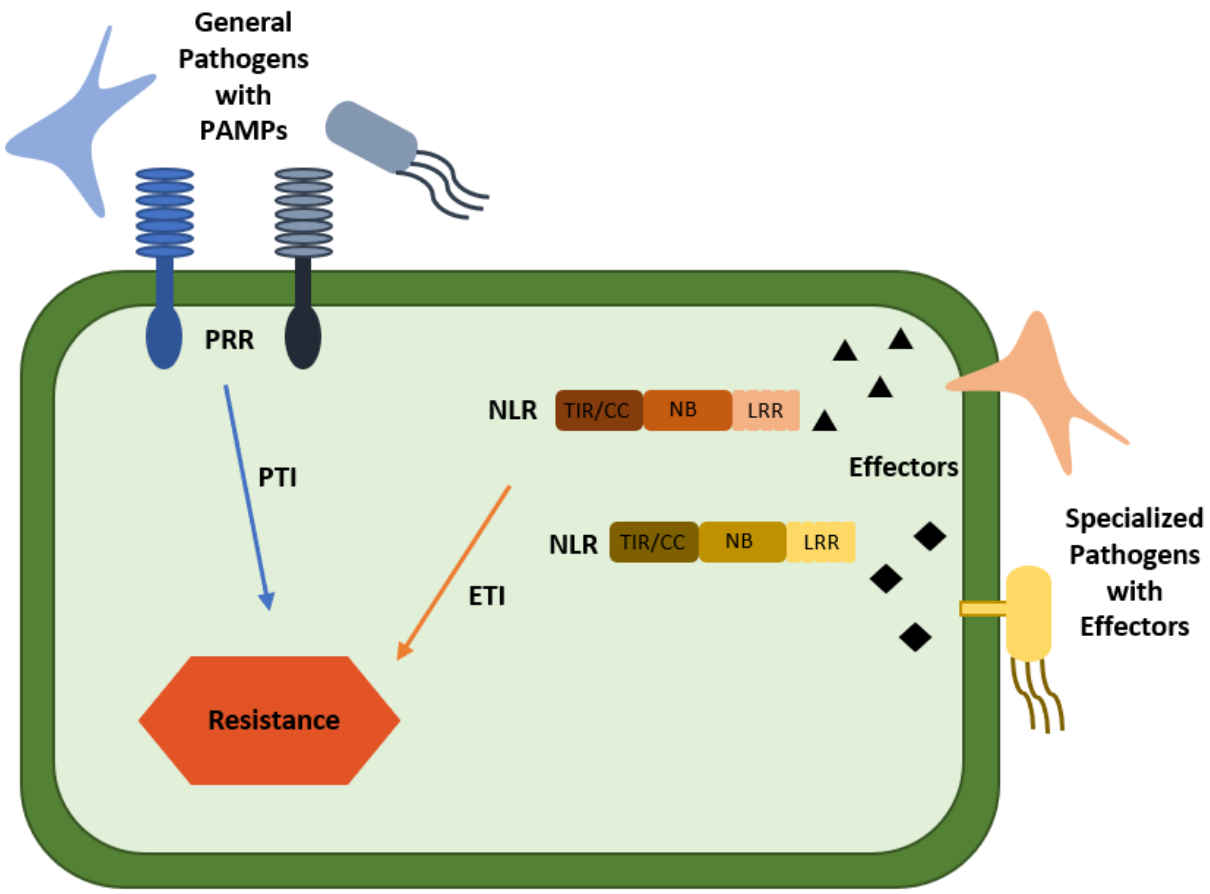


Figure 2: PTI and ETI induce plant defenses Figure depicts the perception of general pathogens with PAMPs through PRRs and the perception of specialized pathogens with effectors through NLRs.

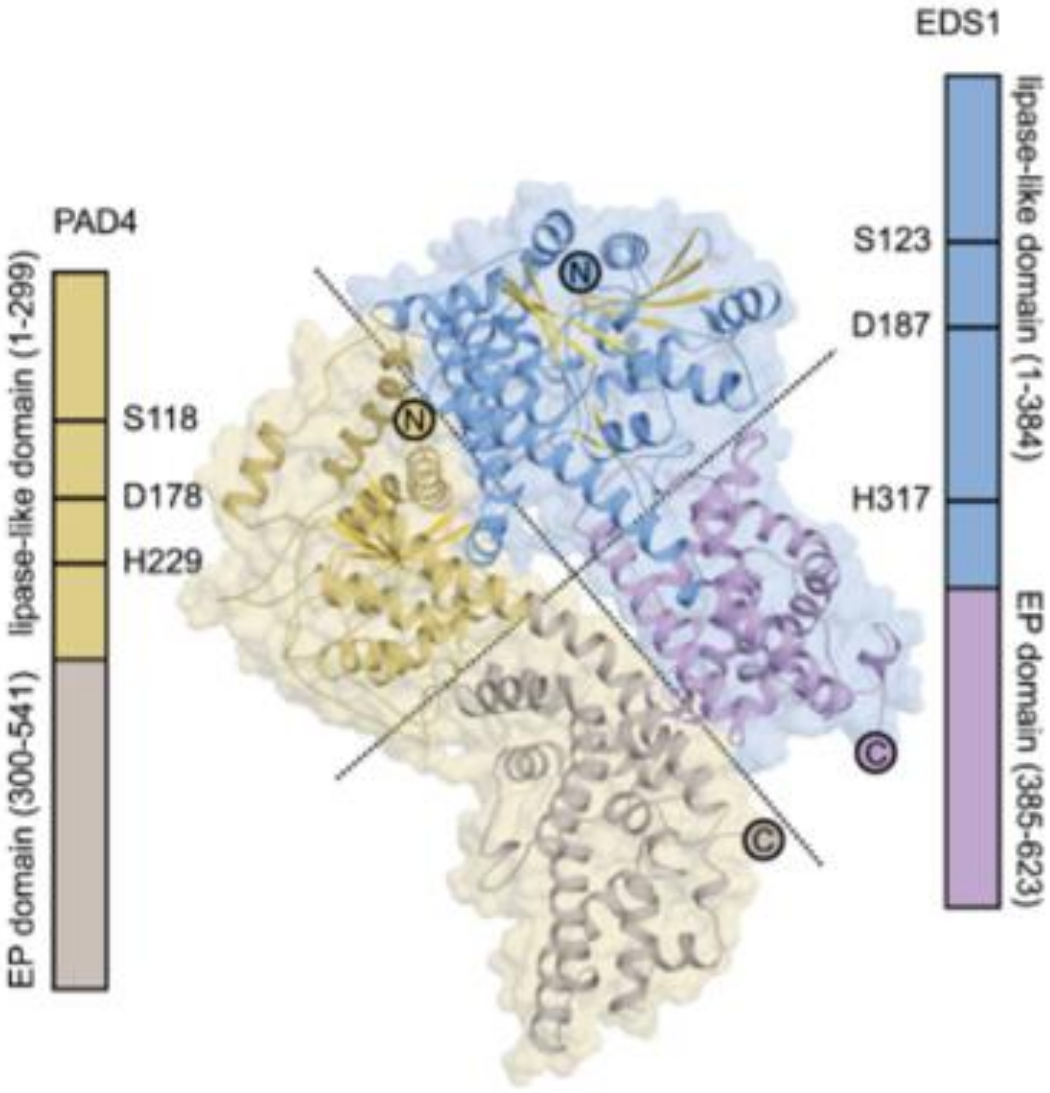


Figure 3: PAD4-EDS1 Heterodimer Image and model from Wagner et al., 2013. PAD4 protein is depicted on the left, and the EDS1 protein on the right.

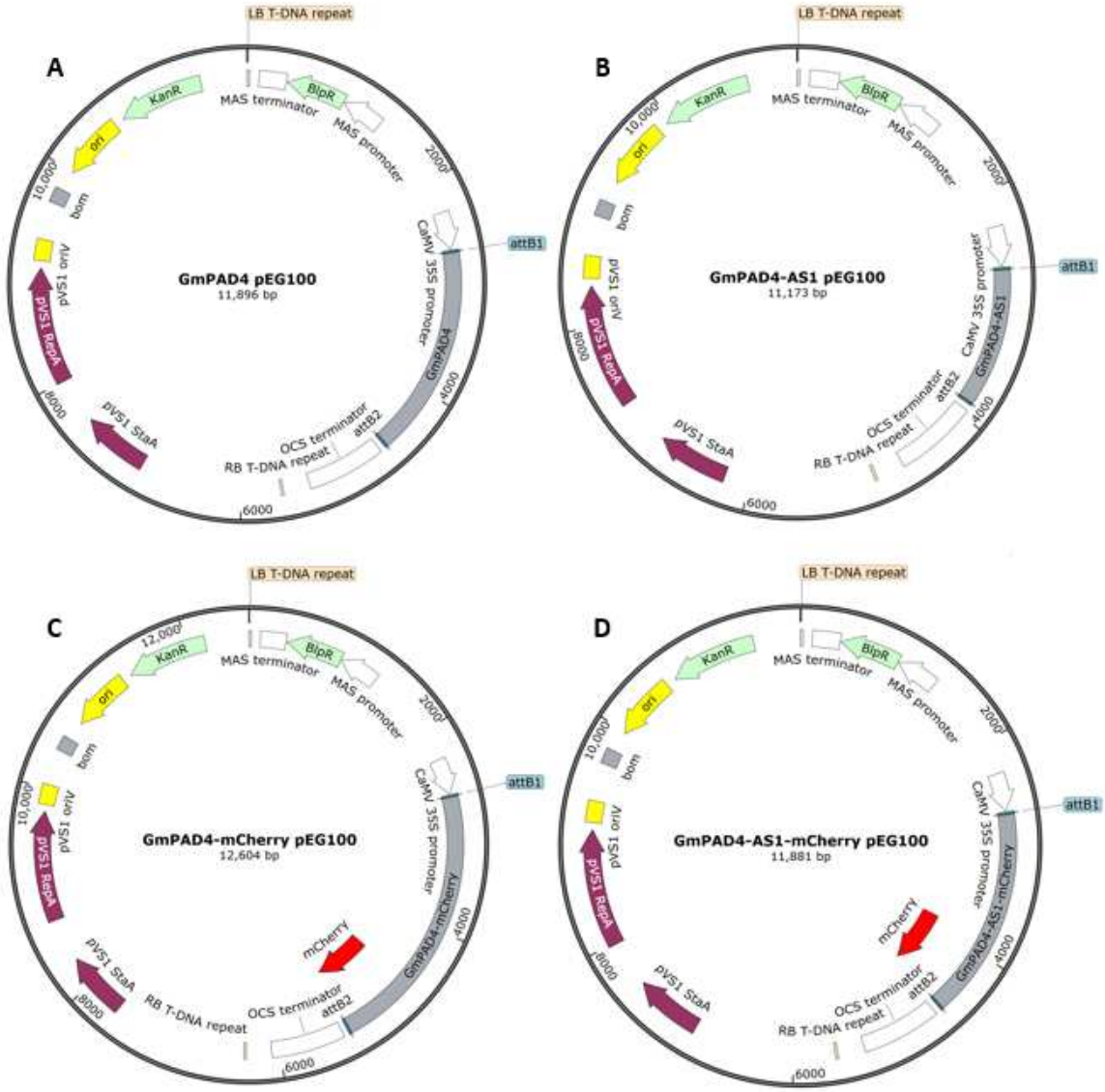


Figure 4: GmPAD4 pEARLEYGATE 100 constructs Complete pEG100 vector shown. GmPAD4 insertions in grey and mCherry fusion in red. A) GmPAD4 pEG100 construct B) GmPAD4-AS1 pEG100 construct C) GmPAD4:mCherry pEG100 construct D) GmPAD4-AS1:mCherry construct.

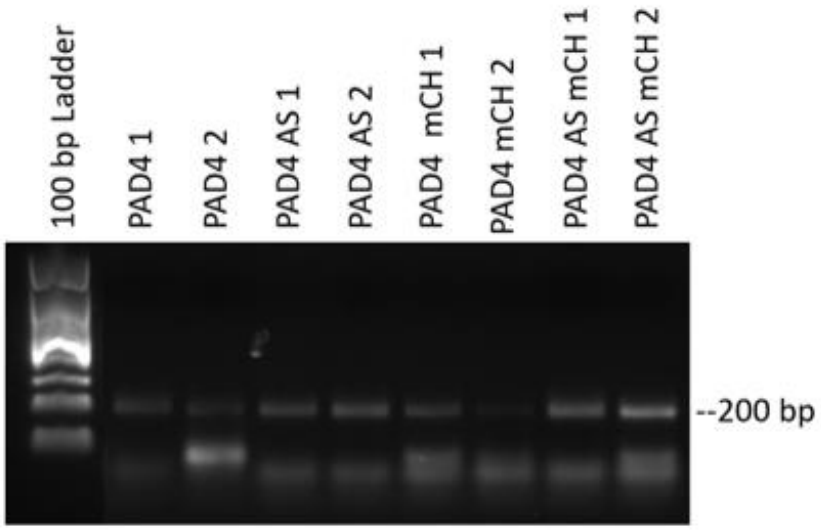


Figure 5: Colony PCR Confirmation of pTWIST pENTR GmPAD4 construct transformation in *E. coli* Expected product size for confirmation of transformation was 179 bp with the primer pair attB1-PAD4 F and PAD4seq1R.

