12-2014

Characterization of a Putative Phospholipase D ′ Like Gene as a Lipid Signaling Modulator and Its Role in Salicylic Acid Mediated Defense Pathway in Nicotiana tabacum

Phillip T. Dean
East Tennessee State University

Characterization of a Putative Phospholipase D ´ Like Gene as a Lipid Signaling Modulator and Its Role in Salicylic Acid Mediated Defense Pathway in *Nicotiana tabacum*

A thesis

presented to

the faculty of the Department of Biological Sciences

East Tennessee State University

In partial fulfillment

of the requirement for the degree

Master of Science in Biology

---

by

Phillip Dean

December 2014

Dhirendra Kumar, PhD Chair

Karl Joplin, PhD

Shivakumar Devaiah, PhD

Keywords: Plant defense, Salicylic Acid, SABP2, SBIP-436, PLD´, Expression

ABSTRACT

Characterization of a Putative Phospholipase D ´ Like Gene as a Lipid Signaling Modulator and Its Role in Salicylic Acid Mediated Defense Pathway in *Nicotiana tabacum*

by

Phillip Dean

Plants are in a perpetual evolutionary arms race with a wide range of pathogens. The salicylic acid (SA) mediated defense pathway has been shown to be one of the major defenses plants initiate in defense against microbial pathogens. Following pathogen attack high levels of methyl salicylate (MeSA) are produced that can be converted to SA by salicylic acid binding protein 2 (SABP2). SBIP-436 is an interacting protein of tobacco SABP2 and showed high homology to phospholipase D-´ (PLD- ´). With an abundance of stimulators PLD- ´ may be a lipid signaling modulator performing various functions in different situations. We present a novel *Nicotiana tabacum PLD- ´* putative gene construct. We demonstrate that the putative PLD- ´ is subject to alternative splicing and its expression is differentially modulated under biotic and abiotic stress. Our results indicate that this putative PLD- ´ may play a role in the SA mediated defense pathway.

ACKNOWLEDGEMENTS

I would like to thank my committee members, Dr. Karl Joplin and Dr. Shivakumar Devaiah, for their guidance throughout this thesis project with support, constructive criticism, and counsel. I owe deep gratitude to my mentor, Dr. Dhirendra Kumar, whose encouragement, direction, and support from initial to the final level helped me develop a good understanding of this project. I would also like to thank my fellow lab members for their constant aid. I would like to thank all the staff members of Biological Sciences for their support. I would like to thank the ETSU Graduate School for the graduate assistantship, ETSU for RDC grant (14-021M to DK), and NSF grant (MCB 1022077 to DK) to support this project. I would also like to thank my family members and friends for their never-ending love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Plants are the essential cornerstone for the success of life on earth. The development of photosynthesis in plants provided an oxygen rich atmosphere where life could thrive, grow, and evolve. They are the foundation of the energy-rich food chain of the biotic world. Plants are used to provide nourishment, for shelter, medicine, climate control, maintaining water and air quality, to provide industrial products (clothing), and many other important things. It has been around nine-thousand years since humankind founded agriculture. It was here that we began to understand the immense importance and astounding versatility of plants. We depend on plants as fundamental parts of our everyday lives, so it is very important that we make an effort to understand plants in order to protect their well-being as well as our own. Agricultural losses due to pathogens can transform local economies and even devastate a primary food resource for a community, thereby impacting negatively on the health of the community (Scholthof 2007).

Plants live in an intricate environment where they are constantly being combated by a broad range of biotic and abiotic attackers including bacteria, viruses, fungi, nematodes, and insect herbivores. Plants do not have the simple luxury of standing up and moving away, so they must defend themselves in a different manner. Without an adaptive immune system found in animals, plants must use special mechanisms to equate a defense against potential harm. Plants, over time, have evolved an innate immune system, also found in animals, that detects pathogen molecules and creates a specific defense for that specific invader. Plants are in a perpetual evolutionary arms race with a wide range of pathogens. When plants evolve a certain defense mechanism, the pathogen evolves a response to nullify the plants defense, and in response to the

pathogen, the plant does the same. It is important to understand how plant immunity works in order to increase plant resistance and decrease their susceptibility to any organism that poses potential harm.

In the field of plant pathology, specifically epidemiology, an epidemiological model came to light in the 1960s. It was formulated in order to predict and control the spread of disease. The disease triangle is a conceptual model that shows the interactions between the environment, the host, and an infectious (or abiotic) agent (Scholthof 2007). The disease triangle is formulated on the strength of 3 different parameters: the susceptibility of the host, the pervasiveness of the pathogen, and the influence of the environment. These 3 parameters create a triangle with disease in its interior. Figure 1 demonstrates the manipulation of these 3 parameters that will influence the amount disease in the population being studied.

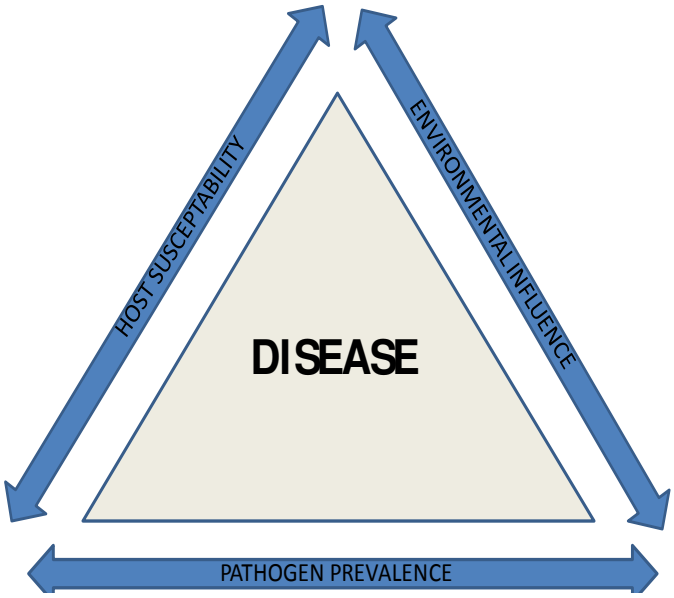


Figure 1: The Disease Triangle: Manipulation of the three surrounding parameters- Host Susceptibility, Environmental Influence, and Pathogen Prevalence- will affect the amount of disease in a population.