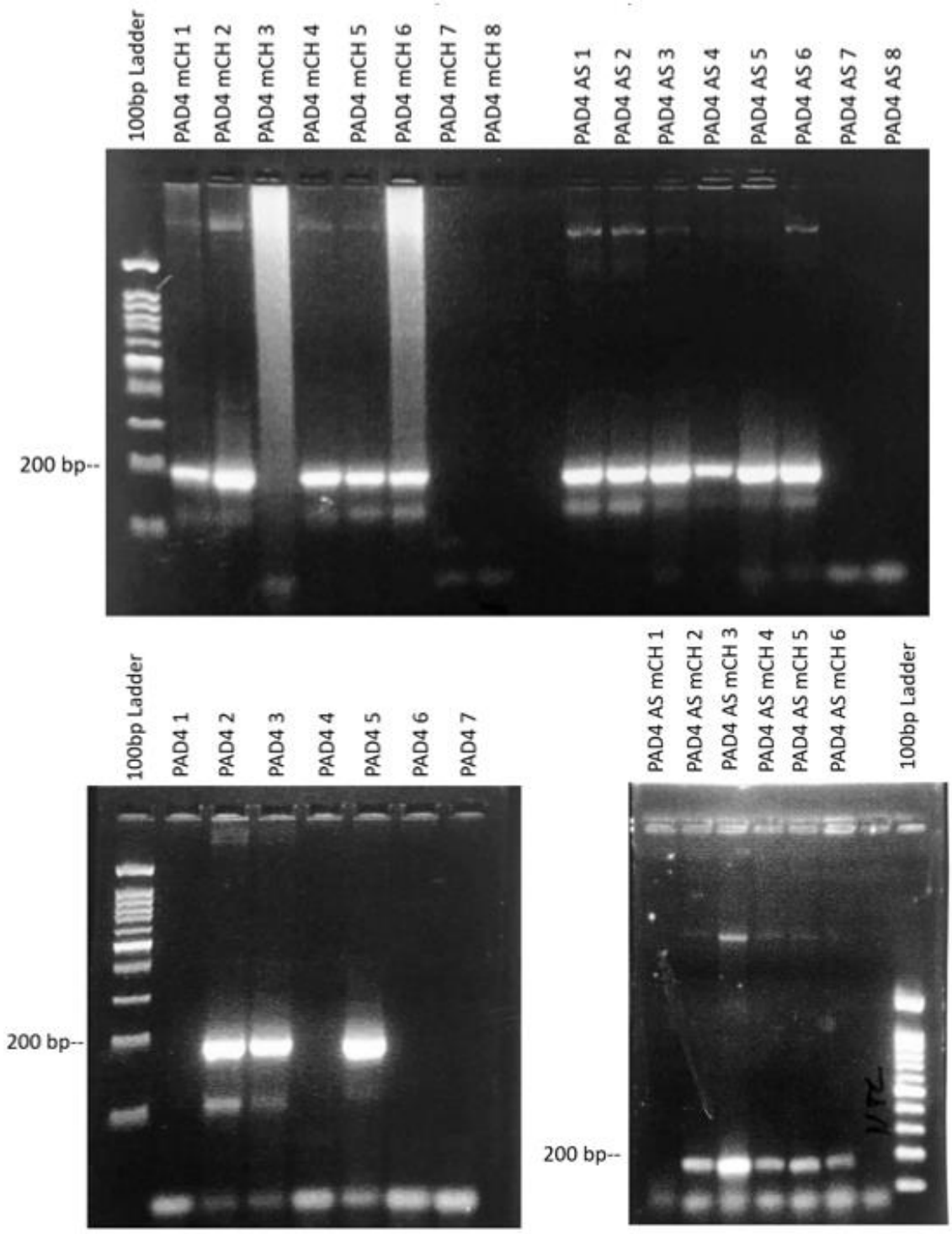


Figure 6: Colony PCR Confirmation of pEarleyGate 100 GmPAD4 construct transformation in *E. coli* Expected product size for confirmation of transformation was 179 bp with the primer pair attB1-PAD4 F and PAD4seq1R. A) pEG100 GmPAD4:mCherry and GmPAD4-AS1 colony PCR B) GmPAD4 colony PCR C) GmPAD4-AS1:mCherry colony PCR

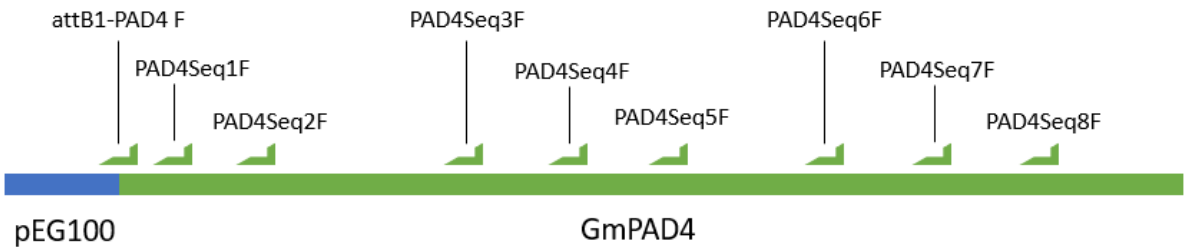


Figure 7: Schematic depicting primers used for sequencing of GmPAD4 insertions in pEG100 plasmid backbone Sequencing was performed via sanger sequencing

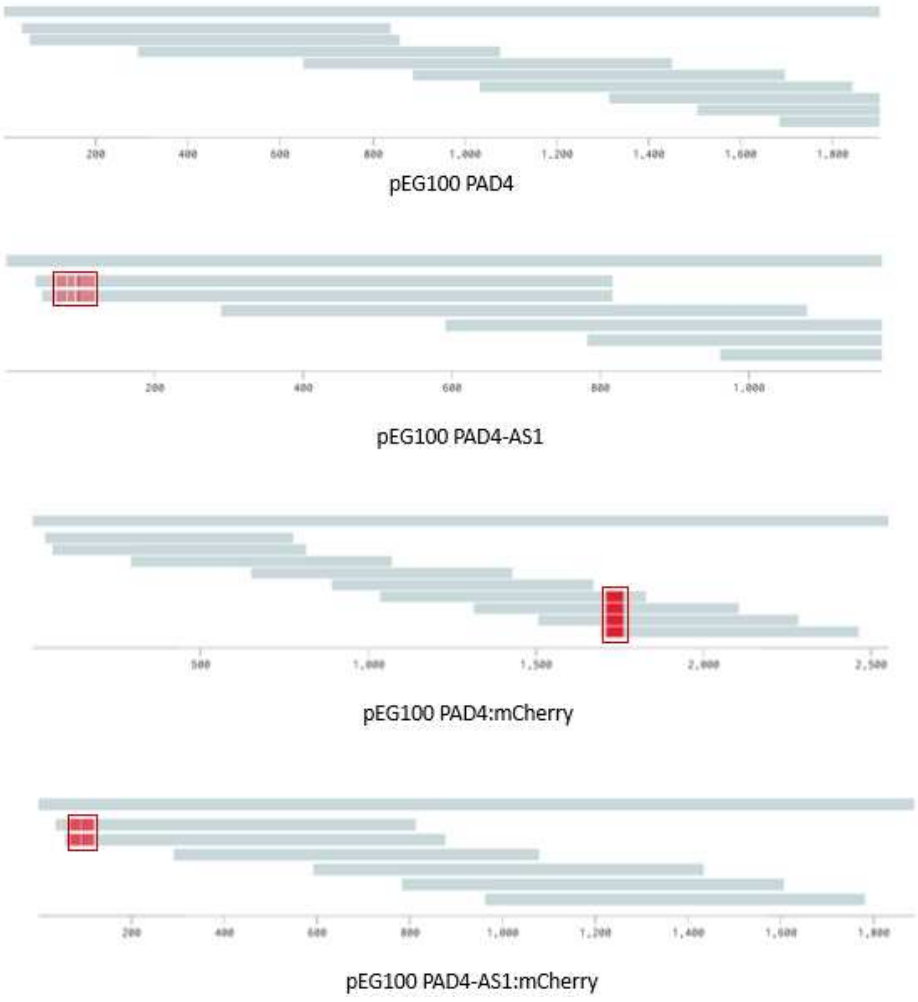


Figure 8: GmPAD4 construct sequencing results aligned sequencing was performed via sanger sequencing, and alignments were produced on the Benchling platform. Red boxes indicate regions that were modified by Twist Biosciences during construct synthesis and were reported not to make protein alterations.

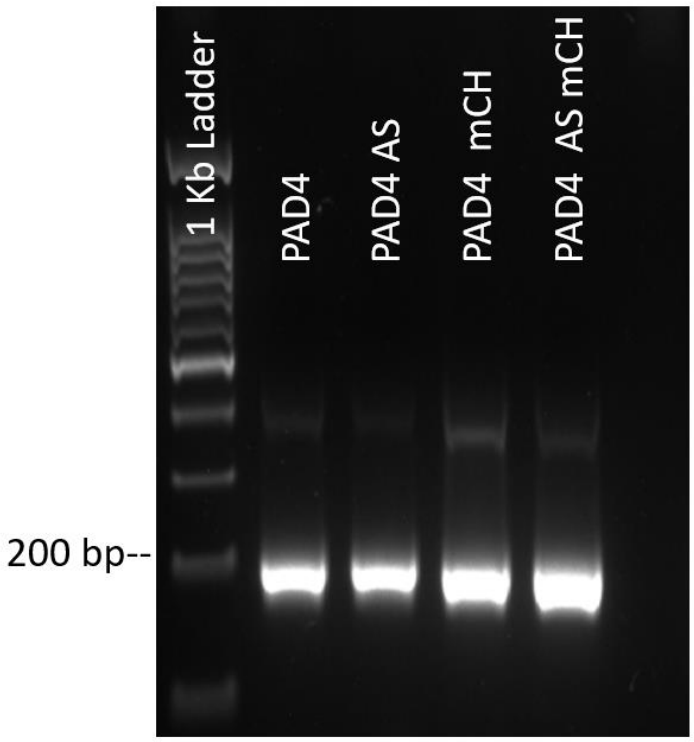


Figure 9: Colony PCR Confirmation of pEarleyGate 100 GmPAD4 construct transformation in *Agrobacterium* strain GV3101 Expected product size for confirmation of transformation was 179 bp with the primer pair attB1-PAD4 F and PAD4seq1R.

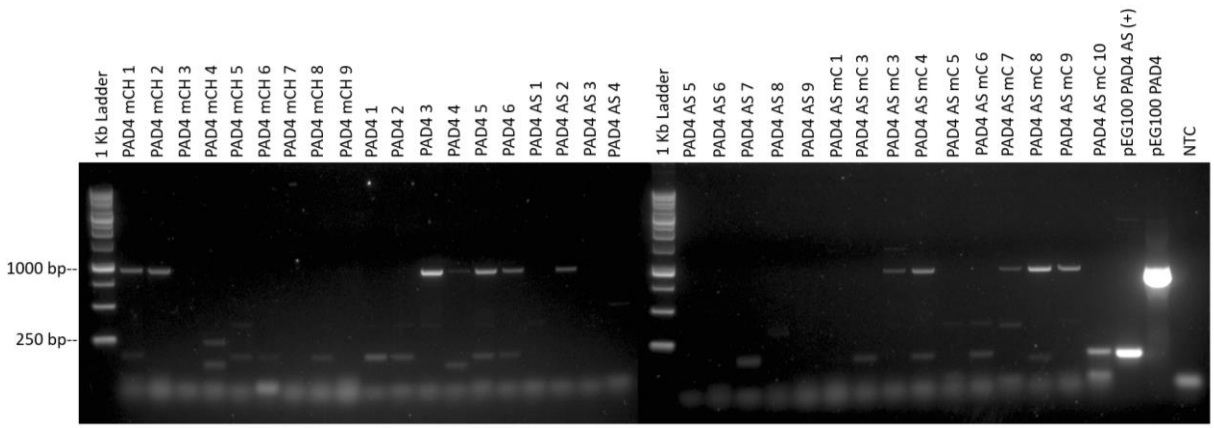


Figure 10: Characterizing GmPAD4 insertions in T1 Arabidopsis plants Expected product size for confirmation of transformation after floral dip was 905 bp for GmPAD4 insertions and 185 bp for GmPAD4-AS1 insertions.

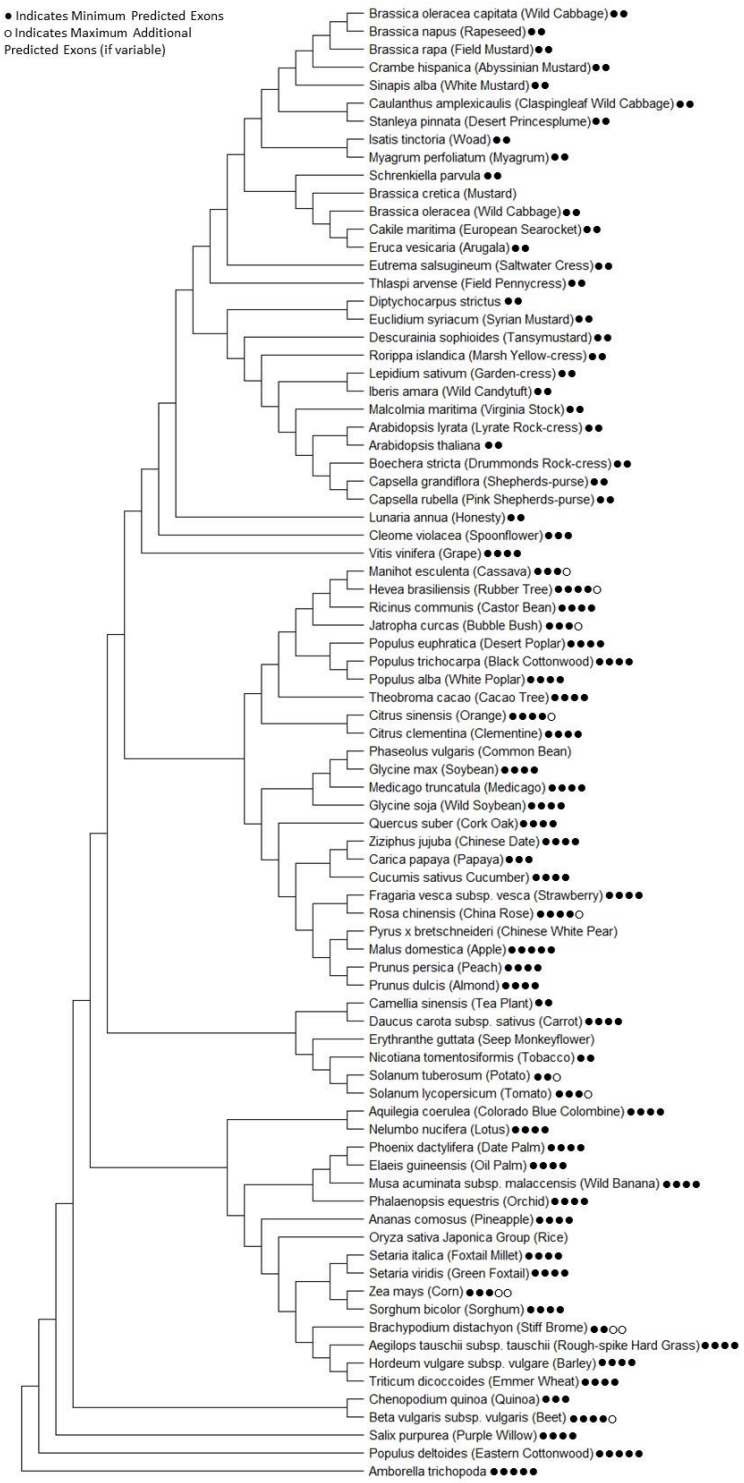


Figure 11: Phylogenetic analysis of PAD4 protein sequences. Sequences were aligned using MEGA-X with the maximum likelihood model. The predicted exon number is on the right, with each black dot representing an exon. Proteins with multiple exon number predictions include outlined circles to indicate the additional exons predicted. Proteins with no exon prediction have no circles included to the right.

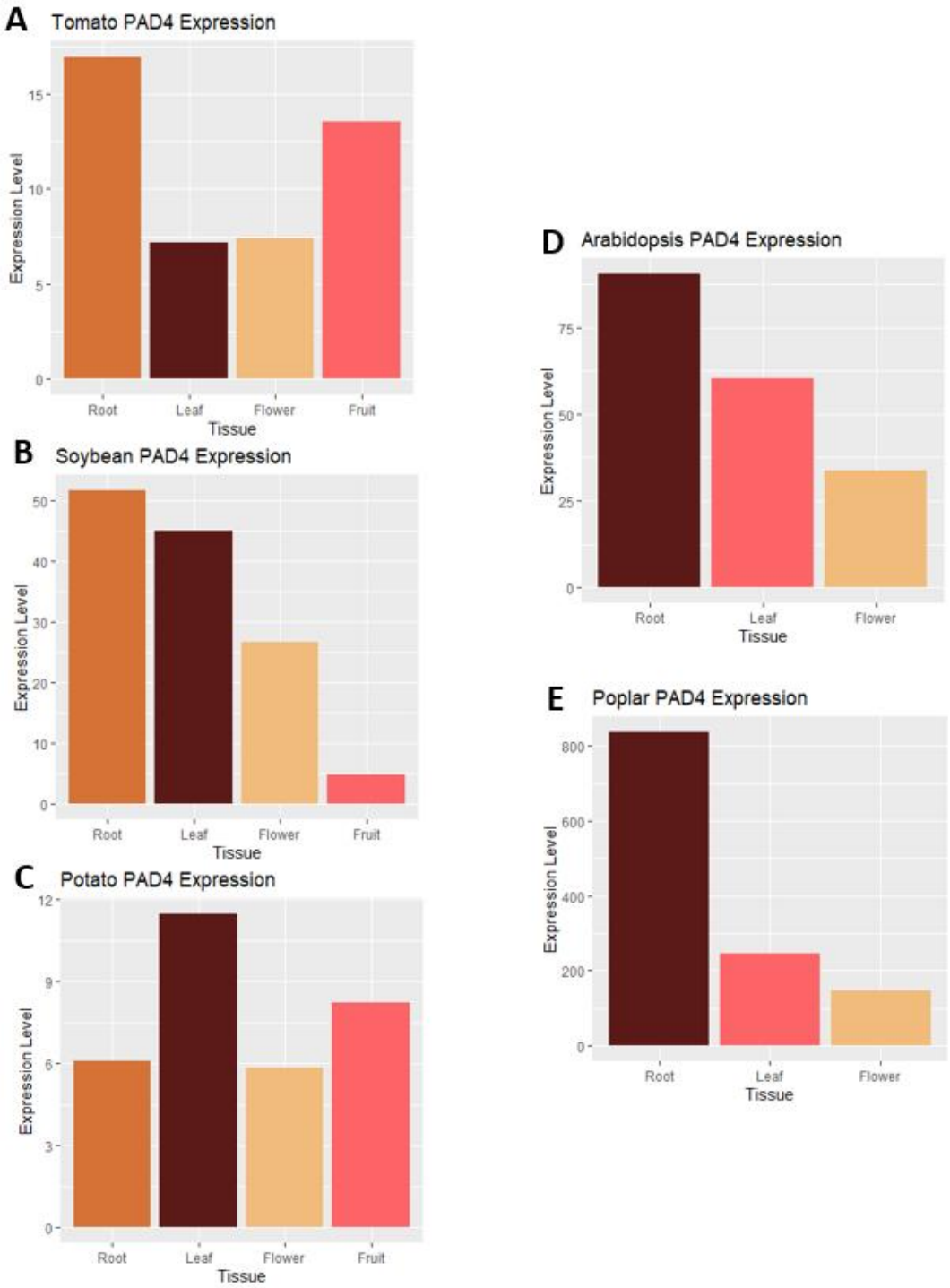


Figure 12: *PAD4* expression is tissue specific Expression data collected from BAR expression viewer (https://bar.utoronto.ca/) (A) SlPAD4 expression in Tomato tissues (B) *GmPAD4* expression in soybean tissues C) *StPAD4* expression in Potato tissues D) *AtPAD4* expression in Arabidopsis tissues E) Poplar *PAD4* expression in Poplar tissues

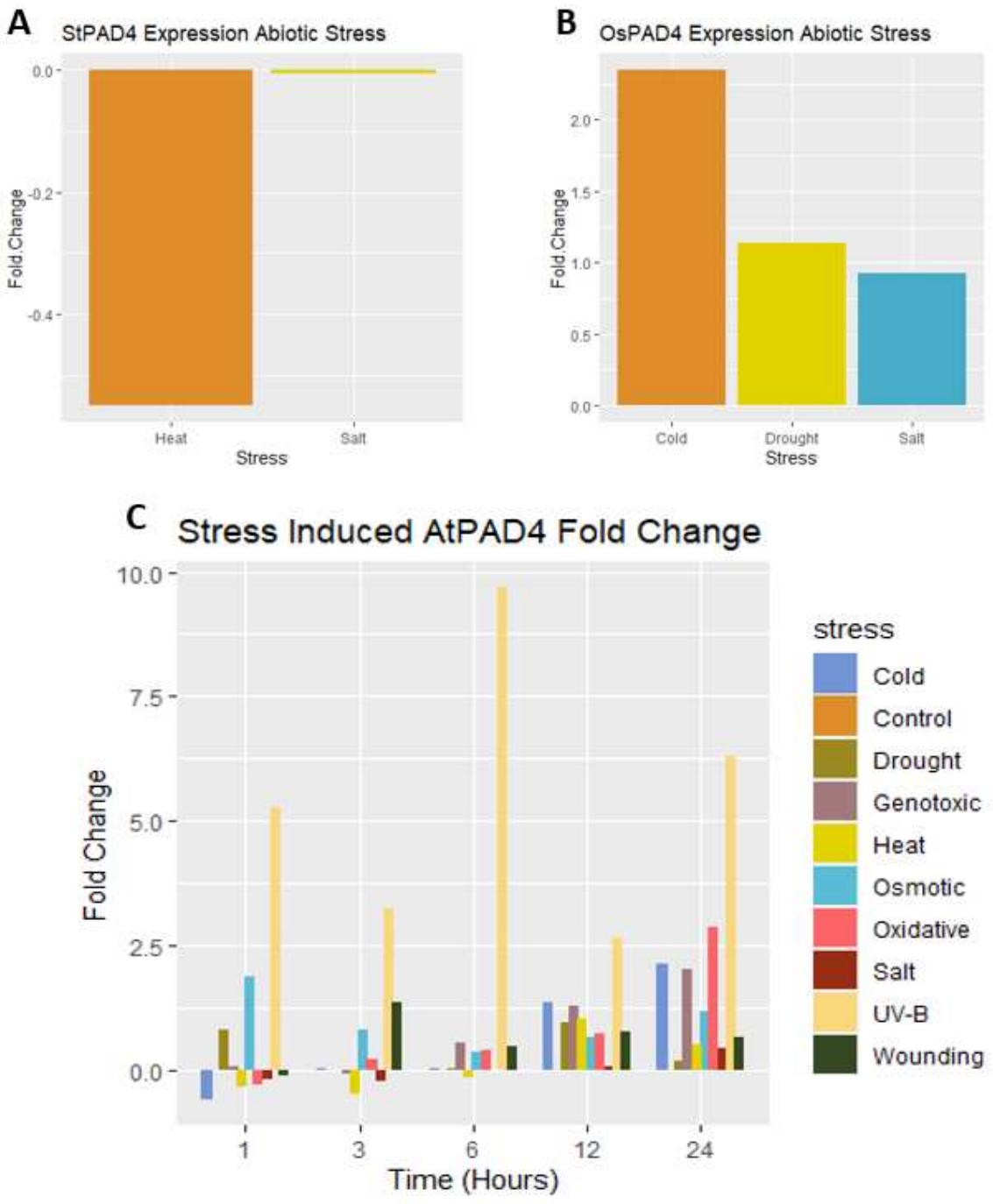
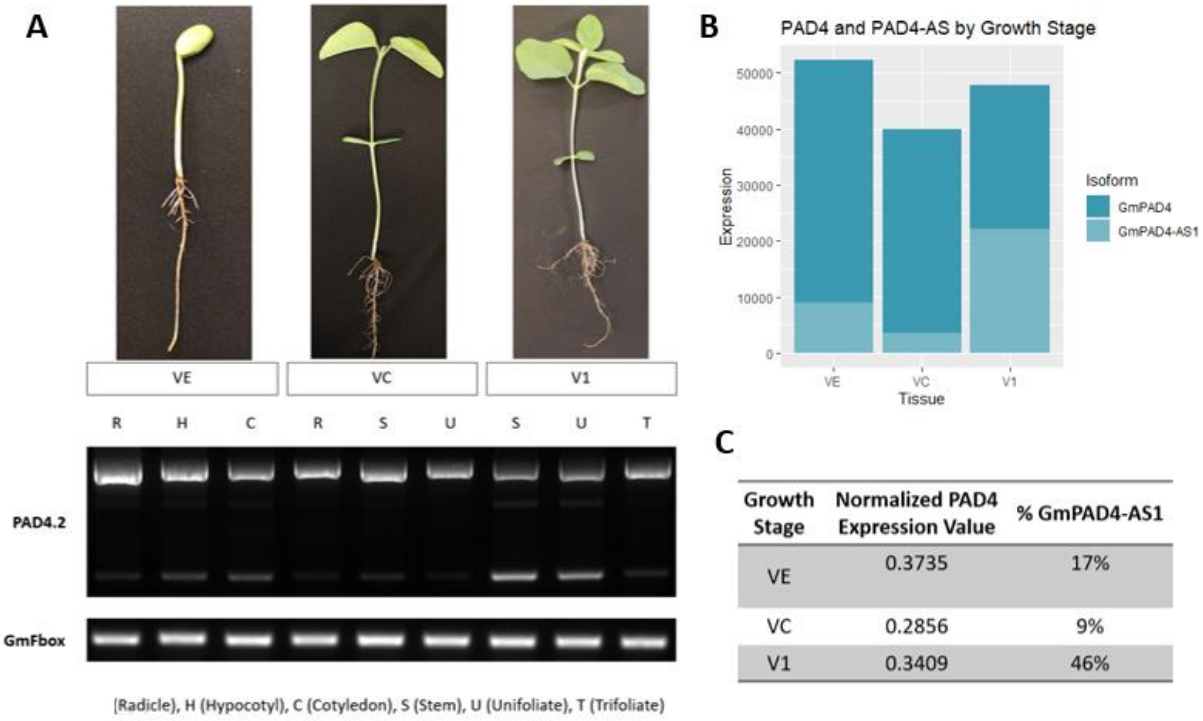


Figure 13: Abiotic stress-induced changes in PAD4 expression Expression data collected from BAR expression viewer (https://bar.utoronto.ca/). Expression changes are depicted as fold changes. (A) Potato *PAD4* expression changes after abiotic stress (B Rice *PAD4* expression changes after abiotic stress C) Arabidopsis *PAD4* changes after abiotic stress



Growth Stage	Normalized PAD4 Expression Value	% GmPAD4-AS1
VE	0.3735	17%
VC	0.2856	9%
V1	0.3409	46%

Figure 14: Semi-quantitative analysis of GmPAD4 and GmPAD4-AS1 expression by growth stage Quantification of gel images was performed in ImageJ A) Plants in the growth stages VE, VC, and V1 with corresponding RT-PCR results for GmPAD4 and GmPAD4-AS1 amplification B) Bar graph with quantified GmPAD4 and GmPAD4 AS1-expression by growth stage C) Table that includes normalized GmPAD4 expression and percentage of total expression that was GmPAD4-AS1 by growth stage.

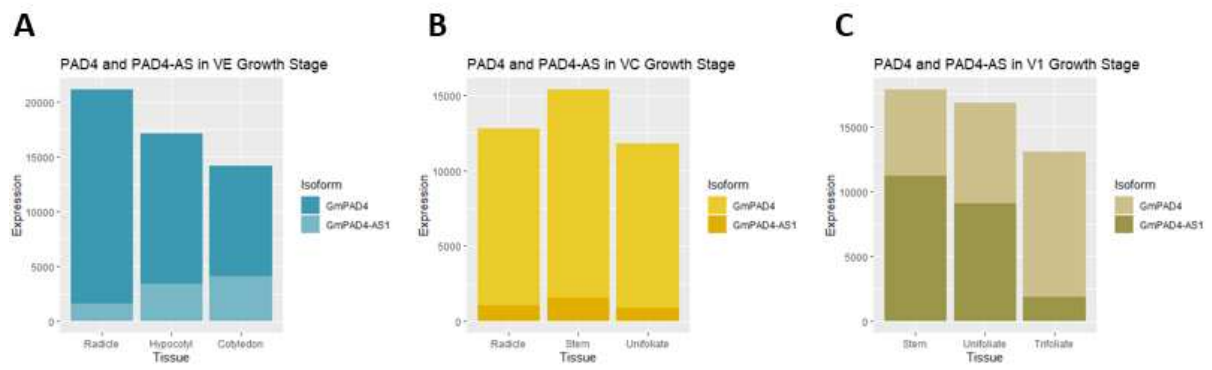
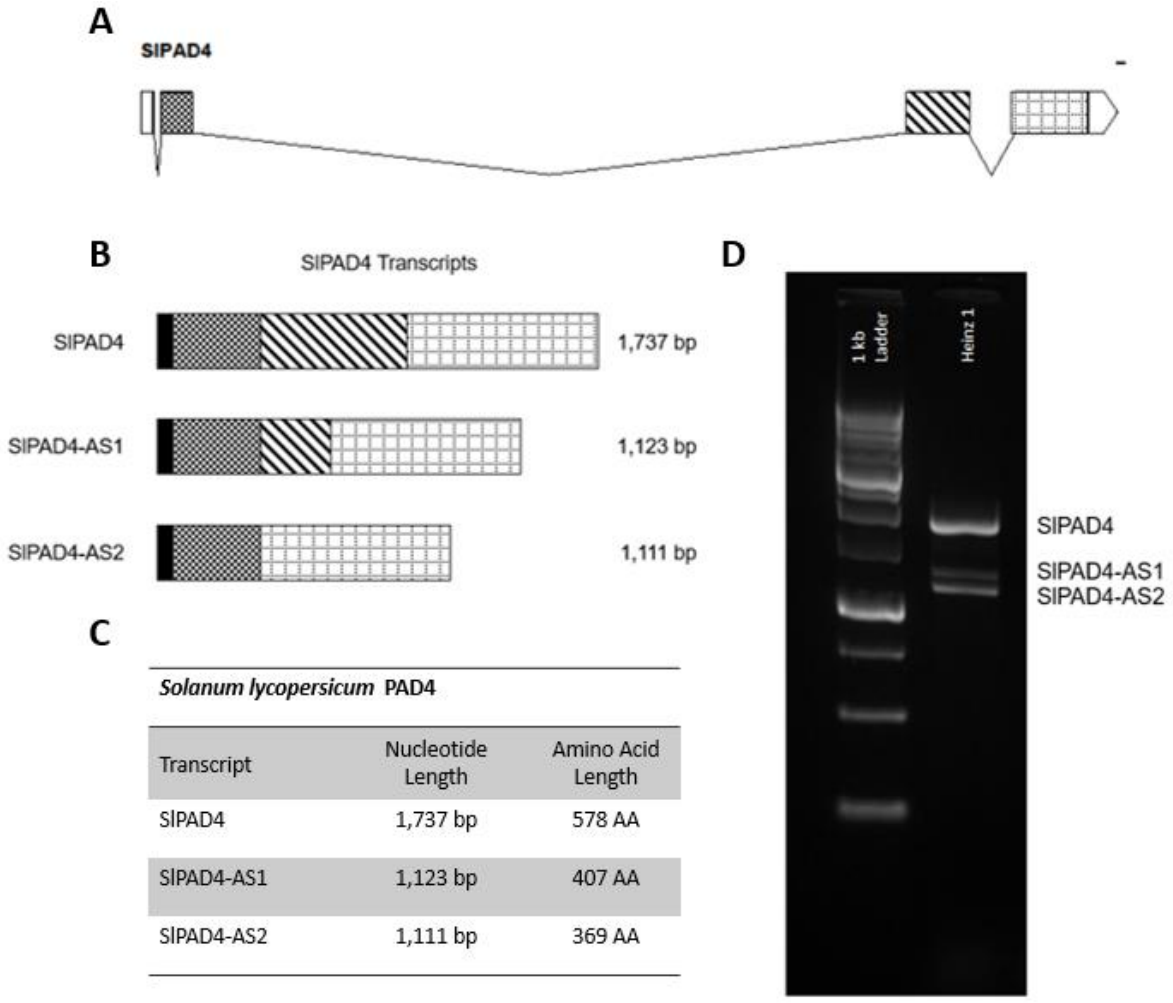