The reduction or interference of any of these parameters should result in less disease and better well-being within a plant population. For disease to be significant all 3 factors must be present. Ergo, our goal as researchers is to better understand these factors and find safe, effective ways to limit them.

Plant Immune System

In order to survive the continuous threat of diverse pathogenic microorganisms in their immediate vicinity, plants have developed efficient defense mechanisms (Broekaert et al. 2006). It is important to remember that plants are in a perpetual evolutionary arms race with their attacking pathogens. After pathogen resistance is increased in the host, the invading pathogen increases its pervasiveness towards the host, genetically surpassing the initial increase in resistance that was put forth.

The innate immune system relies on the immunity of each cell and a system of signals that are created from the site of infection. Plants use pattern-recognition receptors (PRR) to perceive different pathogens or microbes that have the potential to cause harm. Two important innate immunity constituents are PAMP-triggered immunity (PTI) and effector triggered immunity (ETI) (Tsuda and Katagiri 2010). PRRs activate an innate immune response upon detection of PAMPs, so-called PAMP-triggered immunity (PTI). Successful pathogens are able to overcome PTI by means of secreted effectors that suppress PTI responses, resulting in (ETS) effector-triggered susceptibility (Thomma 2011). If the PTI responses are suppressed, the plant turns to its secondary branch of defense, ETI. Plants have responded to these effectors through the development of cytoplasmic R proteins that recognize (the presence or activity of) single effectors (Thomma et al., 2011). ETI is an accelerated and amplified PTI response, resulting in

disease resistance and, usually, a hypersensitive cell death response (HR) at the infection site (Jones and Dangl, 2006).

Phytohormones

Systemic acquired resistance (SAR) is a form of inducible resistance that is triggered in systemic healthy tissues of locally infected plants (Vlot et al. 2009). Plant defense in response to microbial attack is regulated through a complex network of signaling pathways that involve 3 signaling molecules: SA, jasmonic acid (JA), and ethylene (Kunkel & Brooks 2002). These pathways do not function independently but rather influence each other through a complex network of regulatory interactions (Kunkel & Brooks 2002). Several kinds of plant–pathogen interactions result in the generation and emission of long-distance signals from the site of infection to healthy uninfected parts of the plant where subsequent resistance is induced (Vlot et al. 2009). There are other phytohormones involved in defense signaling, such as abscisic acid (ABA) and giberellic acid (GA). ABA has been proposed to act as a mediator in plant responses to a range of stresses, including drought and salt stress (Javid et al. 2011), while GAs are generally involved in growth and development. They control seed germination, leaf expansion, stem elongation, and flowering (Magome et al. 2004). Auxin or IAA (Indole-3-acetic acid) is a phytohormone that is an integral part of growth and development (Spoel and Dong 2008). IAA is produced in the apical meristem of the plant. In regards to plant defense, IAA is known to modulate pathogen resistance during plant defense behavior. Detection of a bacterial elicitor can decrease IAA sensitivity and elevate resistance to *Pseudomonas syringae*. It is also possible for pathogens to modulate IAA levels to decrease host resistance (Erb et al. 2012). Cytokinins are another important phytohormone in plants. They too play an important role in growth and

development. IAA and cytokinins have an antagonistic relationship, where cytokinins are suppressed by IAA (Shimizu-Sato et al. 2009). It has recently been suggested that cytokinin pathways may play a role in SA accumulation and defense signaling (Choi et al. 2011). Animals have secretory glands that produce hormones and release them into their system. Plants do not have these secretory structures. Instead, in plant systems the tissues and cells have the ability to produce these phytohormones. Also, unlike animals, plants do not have circulatory systems that transport important molecular structures to other sites where they are required. Plants make use of the phytohormones, which can easily move from cell to cell, tissue to tissue.

Salicylic Acid and Defense Signaling

SA was named after *Salix* plant (willow) and was first discovered as a major component in the extracts from willow tree bark that had been used as a natural anti-inflammatory drug (Wick 2012). SA has been known to play an integral central role in plant defense signaling (An and Mou 2011). Genetic studies have shown that SA is required for the rapid activation of defense responses that are mediated by several resistance genes, for the induction of local defenses that contain the growth of virulent pathogens, and for the establishment of systemic acquired resistance (Durrant and Dong 2004). Salicylic acid derivation occurs through the Shikimate/Phenypropanoid Pathway via 2 differing routes: 1. phenylalanine is converted to cinnamic acid via phenylalanine ammonia lyase (PAL) –cinnamic acid is converted to either coumaric acid or benzoic acid [benzoate intermediates] (Métraux 2002; Dempsey et al. 2011) – Salicylic Acid; 2. Chorismate is converted to Isochorismate via isochorismate synthase (ICS) –

Isochorismate is converted to SA via isochorismate pyruvate lyase (IPL)(Wildermuth et al. 2001; Strawn et al. 2007;Dempsey et al. 2011). The Shikimate pathway can be seen in Figure 2.