Figure 15: Semi-quantitative analysis of GmPAD4 and GmPAD4-AS1 expression by tissue
Quantification of gel images was performed in ImageJ. Quantified GmPAD4 and GmPAD4 AS1-expression. A) Expression in VE growth stage B) Expression in VC growth stage C) Expression in V1 growth stage



Solanum lycopersicum PAD4

Transcript	Nucleotide Length	Amino Acid Length
SlPAD4	1,737 bp	578 AA
SlPAD4-AS1	1,123 bp	407 AA
SlPAD4-AS2	1,111 bp	369 AA

Figure 16: *Solanum lycopersicum* PAD4 (*SlPAD4*) produces truncated transcripts A) Gene structure of *SlPAD4* depicting four exons and three introns B) Schematic of mature transcript for full length and alternatively spliced transcripts in *SlPAD4* C) Table including splice variant length in bp and length of resulting protein in AA D) Gel image depicting full-length *SlPAD4*, *SlPAD4-AS1,* and *SlPAD4-AS2*

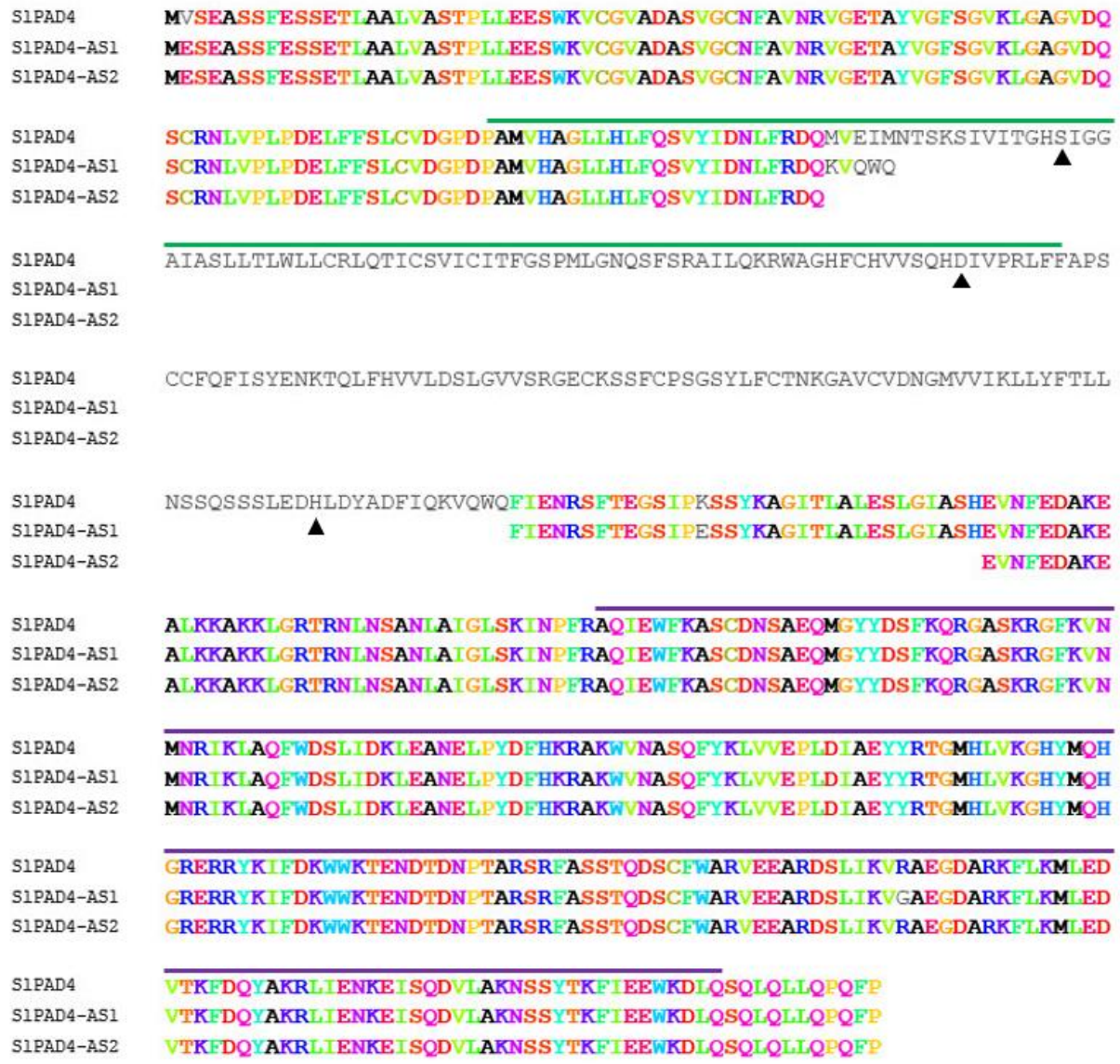


Figure 17: *SlPAD4* splice variants produce truncated proteins. Amino acid alignment was performed with EMBL-EBI Clustal Omega Multiple Sequence Alignment and STRAP alignment annotation (Gille et al., 2014). Black triangles indicate the conserved catalytic residues S129, D188, and H275. The green line indicates the predicted lipase domain, and the purple line indicates the predicted EP domain.

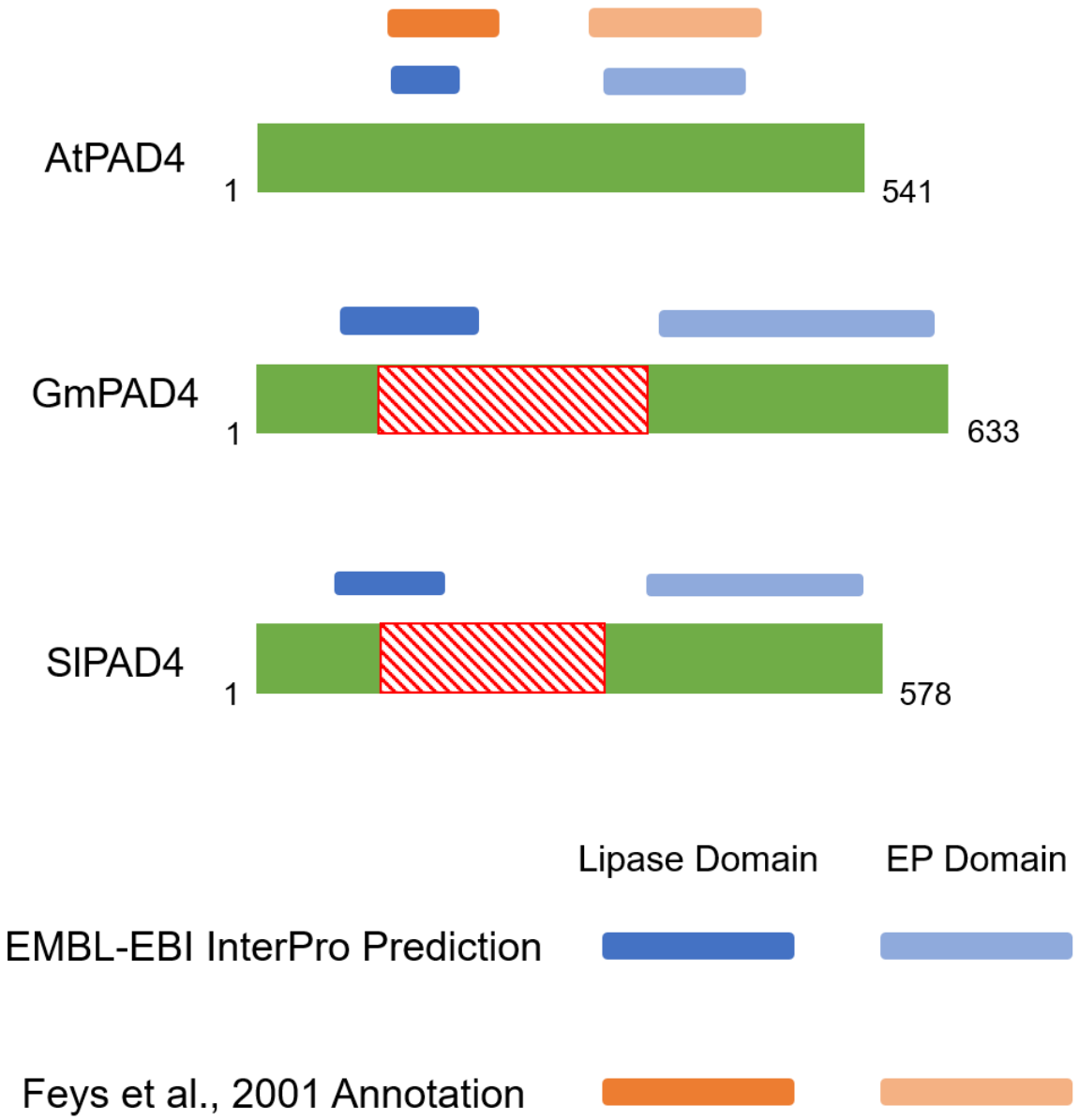


Figure 18: SlPAD4, GmPAD4 and AtPAD4 have conserved LP and EP domains Figure depicts AtPAD4, GmPAD4 and SlPAD4 with domain annotations provided from Feys et al., 2001 in orange and annotations predicted by EMBL-EBI InterPro in blue. Red regions indicate regions that were spliced out in GmPAD4-AS1 and SlPAD4-AS2.

Figure 19: SlPAD4, GmPAD4, and AtPAD4 share conserved catalytic residues. Amino acid alignment was performed with EMBL-EBI Clustal Omega Multiple Sequence Alignment and STRAP alignment annotation (Gille et al., 2014). Annotations for AtPAD4 catalytic residues are from Wagner et al., 2013. Black boxes indicate conserved catalytic residues S-D-H. The red triangle indicates H224, previously identified by Selig 2017.

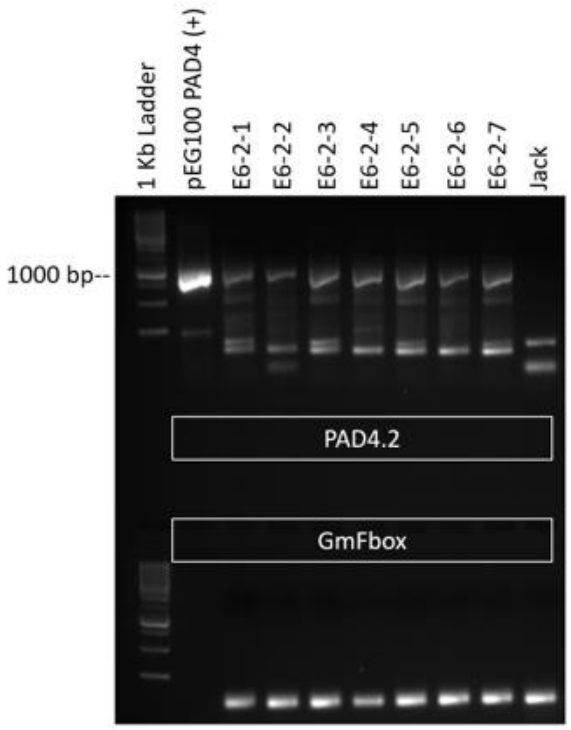


Figure 20: Confirmation of 35S:GmPAD4 insertion in *Glycine max* overexpression line E6-2 Gel image shows confirmation of 35S:GmPAD4 insertion in the E6-2 line of *Glycine max.* Expected product size for confirmation was 900 bp.

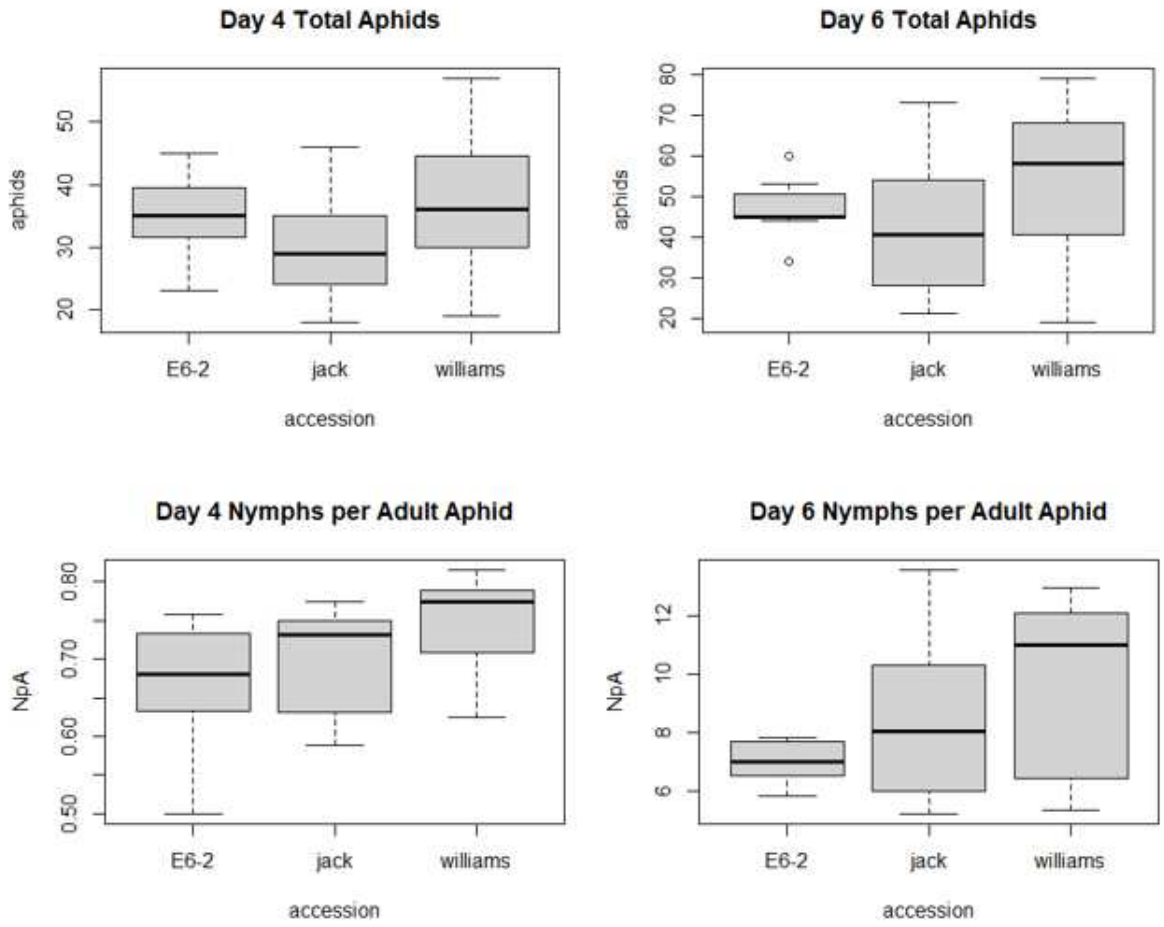


Figure 21: Soybean Aphid Assay Adult and nymph soybean aphids counted each day. Mean differences were calculated using one-way ANOVA, no significant differences in mean total aphids or nymphs per adult aphid between soybean lines. E6-2 n=7, Jack n=8, Williams n=10 A) Day 4 mean total aphids B) Day 6 mean total aphids C) Day 4 mean nymphs per adult aphid D) Day 6 mean nymphs per adult aphid

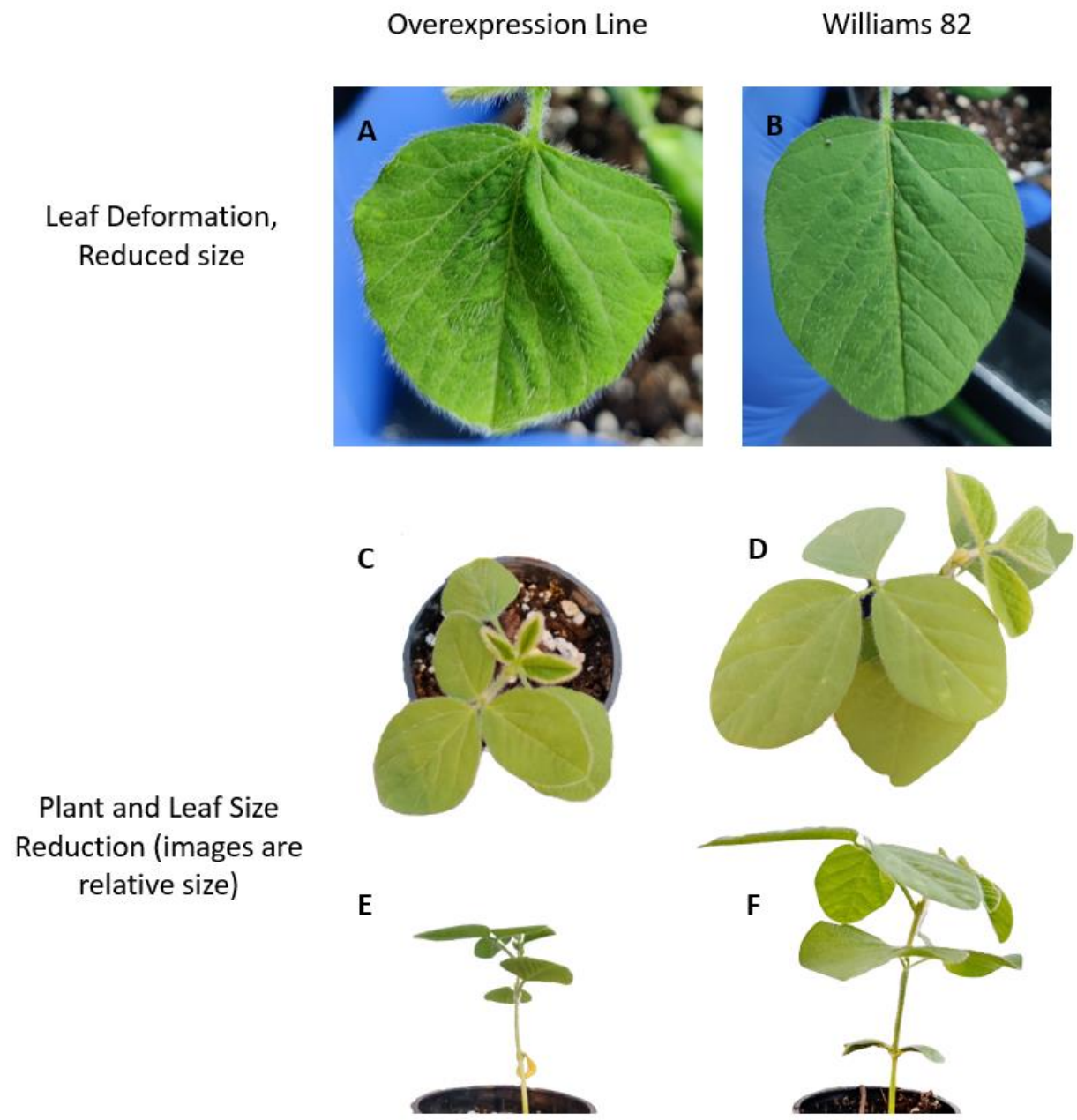


Figure 22: Phenotypic differences in *Glycine max* PAD4 overexpression lines and Williams 82 (A) Overexpression line, shows wavy margins and reduced leaf size (B) Williams 82, shows typical leaf margins and size (C & E) Overexpression line, shows reduced leaf and plant size (D & F) Williams 82, shows typical leaf and plant size.

CHAPTER 2: MOLECULAR PLANT DIAGNOSTIC METHODS

Introduction

Effective plant pathogen identification is increasingly important as pathogen management strategies transform to account for changing pathogen emergence and environmental concerns. Plant pathogen management strategies are shifting to include plant pathogen identification as stepping stone to guide further management choices. Crop losses from pests and pathogens are of concern for all crops, (Savary et al., 2019). Annual crop losses due to pests and pathogens for rice are estimated to be as high as 30%, losses for corn are estimated to be 22.5%, followed by wheat with losses at 21.5% (Savary et al., 2019). Traditional methods of pathogen control may require a longer time to perform and be less specific, leading to improper or misuse of expensive and environmentally troublesome chemicals. Production of crops is also carried out worldwide, where pest and pathogen control chemical treatments are less readily available, therefore, plant molecular diagnostics may help to tailor chemical treatments to only those necessary, and reduce cost. In addition, without pathogen identification treatments cannot be tailored for a specific pathogen and may not be as effective.

Within the US some crops cannot be sent to shipped to labs to be tested due to federal regulations. One such example is cannabis, which has major pathogens that impact the crop including viroids, bacterial and pre and post-harvest fungal infections (Punja, 2021). Broadening education for molecular diagnostic methods will make growers aware of their options and the basic science and methodology behind them.

Plant molecular diagnostics are an important part of the services that land grand extension offices to provide to the public and work in conjunction with other services such as education. The topic of plant molecular diagnostics was chosen as a topic for this extension publication because these molecular methods are widely used, or have potential innovative use, and lack easily understandable and accessible information.

In this extension publication we provide information on two common plant diagnostic methods, Enzyme-linked immunosorbent assay (ELISA), and Polymerase Chain Reaction (PCR). In addition, we provide information about three less common, but more modern or rapid methods, Lateral Flow Assay, RPA, and a general workflow for next-generation sequencing. A recent survey of extension plant diagnostic laboratories indicates that Lateral Flow Assays are being used in 85% of laboratories, ELISA is being used in 72% laboratories and PCR is being used in 60% of laboratories (Iles et al. 2021). The same survey indicates that RPA and next-generation sequencing are not widely utilized, with only 9% and 4% of labs using these methods respectively (Iles et al. 2021). However, 21% of the labs indicated that they would like to use RPA, and 29% of labs had interest in using next-generation sequencing (Iles et al. 2021).

General Aims and Goals

This extension publication is prepared for the CSU Extension, will undergo peer review, and will be published through the CSU Extension Service. In this extension publication, we had the following general aims and specific goals:

1. Provide approachable and accurate information on molecular plant diagnostic methods.
2. Provide diagrams for molecular methods that include understandable descriptions of each step

3. Explain little-emphasized aspects of diagnostics, such as proper sampling techniques and interpretation of results.

Goals:

1. Emphasize the importance of plant disease diagnostics
2. Familiarize readers with standard and upcoming diagnostic techniques
    a. Readers can identify the target of each assay
    b. Readers can identify practical considerations for each assay, such as sensitivity, time, cost, and expertise
3. Emphasize the importance of, and demonstrate proper sampling techniques
    a. Readers can adequately identify where to sample for accurate results
4. Familiarize readers with general steps for standard and upcoming diagnostic techniques
    a. Readers should be able to understand what is occurring at each step functionally
    b. Readers should be aware of the method output type (Spectroscopy, Visual, ect.)
    c. Readers should be able to understand the basic positive and negative outputs of each method

Publication Content

Title: Molecular Plant Diagnostic Methods

Overview

- Effective sampling
- ELISA
- Lateral Flow Assay

- PCR
- RPA
- Sequencing
- Practical Considerations

Why are plant diagnostic methods important?

Reliable plant disease diagnostics are essential for developing specific and effective treatment strategies that reduce crop loss. Accurate identification of plant pathogens can enable a grower to develop specific and cost-effective management strategies. Molecular methods, though more expensive, can provide information about pathogens rapidly and can differentiate between pathogens causing similar disease symptoms. Molecular methods can detect pathogens at a lower level in samples, allowing for the earlier identification of pathogens. In addition, they can provide results that have higher accuracy which allows for management strategies to be tailored specifically to the pathogen of concern and can reduce the need for excess application of fungicides and pesticides.

Effective Sampling

Sampling from tissue that is ***symptomatic*** but ***living*** is essential. Sampling from ***diverse tissues*** such as root, stem, and leaves can ensure that regardless of pathogen or life stage that the sample will have pathogen to allow for identification (Figure 23). Pathogens can have complex movements between plant tissues throughout their life cycle. Therefore, it is essential to sample from a variety of tissues.

ELISA (Enzyme-linked Immunosorbent Assay)

ELISA identifies a pathogen protein within the sample and provides an approximate quantification of a pathogen within the sample (Figure 24). Antibodies used in ELISA are highly specific and can differentiate between similar pathogens with high accuracy. Sample preparation for this method is minimal and can be performed by grinding plant tissue in an extraction buffer. Though ELISA is inexpensive and suitable for numerous samples, it is time intensive. Results are interpreted through spectroscopy, meaning that specialized expertise and equipment are required.

Lateral Flow Assay

The technique uses antibodies to identify pathogen protein or genomic material in the sample. After grinding sample tissue in the extraction buffer, the liquid is drawn down the lateral flow test strip where pathogen protein/genomic material binds to a labeled antibody (Figure 25)(Figure 26). The sample is drawn further down the test strip where labeled antibodies bound to pathogen protein will bind to an antibody at the test line, which results in a positive test line (Figure 26). Antibodies used in this method are highly specific and can differentiate between similar pathogens with high accuracy. Sample preparation for this method is minimal and can be performed by grinding plant tissue in an extraction buffer. However, this method is more costly than other antibody based methods, such as ELISA, and cannot easily be altered to accommodate multiple samples. Lateral Flow Assay is a simple and rapid procedure that can be performed in under 30 minutes. Results can be interpreted visually on the test strip, and all steps can be performed at room temperature, which means this assay can be performed in the field.

PCR (Polymerase Chain Reaction)

PCR utilizes a polymerase enzyme to amplify specific regions of pathogen genomes (Figure 27). The PCR reaction is carried out in a thermocycler, which allows each step of amplification, denaturation, annealing, and elongation to occur at optimal temperatures (Figure 27). Samples that contain the pathogen will show DNA amplification, while samples that do not have the pathogen will not. A PCR reaction contains several components: (1) Primers, which target a specific part of the pathogen genome, (2) polymerase, an enzyme that elongates the DNA, (3) dNTPs, nucleotides that the polymerase utilizes while elongating, and (4) salts, which provide optimal conditions for the reaction. Because primers are designed for almost any pathogen with genetic material, PCR can identify various pathogens, from bacteria to viruses. However, it may be difficult to distinguish between different pathogen strains due to their high degree of genomic similarity. PCR assay is more sensitive because target DNA from pathogens is amplified, allowing pathogen detection at low infection levels. Sample preparation for this method requires extraction of genomic material, DNA, or RNA; therefore, it must be performed in a laboratory. In addition, performing the PCR, or amplification reaction, requires samples to be incubated at multiple temperatures and therefore requires equipment and some expertise.

RPA (Recombinase Polymerase Amplification)

RPA utilizes a recombinase enzyme to amplify specific regions of pathogen genomes (Figure 28). Regions of amplification are designated by primers and typically designed to identify a specific pathogen. Samples that contain the pathogen will show DNA amplification while samples that do not have the pathogen will not. This assay is more sensitive thans ELISA, or Lateral Flow Assay, because target DNA from pathogens is amplified, allowing for a lower

amount of pathogen to be detectable. This method does not require DNA extraction and amplification steps can be performed at a wide range of temperatures. Reagents for the amplification reaction can be dried, allowing for use outside of laboratories.. Results can be visualized through several methods, each of which requires different equipment. RPA is a versatile assay that can be performed rapidly and with reagents that do not require special storage. This assay can easily be adapted for use with multiple samples and in the field.

Sequencing

Sequencing is utilized to identify a pathogen through its complete genomic material. Other molecular methods, which target pathogen protein, or small portions of pathogen DNA, can result in false negatives if a pathogen can alter the assay target, be it protein or DNA. Though there are many sequencing technologies, most follow a similar workflow (Figure 29). After sequencing, the generated DNA or RNA sequences can be aligned to known sequences available in public databases. This technique requires high-quality DNA or RNA extraction and preparation of fragmented DNA or RNA. . Pathogen identification can be made through direct comparison to known sequences. Sequencing does not require a known genetic sequence or pathogen-specific protein. This is especially helpful in pathogens such as viruses, which can recombine quickly and be difficult to identify confidently. In addition, sequencing allows for a complete view of pathogens in a sample. Sequencing can be performed through various methods, which are diverse and provide different advantages that can significantly change how sequencing is performed and completixty of computational analysis. Sequencing can be costly, and sample preparation and analysis typically require personnel with expertise.

Practical Considerations

Molecular diagnostics can identify many pathogens, but each method has practical considerations (Figure 30). A method's sensitivity determines how low pathogen content a method can reliably detect. Time is also a consideration when choosing a method. Some methods, such as RPA and Lateral Flow Assay, can be performed in a short period, typically under 30 minutes, and are therefore more appropriate for on-site and field testing. In addition, some methods, such as PCR and ELISA, are more time-consuming for fewer samples but can be more efficient when more samples are being analyzed. Molecular diagnostic techniques are likely to be more expensive than other diagnostic methods due to their requirement for equipment and materials. However, within the molecular methods, there is still a wide range of costs associated with individual methods. Finally, some methods require less expertise and equipment, such as Lateral Flow Assay and RPA, and therefore are viable options for field diagnostics. Other methods, such as ELISA, PCR, and Sequencing, require equipment or trained personnel and, therefore, typically must be performed in a laboratory.

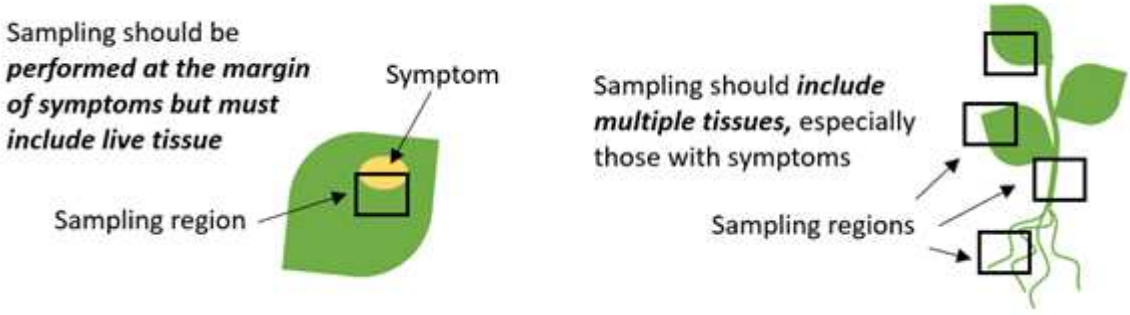


Figure 23: Effective sampling should include living tissue from symptomatic areas, and include a variety of tissues.