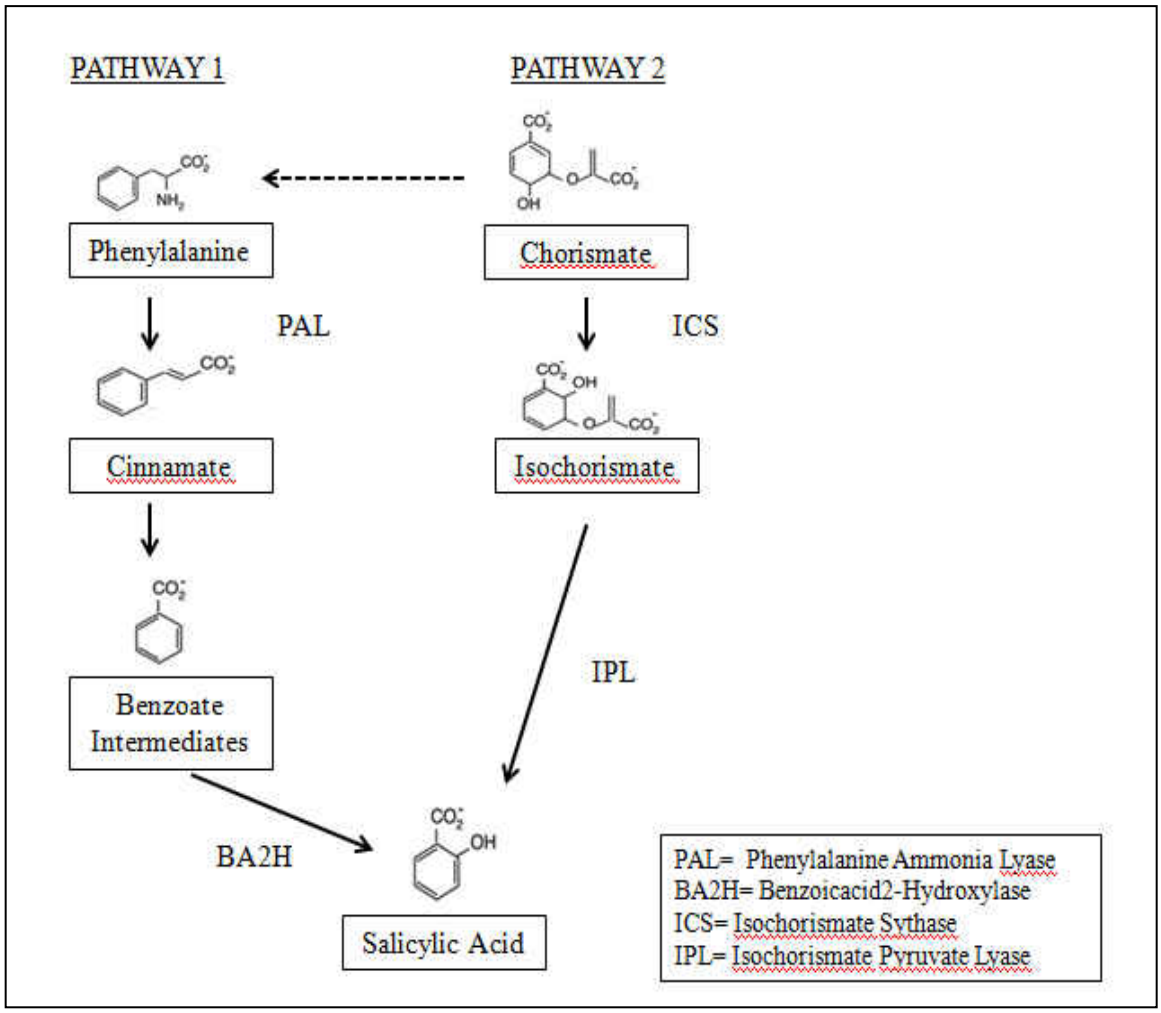


Figure 2: The Shikimate pathway – resulting in the biosynthesis of Salicylic Acid

Production of SA via these routes is required for local resistance as well as systemic acquired resistance. To acquire SAR, the SA must be mobilized to other parts of the plant. To become

mobilized, SA is methylated to form methyl salicylate (MeSA) by SA carboxyl methyltransferase (SAMT) (Dempsey et al. 2011). MeSA is a volatile ester that is dramatically increased upon pathogen infection and has been proposed to be an airborne signal (Kumar and Klessig 2003; Loake and Grant, 2007). Once MeSA has reached its target tissue SABP2/methyl esterase helps to convert SA from MeSA. The conversion of SA to MeSA to SA is demonstrated in Figure 3.

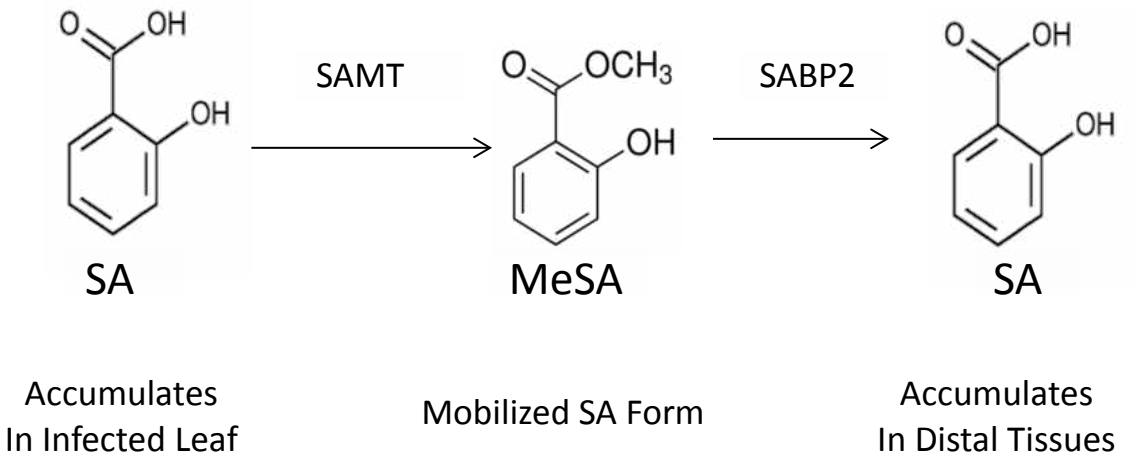


Figure 3: Conversion of SA to MeSA to SA. Salicylic acid accumulates and is converted to MeSA by salicylic acid methyltransferase (SAMT). MeSA is then converted back to SA by salicylic acid binding protein protein 2 (SABP2) in distal uninfected tissue.

Salicylic Acid Binding Protein 2

Salicylic acid binding protein 2 (SABP2) is a SA receptor that is required for the plant immune response. SABP2 is a 29kDa enzyme that converts MeSA to SA. Kumar and Klessig were the first to identify and implicate SABP2's role in plant innate immunity (Kumar and Klessig 2003). Although SABP2 binds SA with high affinity, its abundance is exceptionally low in plants. SABP2's catalytic function is to convert the volatile ester MeSA to SA in local

infected and systemic uninfected tissues to induce SAR (Kumar and Klessig 2003; Forouhar et al. 2005). In order to better understand the signaling factors surrounding SABP2, a Yeast-2-Hybrid (Y2H) screening was attempted. Several proteins were demonstrated to be successfully interacting with SABP2. These proteins were termed Salicylic Acid Binding Protein 2 Interacting Proteins (SBIP).