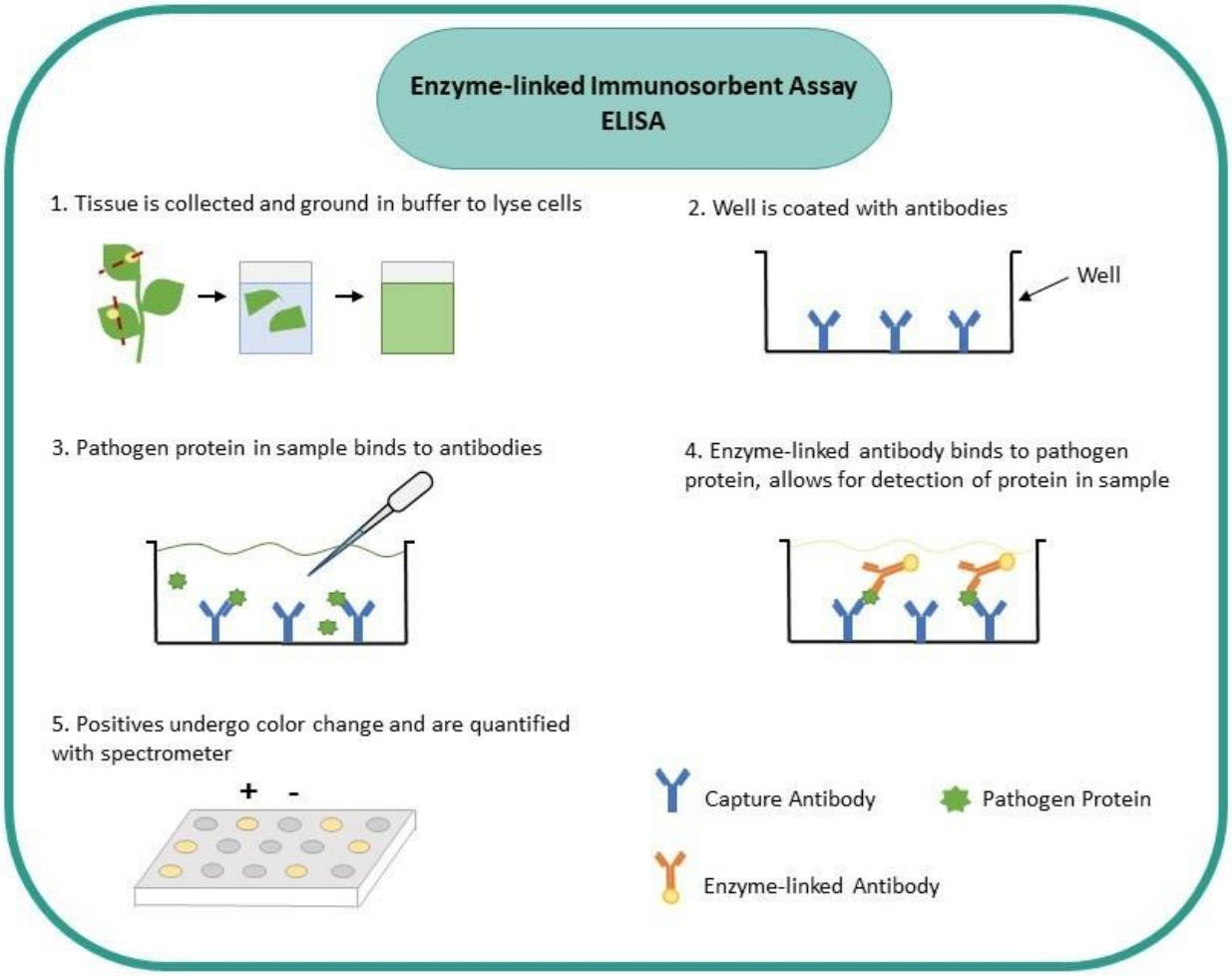


Figure 24: Enzyme-linked Immunosorbent Assay (ELISA) ELISA utilizes pathogen-targeting antibodies to identify pathogens within a sample. Positive tests are quantified with a spectrometer.

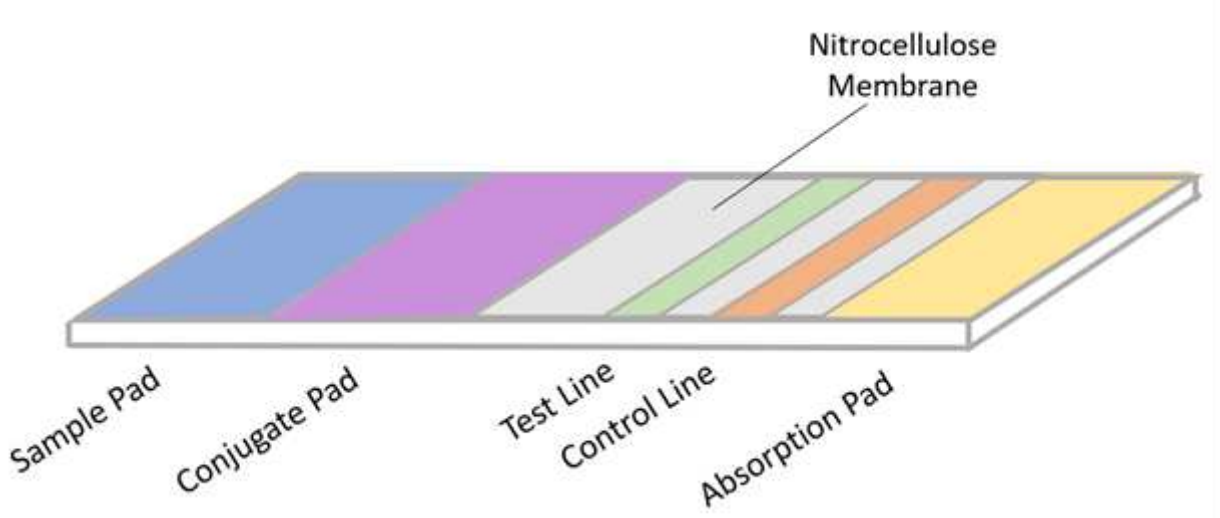


Figure 25: Structure of Lateral Flow Assay test strip Lateral Flow Assay is performed on a test strip. Sample will be wicked into the test strip at the sample pad and pass through each region, the conjugate pad, test line, and control line, finally ending at the absorption pad.

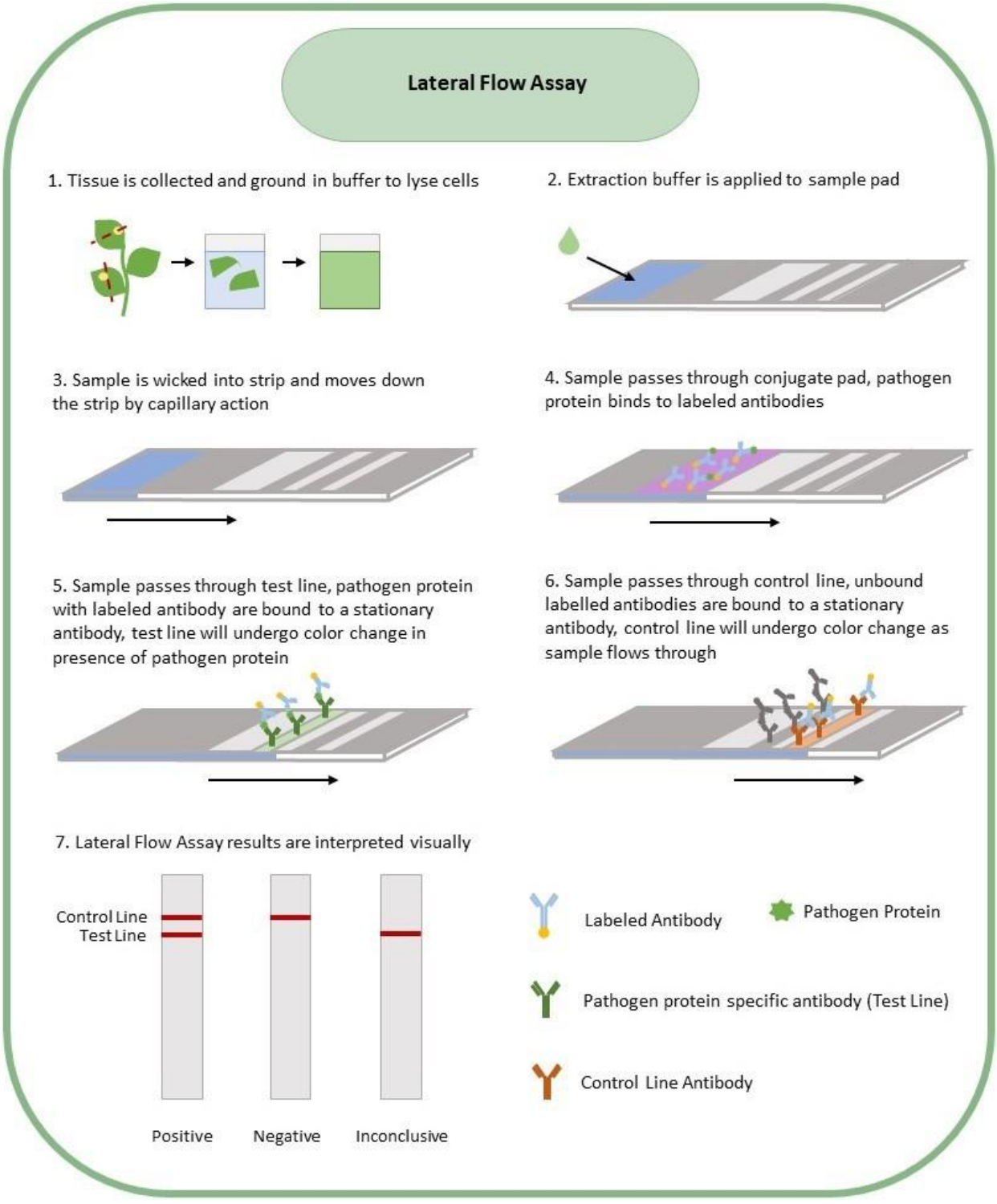


Figure 26: Lateral Flow Assay Lateral Flow Assay utilizes pathogen-targeting antibodies to identify pathogens within a sample. Positive tests are quantified visually on the test strip.

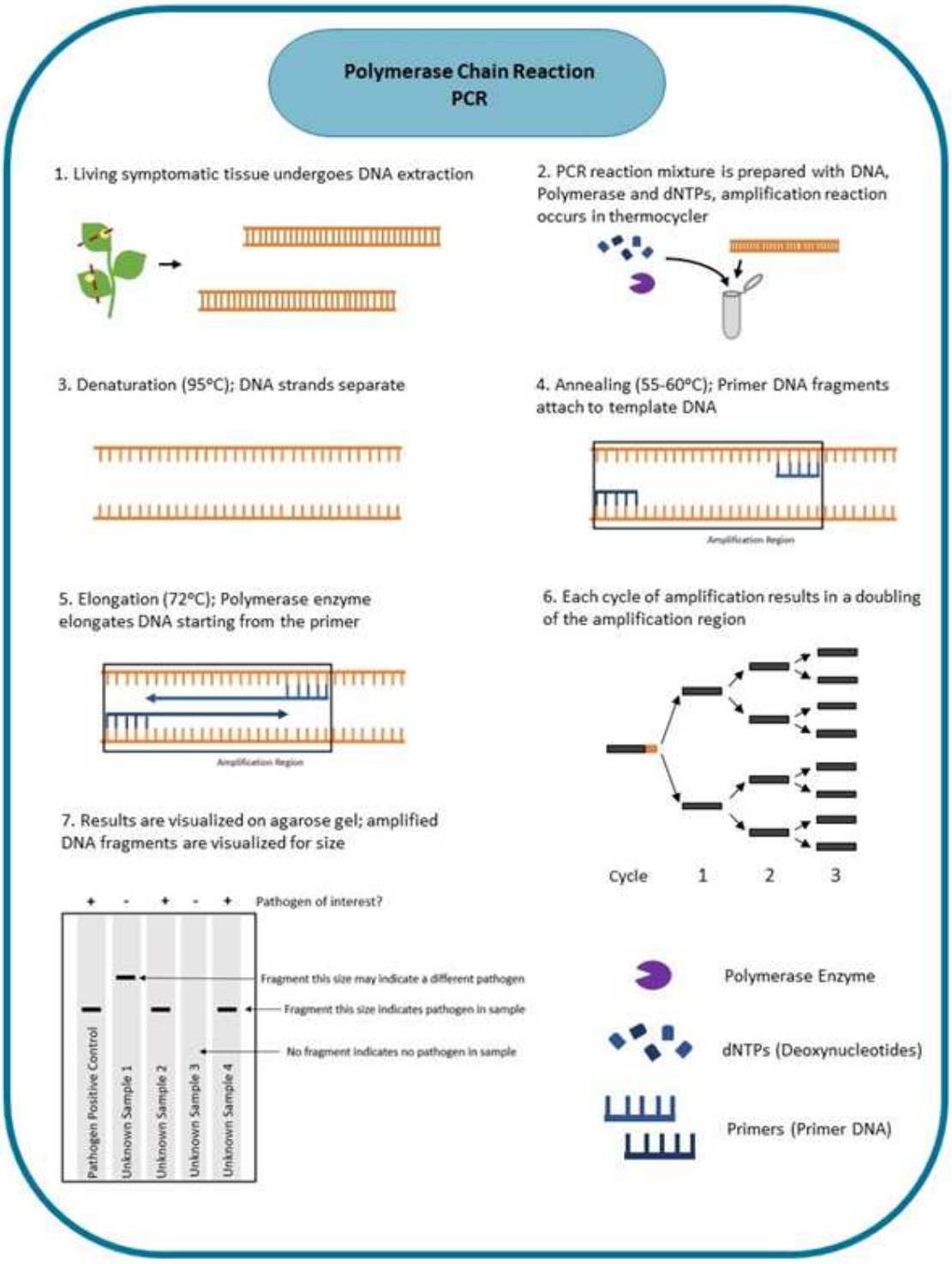


Figure 27: Polymerase Chain Reaction (PCR) PCR uses pathogen-genome specific primer DNA and a polymerase enzyme to amplify pathogen DNA fragments. Amplified DNA fragments can be visualized with an agarose gel, where banding patterns can indicate pathogen presence in sample.

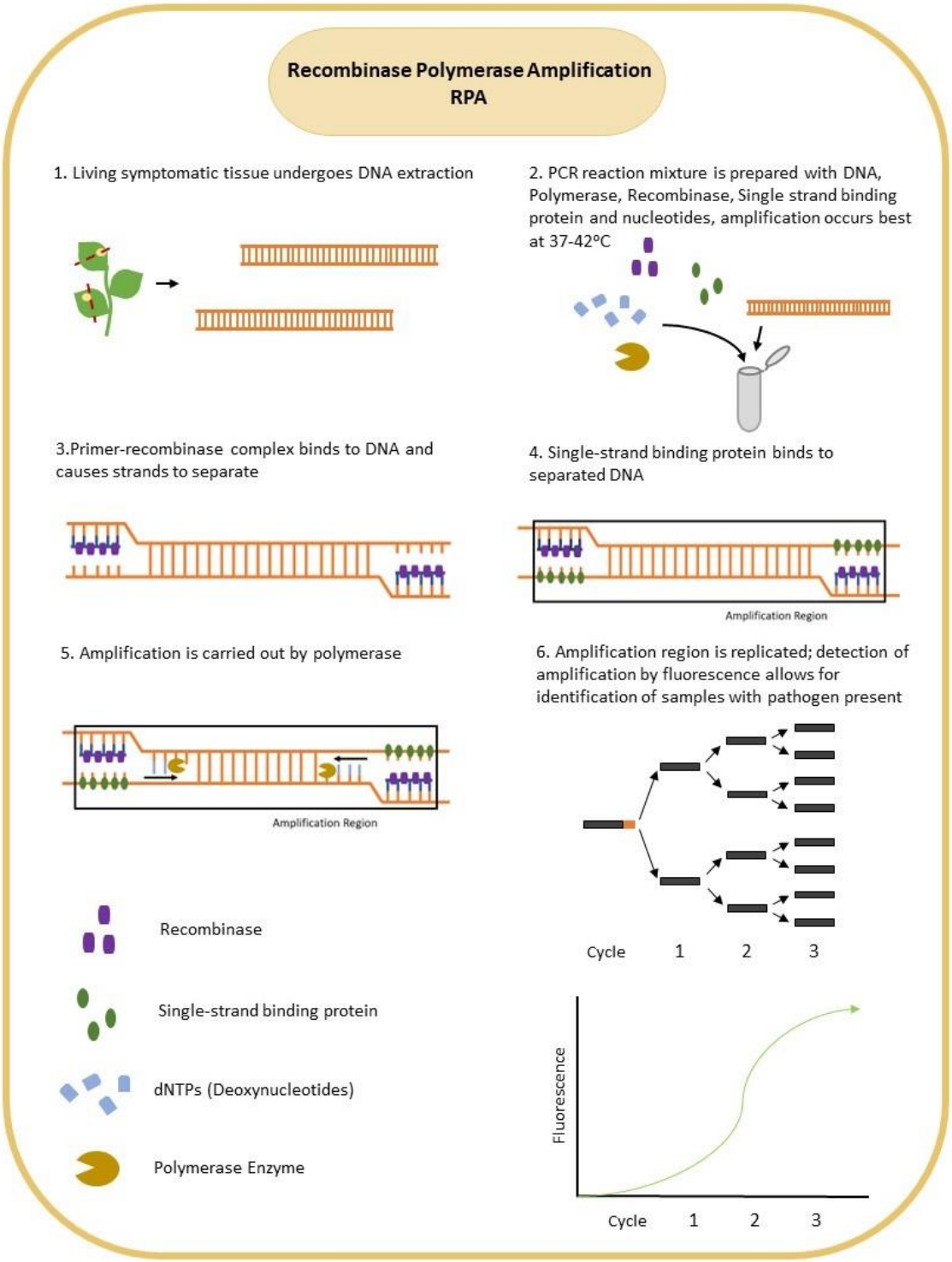


Figure 28: Recombinase Polymerase Amplification (RPA) RPA uses pathogen-genome specific primer DNA and a recombinase and polymerase enzyme to amplify pathogen DNA fragments. Amplified DNA fragments can be visualized with spectroscopy.

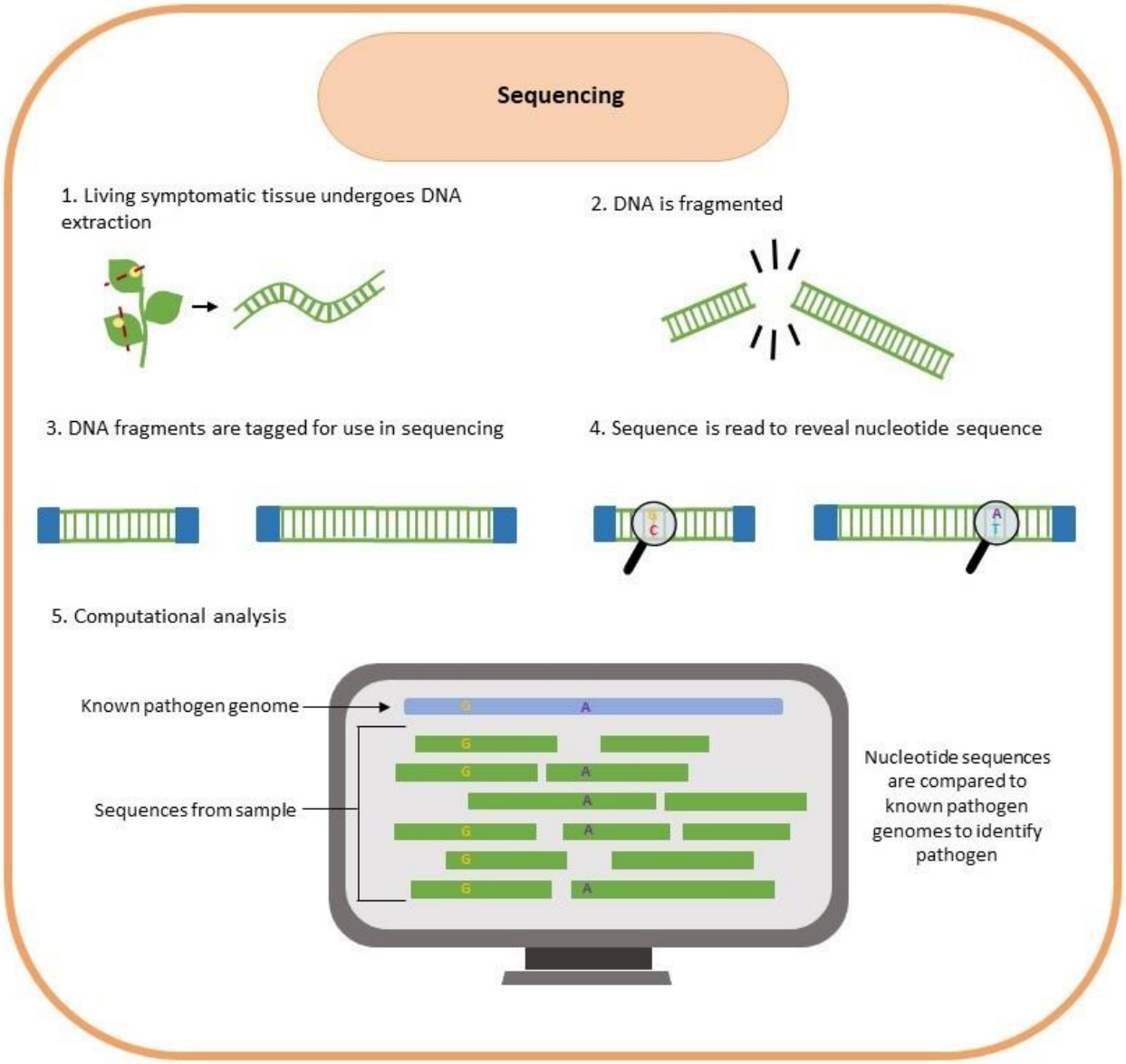


Figure 29: Sequencing Workflow Specific sequencing technologies follow individual workflows, but in general sequencing requires the extraction of genomic material (DNA or RNA), the tagging of DNA fragments, the reading of each fragments nucleotide sequence, and then computational analysis to compare fragment sequences to known pathogen genomes.

Figure 30: Practical Considerations Each molecular plant diagnostic method has practical considerations that can be important when choosing a method for diagnostic use. These practical considerations include what the method targets, the sensitivity of the method, time, cost, and expertise.

References

Baggs, E. L., Monroe, J. G., Thanki, A. S., O'Grady, R., Schudoma, C., Haerty, W., & Krasileva, K. V.. (2020). Convergent Loss of an EDS1/PAD4 Signaling Pathway in Several Plant Lineages Reveals Coevolved Components of Plant Immunity and Drought Response. *The Plant Cell*, *32*(7), 2158–2177. https://doi.org/10.1105/tpc.19.00903

Barbazuk, W. B., Fu, Y., & Mcginnis, K. M.. (2008). Genome-wide analyses of alternative splicing in plants: Opportunities and challenges. *Genome Research*, *18*(9), 1381–1392. https://doi.org/10.1101/gr.053678.106a

Bari, R., & Jones, J. D. G.. (2009). Role of plant hormones in plant defense responses. *Plant Molecular Biology*, *69*(4), 473–488. https://doi.org/10.1007/s11103-008-9435-0

Berens, M. L., Berry, H. M., Mine, A., Argueso, C. T., & Tsuda, K.. (2017). Evolution of Hormone Signaling Networks in Plant Defense. *Annual Review of Phytopathology*, *55*(1), 401–425. https://doi.org/10.1146/annurev-phyto-080516-035544

Bernacki, M. J., Czarnocka, W., Rusaczonek, A., Witoń, D., Kęska, S., Czyż, J., Szechyńska-Hebda, M., & Karpiński, S.. (2019). LSD1-, EDS1- and PAD4-dependent conditional correlation among salicylic acid, hydrogen peroxide, water use efficiency and seed yield inArabidopsis thaliana. *Physiologia Plantarum*, *165*(2), 369–382. https://doi.org/10.1111/ppl.12863

Brodersen, P., Petersen, M., Bjørn Nielsen, H., Zhu, S., Newman, M.-A., Shokat, K. M., Rietz, S., Parker, J., & Mundy, J.. (2006). Arabidopsis MAP kinase 4 regulates salicylic acid- and jasmonic acid/ethylene-dependent responses via EDS1 and PAD4. *The Plant Journal*, *47*(4), 532–546. https://doi.org/10.1111/j.1365-313x.2006.02806.x

Brown, J. W. S., Simpson, C. G., Marquez, Y., Gadd, G. M., Barta, A., & Kalyna, M.. (2015). Lost in Translation: Pitfalls in Deciphering Plant Alternative Splicing Transcripts. *The Plant Cell*, *27*(8), 2083–2087. https://doi.org/10.1105/tpc.15.00572

Calixto, Guo, W., James, A. B., Tzioutziou, N. A., Entizne, J. C., Panter, P. E., Knight, H., Nimmo, H. G., Zhang, R., & Brown, J. W. (2018). Rapid and Dynamic Alternative Splicing Impacts the Arabidopsis Cold Response Transcriptome. *The Plant Cell*., 30(7), 1424–1444. https://doi.org/10.1105/tpc.18.00177

Chen, Q.-F., Xu, L., Tan, W.-J., Chen, L., Qi, H., Xie, L.-J., Chen, M.-X., Liu, B.-Y., Yu, L.-J., Yao, N., Zhang, J.-H., Shu, W., & Xiao, S.. (2015). Disruption of the Arabidopsis Defense Regulator Genes SAG101, EDS1, and PAD4 Confers Enhanced Freezing Tolerance. *Molecular Plant*, *8*(10), 1536–1549. https://doi.org/10.1016/j.molp.2015.06.009

Clough, S. J., & Bent, A. F.. (1998). Floral dip: a simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana. *The Plant Journal*, *16*(6), 735–743. https://doi.org/10.1046/j.1365-313x.1998.00343.x

Dongus, J. A., Bhandari, D. D., Patel, M., Archer, L., Dijkgraaf, L., Deslandes, L., Shah, J., & Parker, J. E.. (2020). The Arabidopsis PAD4 Lipase-Like Domain Is Sufficient for Resistance to Green Peach Aphid. *Molecular Plant-microbe Interactions®*, *33*(2), 328–335. https://doi.org/10.1094/mpmi-08-19-0245-r

EDWARDS, JOHNSTONE, C., & THOMPSON, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Research, 19(6), 1349–1349. https://doi.org/10.1093/nar/19.6.1349

Feys, B. J., Wiermer, M., Bhat, R. A., Moisan, L. J., Medina-Escobar, N., Neu, C., Cabral, A., & Parker, J. E.. (2005a). Arabidopsis SENESCENCE-ASSOCIATED GENE101 Stabilizes and Signals within an ENHANCED DISEASE SUSCEPTIBILITY1 Complex in Plant Innate Immunity. *The Plant Cell*, *17*(9), 2601–2613. https://doi.org/10.1105/tpc.105.033910

Feys, B. J.. (2001). Direct interaction between the Arabidopsis disease resistance signaling proteins, EDS1 and PAD4. *The EMBO Journal*, *20*(19), 5400–5411. https://doi.org/10.1093/emboj/20.19.5400

Filichkin, S. A., Priest, H. D., Givan, S. A., Shen, R., Bryant, D. W., Fox, S. E., Wong, W.-K., & Mockler, T. C.. (2010). Genome-wide mapping of alternative splicing in Arabidopsis thaliana. *Genome Research*, *20*(1), 45–58. https://doi.org/10.1101/gr.093302.109

Flavell, R. B.. (1994). Inactivation of gene expression in plants as a consequence of specific sequence duplication.. *Proceedings of the National Academy of Sciences*, *91*(9), 3490–3496. https://doi.org/10.1073/pnas.91.9.3490

Glazebrook, J.. (2005). Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annual Review of Phytopathology*, *43*(1), 205–227. https://doi.org/10.1146/annurev.phyto.43.040204.135923

Glazebrook, Rogers, E. ., & Ausubel, F. . (1996). Isolation of Arabidopsis mutants with enhanced disease susceptibility by direct screening. *Genetics (Austin)*, 143(2), 973–982. https://doi.org/10.1093/genetics/143.2.973

Grant, & Lamb, C. (2006). Systemic immunity. Current Opinion in Plant Biology, 9(4), 414–420. https://doi.org/10.1016/j.pbi.2006.05.013

HÖFGEN, & WILLMITZER, L. (1988). Storage of competent cells for Agrobacterium transformation. Nucleic Acids Research, 16(20), 9877–9877. https://doi.org/10.1093/nar/16.20.9877

Iles, L. C., Fulladolsa, A. C., Smart, A., Bonkowski, J., Creswell, T., Harmon, C. L., Hammerschmidt, R., Hirch, R. R., & Rodriguez Salamanca, L.. (2021). Everything Is Faster: How Do Land-Grant University–Based Plant Diagnostic Laboratories Keep Up with a Rapidly Changing World?. *Annual Review of Phytopathology*, *59*(1), 333–349. https://doi.org/10.1146/annurev-phyto-020620-102557

Jirage, D., Tootle, T. L., Reuber, T. L., Frost, L. N., Feys, B. J., Parker, J. E., Ausubel, F. M., & Glazebrook, J.. (1999). Arabidopsis thaliana PAD4 encodes a lipase-like gene that is important for salicylic acid signaling. *Proceedings of the National Academy of Sciences*, *96*(23), 13583–13588. https://doi.org/10.1073/pnas.96.23.13583