Yeast-2-Hybrid Screening

In the Yeast-2-Hybrid screening, SABP2 was used as the "Bait" protein (fused in frame to a binding domain), and an expressed protein library from tobacco used as "Prey" proteins (expressed as a fusion partner to an activation domain). When proteins interact with SABP2, the activation domain comes in close proximity to the binding domain leading to activation of a transcription factor upstream of the reporter gene that allows the yeast to grow, change in color, and exhibit antibiotic resistance. The yeast-2-hybrid system is shown in detail in Figure 4.

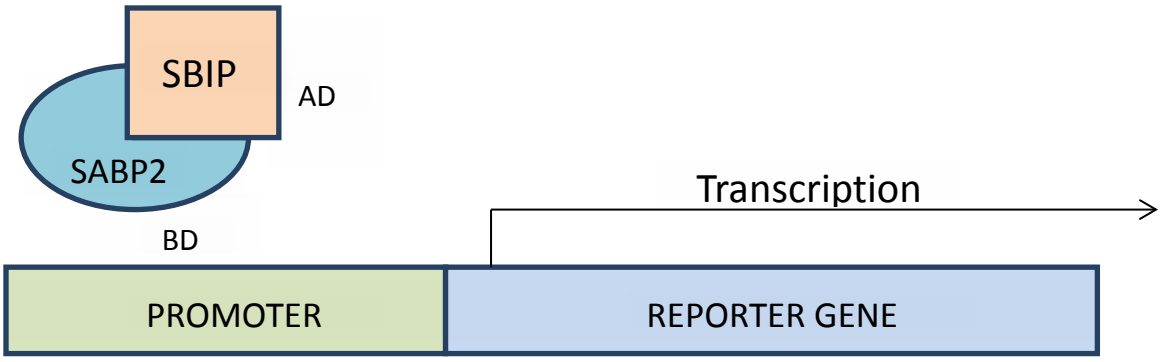


Figure 4: Yeast-2-Hybrid System: When the activating domain partner comes in close proximity to the binding domain, a transcription factor is activated upstream of the reporter gene.
Bait Protein- SABP2, BD - Binding Domain, AD - Activating Domain,

Interacting Protein- SBIP (Salicylic Acid Binding Protein 2 Interacting Protein)

One of the SBIPs, identified as SBIP-436, was shown to be a putative phospholipase D-´ like protein.

SBIP-436 Preliminary Sequence Analysis

A Yeast-2 Hybrid screening, using SABP2 as bait, yielded multiple positive interacting proteins. SBIP-436 was among these interacting proteins. The SBIP-436 sample was sequenced and analyzed. The sequenced cDNA revealed a clone 228bp in length. These can be seen in Figure 5.

A.

```
ACTTGAAAATCTCATATACGACTCACTATACCGGCGAGCGACAGCCTGGAGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCA
GTGAATTCCACCCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGGGAGAACTCAGAAGAGTTGCATATTTATTGATGTTGCTGGGGATTT
GAATTGTATCATGGTTGAGGTAAAACTTAAATAA
```

B.

```
SASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQV
```

Figure 5: SBIP-436 sequenced clone from Yeast-2 Hybrid screening. A. Nucleotide sequence. B. Translated protein sequence.

The sequence obtained from the Yeast Two-Hybrid screening was subjected to BLAST analysis in the NCBI database to determine its identity and similarity with other know genes. It was found to have high homology to PLD Superfamily of proteins. The BLAST result can be seen in Figure 6.

NCBI Nucleotide BLAST

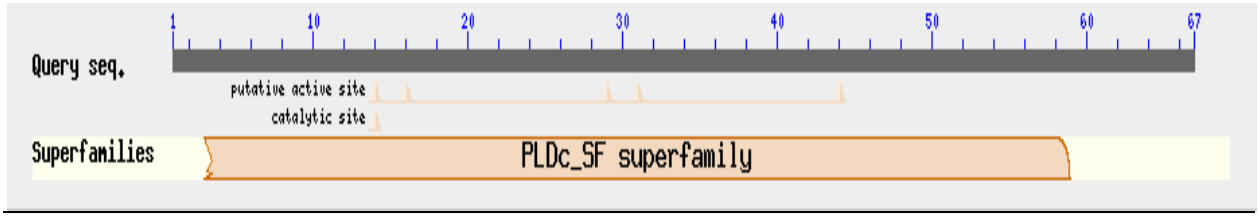


Figure 6: SBIP-436 NCBI BLAST: A nucleotide blast using the interacting protein sequence from yeast two-hybrid shows homology to PLD superfamily. (NCBI BLAST;1990)

Next, a multiple alignment was performed using the SBIP-436 corresponding 3rd frame amino acid sequence and *Arabidopsis thaliana* (AtPLD-´) as well as *Ricinus communis* (RcPLD- ´). This showed that SBIP-436 has high similarity with Phospholipase D-´. These species were chosen due to their high similarity as well as their involvement in previous Phospholipase D ´ studies. SBIP-436 showed 70% similarity with AtPLD-´ and 71% similarity with RcPLD- ´. The clustal alignment can be seen in Figure 7.

NCBI Multiple Alignment with AtPLD-´ and RcPLD- ´

```
SBIP-436   1    NYFCLMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDD  75
ATPLD- δ   709  NFQRFMIYVHAKGMIVDDEYVLMGSANINQRSMAGTKDTEIAMGAYQPNHTWAHKGRHPRGQVYGYRMSLWAEHLGKTGD  788
RCPLD- δ   689  KHQRFMIYVHAKGMIVDDEYVIMGSANINQRSMAGSRDTEIAMGAYQPNHTWGNKKRHPRGQVYGYRMSLWAEHLGLVDS  768

SBIP-436   76   CFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGART-NLPDALTT  154
ATPLD- δ   789  EFVEPSDLECLKKVNTISEENWKRFIDPKFSELQGHLIKYPLQVDVDGKVSPLPDYETFPDVGGKIIGAHSmALPDTLTT  868
RCPLD- δ   769  LFDEPETLDCVKTVNKIAEDNWRRFTEEDFTPLQGFLLKYPLEVDRNGKVSPLTGQENFPDVGGKVLGARS-TFPDSLTT  847
```

Figure 7: SBIP-436 NCBI Multiple Alignment: Amino acids from SBIP436 (First Row), *Arabidopsis thaliana* PLD ´ – ATPLD (Second Row), and *Ricinus communis* PLD ´ – RCPLD (Third Row) are aligned, displaying their similarity. Highlighted in red is a conserved HxKxxxxD motif representative of the PLD catalytic site.

Further analysis of the alignment showed that an HxKxxxxD Binding Motif associated with PLDs was conserved in the SBIP-436 sequence.

Phospholipases

Phospholipases are present in all living organisms and are known to play important roles in biological membranes. Phospholipases are a group of enzymes that hydrolyze phospholipids producing a wide array of various lipids and molecules. The phospholipases are classified by their action site on the phospholipid molecule, their regulation, function, and mode of action. The phospholipases A (PLAs) are acyl hydrolases classified according to their hydrolysis of the l-acyl ester (PLAI) or the 2-acyl ester (PLA2). Phospholipase C (PLC) cleaves the glycerophosphate bond. Phospholipase D cleaves the head group from the phospholipid. Both phospholipases C and D are considered phosphodiesterases (Vance 2008). These 4 enzymatic groups are considered the phospholipase subfamilies. The catalytic function of phospholipases is outlined in Figure 8.

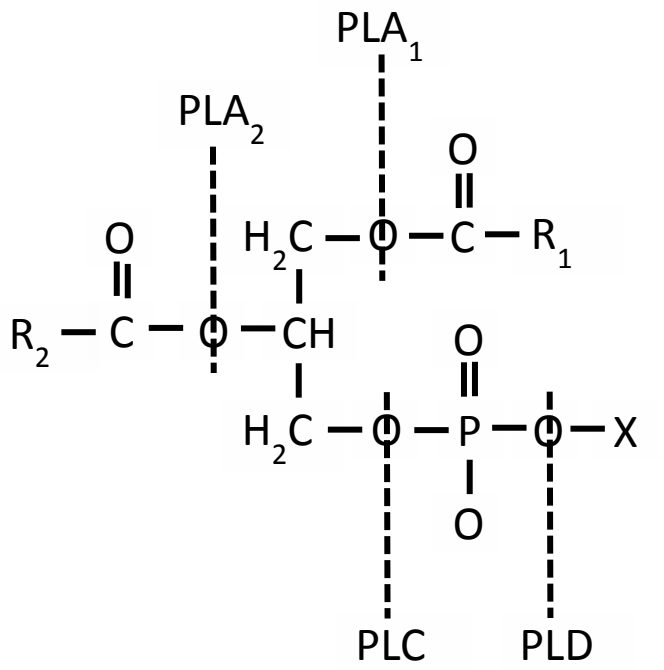


Figure 8: Phospholipase subfamily cleavage sites. Phospholipase A_1, Phospholipase A_2, Phospholipase C, Phospholipase D.

Phospholipase D

The subfamily Phospholipase D (PLD) can be separated into further classes based on structural design, including domain structure, biochemical properties, and sequence similarities. Although diverse, the PLD subfamily share a common link with 2 cellular regulators: Ca^{2+} and phosphatidylinositol 4,5-bisphosphate (PIP_2). There are 12 genes that have been isolated and characterized into 5 isoform classes in *A. thaliana*: PLD-± {1,2,3,4}, PLD-² {1,2}, PLD-³ {1,2,3}, PLD-´, PLD-¶ {1,2}. While all PLDs have mutual regulators, some can work independently (Wang 2002). PLD-² and PLD-³ are both PIP_2 dependent contrary to PLD-±, which can be activated independent of PIP_2. The newest addition to the PLD family, PLD-¶, does not require Ca^{2+} for activation. PLD-´ is unique in regards to other PLDs in that it is activated by oleic acid-18:1 and is exclusively bound to the plasma membrane (Wang 2003). Another detail that can better distinguish differences between the 5 PLD subclasses is their individual substrate specificity. PLDs hydrolyze a range of common membrane phospholipids such as phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), and phosphatidyl-Ser (Qin and Wang, 2002).

Phospholipase D-´ a Novel PLD

PLD-´ is a phospholipase that hydrolyses structural phospholipids to produce PA and free head groups. Although PLD ´ has a conservative C2 Ca^{2+} binding domain and is stimulated by PIP2 just like other PLDs, PLD ´ has unique qualities that set it apart from other PLDs (Wang, 2002). One unique quality of PLD ´ is that it is activated by Oleic Acid (18:1). It is thought that this distinctive interaction between oleic acid and PLD ´ may affect substrate staging and overall modify its enzyme-substrate interaction (Wang, 2001, 2002). The specificity of oleic acid activation has been used as a determining factor in gene silencing of PLD ´. This

transcends towards another unique quality of PLD ´: i.e. substrate preference. PLD ´ can effectively catalyze both phosphatidylcholine and phosphatidylethanolamine. PLD ´•s kinetic affinity for phosphatidylethanolamine is 7-9 fold higher than that to phosphatidylcholine. PLD ´ catalytically cleaves phosphatidylethanolamine into phosphatidic acid and ethanolamine. It was also shown that PIP2 was not required to activate PLD ´ but did stimulate its activity (Wang 2002). PLD ´ can also be activated by the reactive oxygen species H_2O_2, and this activation enhances plant resistance to H_2O_2-induced cell death (Zhang 2003). PLD ´ is also the only PLD that is specifically membrane bound. It is bound either on the plasma membrane or to microtubules (Zhang 2003; Li et al. 2005).