Kalyna, M., Simpson, C. G., Syed, N. H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., Mcnicol, J., Dinh, H. Q., Barta, A., & Brown, J. W. S.. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. *Nucleic Acids Research*, *40*(6), 2454–2469. https://doi.org/10.1093/nar/gkr932

Karasov, Chae, E., Herman, J. J., & Bergelson, J. (2017). Mechanisms to Mitigate the Trade-Off between Growth and Defense. *The Plant Cell*, 29(4), 666–680. https://doi.org/10.1105/tpc.16.00931

Ke, Y., Liu, H., Li, X., Xiao, J., & Wang, S.. (2014). Rice OsPAD4functions differently from Arabidopsis AtPAD4 in host-pathogen interactions. *The Plant Journal*, *78*(4), 619–631. https://doi.org/10.1111/tpj.12500

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K.. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lei, J., A. Finlayson, S., Salzman, R. A., Shan, L., & Zhu-Salzman, K.. (2014). BOTRYTIS-INDUCED KINASE1 Modulates Arabidopsis Resistance to Green Peach Aphids via PHYTOALEXIN DEFICIENT4. *Plant Physiology*, *165*(4), 1657–1670. https://doi.org/10.1104/pp.114.242206

Louis, J., & Shah, J.. (2015). Plant defence against aphids: the PAD4 signalling nexus. *Journal of Experimental Botany*, *66*(2), 449–454. https://doi.org/10.1093/jxb/eru454

Louis, J., Gobbato, E., Mondal, H. A., Feys, B. J., Parker, J. E., & Shah, J.. (2012). Discrimination of Arabidopsis PAD4 Activities in Defense against Green Peach Aphid and Pathogens. *Plant Physiology*, *158*(4), 1860–1872. https://doi.org/10.1104/pp.112.193417

Louis, J., Leung, Q., Pegadaraju, V., Reese, J., & Shah, J.. (2010a). PAD4-Dependent Antibiosis Contributes to the ssi2-Conferred Hyper-Resistance to the Green Peach Aphid. *Molecular Plant-microbe Interactions®*, *23*(5), 618–627. https://doi.org/10.1094/mpmi-23-5-0618

Makandar, R., Nalam, V. J., Chowdhury, Z., Sarowar, S., Klossner, G., Lee, H., Burdan, D., Trick, H. N., Gobbato, E., Parker, J. E., & Shah, J.. (2015). The Combined Action of ENHANCED DISEASE SUSCEPTIBILITY1, PHYTOALEXIN DEFICIENT4, and SENESCENCE-ASSOCIATED101 Promotes Salicylic Acid-Mediated Defenses to Limit Fusarium graminearum Infection in Arabidopsis thaliana. *Molecular Plant-microbe Interactions®*, *28*(8), 943–953. https://doi.org/10.1094/mpmi-04-15-0079-r

Martín, G., Márquez, Y., Mantica, F., Duque, P., & Irimia, M.. (2021). Alternative splicing landscapes in Arabidopsis thaliana across tissues and stress conditions highlight major

functional differences with animals. *Genome Biology*, *22*(1). https://doi.org/10.1186/s13059-020-02258-y

Mastrangelo, Marone, D., Laidò, G., De Leonardis, A. M., & De Vita, P. (2012). Alternative splicing: Enhancing ability to cope with stress via transcriptome plasticity. Plant Science (Limerick), 185-186, 40–49. https://doi.org/10.1016/j.plantsci.2011.09.006

Mateo, Muhlenbock, P., Rusterucci, C., Chang, C. C. ., Miszalski, Z., Karpinska, B., Parker, J. ., Mullineaux, P. ., & Karpinski, S. (2004). LESION SIMULATING DISEASE 1 is required for acclimation to conditions that promote excess excitation energy. *Plant Physiology (Bethesda),* 136(1), 2818–2830. https://doi.org/10.1104/pp.104.043646

MüHlenbock, P., Plaszczyca, M., Plaszczyca, M., Mellerowicz, E., & Karpinski, S.. (2007). Lysigenous Aerenchyma Formation in Arabidopsis Is Controlled by LESION SIMULATING DISEASE1. *The Plant Cell*, *19*(11), 3819–3830. https://doi.org/10.1105/tpc.106.048843

MüHlenbock, P., SzechyńSka-Hebda, M., Płaszczyca, M., Baudo, M., Mateo, A., Mullineaux, P. M., Parker, J. E., KarpińSka, B., & KarpińSki, S.. (2008). Chloroplast Signaling andLESION SIMULATING DISEASE1Regulate Crosstalk between Light Acclimation and Immunity inArabidopsis. *The Plant Cell*, *20*(9), 2339–2356. https://doi.org/10.1105/tpc.108.059618

Mur, Kenton, P., Atzorn, R., Miersch, O., & Wasternack, C. (2006). Outcomes of Concentration-Specific Interactions between Salicylate and Jasmonate Signaling Include Synergy, Antagonism, and Oxidative Stress Leading to Cell Death. Plant Physiology (Bethesda), 140(1), 249–262. https://doi.org/10.1104/pp.105.072348

Ng, G., Seabolt, S., Zhang, C., Salimian, S., Watkins, T. A., & Lu, H.. (2011). Genetic Dissection of Salicylic Acid-Mediated Defense Signaling Networks in Arabidopsis. *Genetics*, *189*(3), 851–859. https://doi.org/10.1534/genetics.111.132332

Pegadaraju, V., Knepper, C., Reese, J., & Shah, J.. (2005). Premature Leaf Senescence Modulated by the Arabidopsis PHYTOALEXIN DEFICIENT4 Gene Is Associated with Defense against the Phloem-Feeding Green Peach Aphid. *Plant Physiology*, *139*(4), 1927–1934. https://doi.org/10.1104/pp.105.070433

Pruitt, R. N., Locci, F., Wanke, F., Zhang, L., Saile, S. C., Joe, A., Karelina, D., Hua, C., Fröhlich, K., Wan, W.-L., Hu, M., Rao, S., Stolze, S. C., Harzen, A., Gust, A. A., Harter, K., Joosten, M. H. A. J., Thomma, B. P. H. J., Zhou, J.-M., Nürnberger, T.. (2021). The EDS1–PAD4–ADR1 node mediates Arabidopsis pattern-triggered immunity. *Nature*, *598*(7881), 495–499. https://doi.org/10.1038/s41586-021-03829-0

Punja, Z. K.. (2021). Emerging diseases of Cannabis sativa and sustainable management. *Pest Management Science*, *77*(9), 3857–3870. https://doi.org/10.1002/ps.6307

Reddy, A. S. N., Marquez, Y., Kalyna, M., & Barta, A.. (2013). Complexity of the Alternative Splicing Landscape in Plants. *The Plant Cell*, *25*(10), 3657–3683. https://doi.org/10.1105/tpc.113.117523

Rietz, S., Stamm, A., Malonek, S., Wagner, S., Becker, D., Medina-Escobar, N., Corina Vlot, A., Feys, B. J., Niefind, K., & Parker, J. E.. (2011). Different roles of Enhanced Disease Susceptibility1 (EDS1) bound to and dissociated from Phytoalexin Deficient4 (PAD4) in Arabidopsis immunity. *New Phytologist*, *191*(1), 107–119. https://doi.org/10.1111/j.1469-8137.2011.03675.x

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., Mcroberts, N., & Nelson, A.. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, *3*(3), 430–439. https://doi.org/10.1038/s41559-018-0793-y

Selig, P. R. (2017). *Alternative Splicing of a Host Defense Regulatory Gene, 0RW1S34RfeSDcfkexd09rT2 PHYTOALEXIN DEFICIENT41RW1S34RfeSDcfkexd09rT2 (0RW1S34RfeSDcfkexd09rT2PAD41RW1S34RfeSDcfkexd09rT2) in Soybean (0RW1S34RfeSDcfkexd09rT2 Glycine max1RW1S34RfeSDcfkexd09rT2)* (Order No. 10274144). Available from ProQuest Dissertations & Theses Global. (1952752843). https://ezproxy2.library.colostate.edu/login?url=https://www.proquest.com/dissertations-theses/alternative-splicing-host-defense-regulatory-gene/docview/1952752843/se-2

Singh, V., & Shah, J.. (2012). Tomato responds to green peach aphid infestation with the activation of trehalose metabolism and starch accumulation. *Plant Signaling & Behavior*, *7*(6), 605–607. https://doi.org/10.4161/psb.20066

Staiger, D., & Brown, J. W. S.. (2013). Alternative Splicing at the Intersection of Biological Timing, Development, and Stress Responses. *The Plant Cell*, *25*(10), 3640–3656. https://doi.org/10.1105/tpc.113.113803

Szechyńska-Hebda, M., Czarnocka, W., Hebda, M., & Karpiński, S.. (2016). PAD4, LSD1 and EDS1 regulate drought tolerance, plant biomass production, and cell wall properties. *Plant Cell Reports*, *35*(3), 527–539. https://doi.org/10.1007/s00299-015-1901-y

Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J., & Katagiri, F.. (2009). Network Properties of Robust Immunity in Plants. *PLOS Genetics*, *5*(12), e1000772. https://doi.org/10.1371/journal.pgen.1000772

Wagner, S., Stuttmann, J., Rietz, S., Guerois, R., Brunstein, E., Bautor, J., Niefind, K., & Jane. (2013). Structural Basis for Signaling by Exclusive EDS1 Heteromeric Complexes with SAG101 or PAD4 in Plant Innate Immunity. *Cell Host & Microbe*, *14*(6), 619–630. https://doi.org/10.1016/j.chom.2013.11.006

Walters, B., Lum, G., Sablok, G., & Min, X. J.. (2013). Genome-Wide Landscape of Alternative Splicing Events in Brachypodium distachyon. *DNA Research*, *20*(2), 163–171. https://doi.org/10.1093/dnares/dss041

Wiermer, M., Feys, B. J., & Parker, J. E.. (2005). Plant immunity: the EDS1 regulatory node. *Current Opinion in Plant Biology*, *8*(4), 383–389. https://doi.org/10.1016/j.pbi.2005.05.010

Wituszyńska, W., Ślesak, I., Vanderauwera, S., Szechyńska-Hebda, M., Kornaś, A., Van Der Kelen, K., Mühlenbock, P., Karpińska, B., Maćkowski, S., Van Breusegem, F., & Karpiński, S.. (2013). LESION SIMULATING DISEASE1, ENHANCED DISEASE SUSCEPTIBILITY1, and PHYTOALEXIN DEFICIENT4 Conditionally Regulate Cellular Signaling Homeostasis, Photosynthesis, Water Use Efficiency, and Seed Yield in Arabidopsis. *Plant Physiology*, *161*(4), 1795–1805. https://doi.org/10.1104/pp.112.208116

Wubben, M. J. E., Jin, J., & Baum, T. J.. (2008). Cyst Nematode Parasitism of Arabidopsis thaliana Is Inhibited by Salicylic Acid (SA) and Elicits Uncoupled SA-Independent Pathogenesis-Related Gene Expression in Roots. *Molecular Plant-microbe Interactions®*, *21*(4), 424–432. https://doi.org/10.1094/mpmi-21-4-0424

Xiao, S., Brown, S., Patrick, E., Brearley, C., & Turner, J. G.. (2003). Enhanced Transcription of the Arabidopsis Disease Resistance Genes RPW8.1 and RPW8.2 via a Salicylic Acid–Dependent Amplification Circuit Is Required for Hypersensitive Cell Death. *The Plant Cell*, *15*(1), 33–45. https://doi.org/10.1105/tpc.006940

Xing, D., & Chen, Z.. (2006). Effects of mutations and constitutive overexpression of EDS1 and PAD4 on plant resistance to different types of microbial pathogens. *Plant Science*, *171*(2), 251–262. https://doi.org/10.1016/j.plantsci.2006.03.022

Xing, D., & Chen, Z.. (2006). Effects of mutations and constitutive overexpression of EDS1 and PAD4 on plant resistance to different types of microbial pathogens. *Plant Science*, *171*(2), 251–262. https://doi.org/10.1016/j.plantsci.2006.03.022

Youssef, R. M., Macdonald, M. H., Brewer, E. P., Bauchan, G. R., Kim, K.-H., & Matthews, B. F.. (2013). Ectopic expression of AtPAD4broadens resistance of Soybean to soybean cyst and root-knot nematodes. *BMC Plant Biology*, *13*(1), 67. https://doi.org/10.1186/1471-2229-13-67

Zeng, H., Liu, Y., Chen, D., Bao, H., Huang, L., Yin, J., Chen, Y., Xiao, S., & Yao, N.. (2021). The immune components ENHANCED DISEASE SUSCEPTIBILITY 1 and PHYTOALEXIN DEFICIENT 4 are required for cell death caused by overaccumulation of ceramides in Arabidopsis. *The Plant Journal*, *107*(5), 1447–1465. https://doi.org/10.1111/tpj.15393

Zhang, & Gassmann, W. (2007). Alternative Splicing and mRNA Levels of the Disease Resistance Gene RPS4 Are Induced during Defense Responses. Plant Physiology (Bethesda), 145(4), 1577–1587. https://doi.org/10.1104/pp.107.108720

Zhou, N., Tootle, T. L., Tsui, F., Klessig, D. F., & Glazebrook, J.. (1998). PAD4 Functions Upstream from Salicylic Acid to Control Defense Responses in Arabidopsis. *The Plant Cell*, *10*(6), 1021–1030. https://doi.org/10.1105/tpc.10.6.1021

Zhou, N., Tootle, T. L., Tsui, F., Klessig, D. F., & Glazebrook, J.. (1998). PAD4 Functions Upstream from Salicylic Acid to Control Defense Responses in Arabidopsis. *The Plant Cell*, *10*(6), 1021–1030. https://doi.org/10.1105/tpc.10.6.1021

Zhu, S., Jeong, R.-D., Venugopal, S. C., Lapchyk, L., Navarre, D., Kachroo, A., & Kachroo, P.. (2011). SAG101 Forms a Ternary Complex with EDS1 and PAD4 and Is Required for Resistance Signaling against Turnip Crinkle Virus. *PLOS Pathogens*, *7*(11), e1002318. https://doi.org/10.1371/journal.ppat.1002318

List of Abbreviations

AltA, Alternate acceptor

AltD, Alternate donor

AS, Alternative splicing

EDS1, ENHANCED DISEASE SUSCEPTIBILITY 1

ELISA, Enzyme-linked Immunoaorbent Assay

EP, EDS1-PAD4

ES, Exon skipping

ET, Ethylene

ETI, Effector-triggered immunity

FC, Fold Change

GPA, Green Peach Aphid

IR, Intron retention

JA, Jasmonic acid

LP, Lipase

MPK4, MAP Kinase 4

NLR, Nucleotide-binding Lucine Rich

NMD, Non-sense Mediated Decay

PAD4, PHYTOALEXIN DEFICIENT 4

PAMPs, Pathogen Associated Molecular Patterns

PCR, Polymerase Chain Reaction

PTI, PAMP-triggered immunity

ROS, Reactive Oxygen Species

RPA, Recombinase Polymerase Amplification

RPS4, *RESISTANT TO PSEUDOMONAS SYRINGAE4*

SA, Salicylic acid

SAG101, SENESCENCE ASSOCIATED GENE 101

SAR, Systemic Acquired Resistance