PLD ´ is a necessary protein in several defense mechanisms. PLD ´•s mechanistic characteristics predominantly apply to abiotic stressors, but the effects produced from biotic and abiotic attacks in the plant can overlap significantly and be modulated accordingly through the crossing of signal pathways in order to reach the overall goal of SAR. PLD ´ has been shown to be an active player in osmotic stress, high salt stress, heat stress, cold acclimation, and microtubule cytoskeleton reorganization (Li et al. 2005; Zhang et al. 2003). PLD ´•s cold acclimation mechanism was researched by Li•s group. They found that PLD ´ Knock Out (KO) plants were very sensitive to low temperatures. At –7 °C essentially all PLD ´ KO plants were killed while nearly all wild type plants survived. To confirm that PLD ´ was the key factor in the freezing tolerance in *A. thaliana*, they complemented the PLD ´ KO plants with the wild type PLD ´ gene. The complemented-PLD ´ KO plants reverted to wild type and were more tolerant to freezing conditions. To further confirm their postulations, they over-expressed the PLD ´ gene in *A. thaliana*. The PLD ´ Over Expressed (OE) plants and the wild type were subjected to –10

°C (Li et al. 2005). Nearly all the wild type plants were killed while all of the PLD ´ Over-expressing (OE) plants survived (Zhang and Wang 2004). The premise that the insensitivity to H_2O_2 related cell death may be related to the freezing tolerance mechanism of PLD ´ was also tested in the same study. To test this, they compared the effect of H_2O_2 on cell death of PLD ´ OE to that of wild-type and to that of PLD ´ KO plants. They used leaf protoplasts isolated from nonacclimated and cold-acclimated plants. The nonacclimated plants resulted in the same results stated previously except for the PLD ´ OE plants that exhibited 30% less H_2O_2 damage, making them more tolerant not only to abrupt freezing conditions but to H_2O_2 related cell death as well (Zhang and Wang 2004). Possibilities were suggested regarding PLD ´•s freezing tolerance mechanism. Being that PLD ´ is specifically bound to the plasma membrane as well as the microtubule cytoskeleton, it was suggested that it increases tolerance to freezing-dehydration of the cells by reorganizing the membranes and the overall cell structure for better support (Zhang and Wang 2004). In metaphor, it could be viewed as the field-medic in the plant rebuilding the cell structure to the former strength after inflicted damage. Also, if PLD ´ has the ability to choose substrates, it may affect the overall composition of the membranes, then PLD ´'s preferred substrate being phosphatidylethanolamine, the membrane may be more tolerant to stressors than a membrane in deficit of this compound. We believe this to be the same mechanistic process that coincides within the SA mediated signaling pathway resultant by either abiotic or biotic stressors.

PLD ´•s has been shown to be the only PLD to be activated by oleic acid (Wang and Wang 2001). This specific oleic acid activated PLD ´ maintains the ability to play a role in response to various cellular stresses. In particular the oleic acid activated PLD ´ plays a role in

H_2O_2 mediated cell death. In a study by Wang's group this role of PLD ´ was studied in *Arabidopsis*. They found that H_2O_2 actually stimulated PLD ´ and produced a signal; presumably by PLD ´ derived phosphatidic acid (PA) where PA acts as a secondary messenger to down regulate H_2O_2 mediated cell death. The ablation of the PLD ´ rendered the *At*-cells more sensitive to H_2O_2. Complemented PLD ´ KO plants with the PLD ´ gene with its own promoter showed reduced sensitivity to H_2O_2, promoting anticell death characteristics. An anticell death model is proposed in Figure 9.

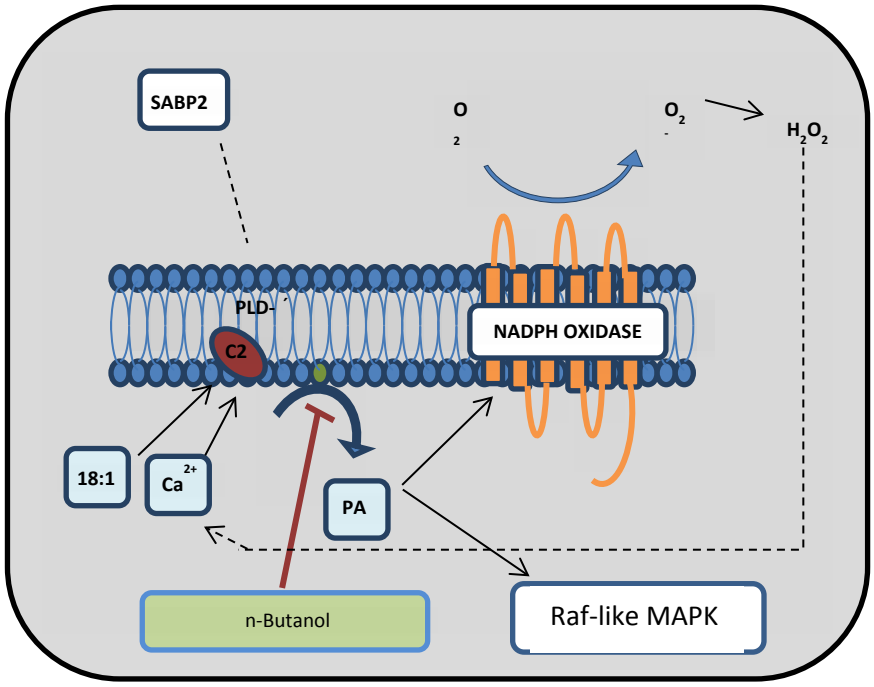


Figure 9: PLD ´ Anticell death model with possible interaction with SABP2. Hypothetical model of PLD ´•s involvement in H_2O_2 regulated anticell death via transphosphatidylation. The presence of a primary alcohol (N-butanol) can cause transphosphatidylation inhibiting the production of the secondary messenger Phosphatidic Acid (PA). The lack of PA reduces activity of NADPH oxidase's production of H_2O_2.

PLD ´ KO plants had an indistinguishable phenotype from wild-type plants (Zhang et al. 2003). This may be because that there are 12 different PLDs performing overlapping functions

compensating for the lack of PLD ´ function and they can all produce PA, the second messenger. Although the PA is being produced by other PLDs, the oleic acid activated PLD-PA is only produced through PLD ´.

PLD ´ distinctive localization creates a special niche for its activity. It is bound to the plasma membrane as well as the microtubule cytoskeleton. A *Nicotiana tabacum* 90kD Phospholipase D (p90) was identified that binds to both the plasma membrane and microtubules, which is a unique characteristic of PLD ´. *At*-PLD ´ was found to be expressed in roots, leaves, stems, and flowers but not in siliques. While under dehydration stress it was strongly expressed in vascular tissues of cotyledons and leaves (Katagiri et al. 2001). The full length gene in *Arabidopsis* was characterized as a polypeptide consisting of 857 amino acids with a molecular mass of 97.7 KD and pI of 6.7 (GenBank accession number AF306345). It was deemed AtPLD98. With the addition of a C2 domain the length of the polypeptide reaches 868 amino acids, which is conserved across all PLDs, except PLD¶ (Katagiri et al. 2001) The deduced amino acid sequence was found to be identical to At- PLD ´ (GenBank accession number AF322228). The AtPLD98 contains a double HxKxxxD motif (HQKCVLVD/HAKGMIVD) that is thought to be the phospholipid-metabalizing catalytic center conserved in all PLDs. These are located at residues 362 to 369 and 707 to 714 respectively (Punting and Kerr 1996; Wang 2000). PLDs have a C2 domain that is a calcium dependent lipid binding domain. This C2 domain was found to lie at residues 75-107 in the *At*PLD98. Immunofluorescence microscopy showed that the *N. tabacum* 90kD Phospholipase D binds to cortical microtubules in BY-2 cells (Gardiner et al. 2001). Another study demonstrated that activation of PLD ´ affected the organization of microtubules. Using BY-2 cells treated with n-butanol, a potent activator of PLD ´, resulted in

release of microtubules from the plasma membrane as well as partial depolymerization of the microtubules. They also showed that only n-butanol activated PLD ´, but not sec- or tert-butanol. This demonstrates that PLD ´ is activated by primary alcohol (Munnik et al. 2003). When PLD ´ uses a primary alcohol it goes through a process known as transphosphatidylation rather than its usual hydrolytic function. In transphosphatidylation the phosphatidyl group of phosphatidylcholine is transferred to an alcohol such as – methanol, ethanol, 2-propanol, glycerol, ethanolamine, and serine. This action produces a phosphatidyl-alcohol and choline in contrast to the hydrolytic products of PA and choline (Yang et al. 1966). This alcohol induced activation of PLD ´ disassociates itself from the microtubules and stays with the plasma membrane allowing for microtubule reorganization. Activation of PLD ´ and microtubule reorganization has been shown to be induced by pathogen elicitors and osmotic stress as well as defense signaling in response to these stressor by SA or ABA (Munnik et al. 2003; Rainteau et al. 2012). It has been shown that when microtubules rearrange following pathogen/fungus attacks, they tend to rearrange in a radial array toward the site of penetration (Kobayashi et al. 2007). The directional reorganization of microtubules coupled with PLD ´•s disassociation from the microtubules, leading to this effect, as well as PLD ´•s continued association with the plasma membrane may lead to incorporation of the pathogen to the host plant and localize its negative effect to this area.

Hypotheses

Hypothesis I:

Tobacco SBIP-436 is *Nicotiana tabacum* PLD-´.

If the full SBIP-436 is cloned, retains accepTable homology to *N. benthamiana* PLD-´, and is shown to be activated by oleic acid, the tobacco SBIP-436 is *Nicotiana tabacum* PLD-´.

If the SBIP-436 is cloned, retains accepTable homology to *N. benthamiana* PLD-´, and is shown to NOT be activated by oleic acid, tobacco SBIP-436 should be considered as an alternate PLD.

Hypothesis II:

SBIP-436 expression is differentially modulated during pathogen infection and abiotic stress.

If SBIP-436 displays a change in expression from its baseline expression after TMV inoculation and changes are also demonstrated after NaCl or wounding treatment, SBIP-436 is involved in pathogen infection and abiotic stress.

If SBIP-436 displays a change in expression from its baseline expression after TMV inoculation and no changes are demonstrated after NaCl or wounding treatment, SBIP-436 is involved in pathogen infection and not abiotic stress.

If SBIP-436 displays a change in expression from its baseline expression after NaCl or wounding treatment and no change is demonstrated after TMV inoculation, SBIP-436 is involved in abiotic stress and not pathogen infection.

CHAPTER 2

MATERIALS AND METHODS

Materials

Plant Materials

The *Nicotiana. tabacum* cv. *Xanthi nc NN* (XNN) was used as a control to the experimental line NahG (expressing *nahG* gene in XNN, that encodes SA hydroxylase that converts SA to catechol). The C3 (contains empty pHANNIBAL vector in XNN) was used as a control to the experimental 1-2 (SABP2 silenced in XNN) line. All plants were grown and maintained in a growth chamber (PGW 36, Conviron, Canada) set to a 16-h day cycle with a constant temperature of 22°C. Seeds were germinated in autoclaved (20 min) soil containing peat moss (Fafard Canadian growing mix F-15, Agawam, MA). After 14 days 2 tobacco seedlings were transferred to each of 4 X 4 inch flats. After 30 days, young plants were transferred to a pot 8 inches in diameter. Fertilizer was added 3 days following the final transfer. Six to 8 weeks old plants were used to perform the experiments.

Chemicals and Reagents

Magnesium sulphate ($MgSO_4$), magnesium chloride ($MgCl_2$), sodium chloride (NaCl), sodium phosphate monobasic (NaH_2PO_4), sodium phosphate dibasic (Na_2HPO_4), ammonium sulfate ($(NH_4)_2SO_4$), Glycerol, Chloroform, Isopropanol, and all other standard chemicals were obtained from Fisher Scientific, Pittsburgh, PA. Proteose peptone 3 (Becton and Dickenson), Agar, Polyvinylpolypyrrolidone (PVPP) were obtained from Acros Organics, Audubon Park, NJ. Semipurified Tobacco mosaic virus (TMV) was available in-house. Oligo dT-20, Taq DNA polymerase (Invitrogen, CA), DNA ladder (New England Biolabs), dithiothreitol (DTT), MMLV

reverse transcriptase, recombinant RNAsin, RNAse free DNAse, (Promega), Agarose (SeaKem), Gel loading dye (Bio-Rad), Tris-Acetate-EDTA (TAE), TRI reagent and Diethyl pyrocarbonate (Sigma-Aldrich), Rifampicin (Phytotechnology), Liquid Nitrogen (Airgas, TN). Qiagen gel extraction kit, QIAprep Spin Miniprep Kit (Qiagen). pGEMT (Promega), TOPO (Invitrogen) were purchased.

Methods

Bioinformatics Analysis

Sequence Alignments and Database Analysis. SBIP-436 sequence analysis was performed using multiple bioinformatics tools. Among the tools used were BLAST (Basic Local Alignment Search Tool) (Altschul et al. 1990) at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi), ExPASy Bioinformatics Resource Portal (http://www.expasy.org/tools/), and NCBI ORF (Open Reading Frame) Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The SBIP-436 yeast -2-hybrid clone was found to be a partial sequence so the Sol Genomics database (http://solgenomics.net/tools/blast/index.pl) was used to build a possible gene construct for the unannotated *N. tabacum* PLD´. The sequences found were aligned to each other and known sequences using the Clustal W2 tool (ClustalW2); http://www.ebi.ac.uk/Tools/msa/clustalw2/). A DNA sequence cleaner was used to organize and format sequences for ease of use and storage (http://www.cellbiol.com/scripts/cleaner/dna_protein_sequence_cleaner.php). Two splice site prediction tools were used to analyze the partial sequence clone of SBIP-436. The first was Wang Computings ASSP (Alternative Splice Site Predictor) (http://wangcomputing.com/assp/evaluation.html) and the second used was Splice Port (http://spliceport.cbcb.umd.edu/) for comparison and validation.

Cloning in pGEMT and TOPO Vectors. The SBIP-436 partial sequence that was amplified using primers DK511 & DK514. Amplified products were separately cloned in PCR cloning vectors e.g. pGEMT and TOPO. Given PLD's nature of multiple isoforms, highly conserved domains, and implications of alternative splicing, it was found prudent to also distinguish their relation to one another, if any.

Gel Extraction

The gene of interest amplified via RT-PCR was analyzed via electrophoresis in a 1.2% agarose gel. Following quick visualization under UV, the DNA fragment was excised from the gel with a clean scalpel and DNA purified using Qiagen gel extraction kit following manufacturer's instructions. Eluted DNA was quantified using Nanodrop.

Ligation into pGEMT vector

The pGEMT vector to insert (Gel purified PCR product) molar ratio was calculated using following equation:

$$\frac{50\text{ng vector} \ * \ (\ )\text{kb insert}}{3.0kb\ vector} * \frac{3}{1} = (x)ng\ insert$$

The pGEMT vector was briefly centrifuged and the 2x rapid ligation buffer was vortexed vigorously. The following reaction was setup in a 1.7 ml microcentrifuge tube. Ligation reaction is shown in Table 1.

Table 1: pGEMT Vector Ligation Reaction

Reaction Compnonent	Amount
2x Rapid Ligation Buffer	5µl
pGEMT Vector	1 µl
PCR Product	(x) µl
T4 DNA Ligase (3Weiss units/µl)	1 µl
Nuclease-Free water to a final volume of -	10 µl

The reaction was mixed by pipetting and incubated at room temperature for 1 hour. Two microliter of the ligation reaction and 100 µl of DH5± competent cells were added to a new 1.7 ml microcentrifuge tube then placed on ice for 20 minutes. The tubes were briefly incubated for 45 seconds in a 42^{o}C waterbath and returned to ice for 2 minutes. Nine hundred microliter of LB (no antibiotic) was added to the tube. The tube was closed and its cap secured by parafilm, was dropped in a 250 ml flask, and placed in a 37^{o}C shaker for 1 hour. After shaking, the sample was centrifuged for 45 seconds at 10,000 rpm. Nine hundred microliter of supernatant was removed, and pellet was resuspended in remaining 100 µl by vortexing. X-gal (20 mg/ml) and 0.1 mM IPTG were spread on LB plate with sterile glass beads at least 20 minutes prior to plating of samples. One hundred microliter of culture was spread with glass beads and placed in 37^{o}C

incubator for 13-16 hours. Ten white and 2 blue isolated colonies were grown in 3 ml LB overnight at 37^{o}C in a shaker.

Isolation of Plasmid DNA - QIAprep Spin Miniprep Kit

The 3 ml overnight culture was pelleted by centrifuging at 10,000 rpm for 5 min in a 1.7ml microcentrifuge tube. Bacterial Pellet was used to isolate plasmid DNA using Qiagen min prep kit following manufacturer's instructions. Plasmid DNA was isolated in 25µl of Elution Buffer. Nanodrop was used to calculate the amount of plasmid DNA. 2-3µl of plasmid DNA was analyzed on an agarose gel and visualized under UV. Plasmid DNA were sent for sequencing.

Preparation and Treatment of Tobacco with Tobacco Mosaic Virus

N. tabacum (XNN) plants were grown for 6 weeks before using for infected with pathogens. The pathogen TMV was prepared to a 10^{-3} dilution of 45µg/ml in 0.05 M sodium phosphate buffer, pH 7.0. Carborundum was used as an abrasive to allow entry of TMV into the leaf. The upper fully expanded leaves were dusted uniformly with carborundum on the adaxial surface. A prewashed 4x4 inch cheesecloth (4-6 layers) soaked in a diluted TMV solution was used to gently rub against the carborundum dusted leaf. XNN plants infected with TMV were isolated from uninfected plants and maintained at a regular regimen of light (16 hr/day) and water (once every 2 days). Leaf samples were taken at desirable time points.

Preparation of the Bacterial Inoculums

The *N. tabacum* host-pathogen *Pseudomonas syringae* pv. *tabaci (Pst)* and the nonhost pathogen *P. syringae* pv. *phaseolicola (Psp)* were grown on King's B medium in a sterile petri dish. 25µg/ml of the antibiotic rifampicin was present in King's B medium used for growing *Psp.* Bacterial plates were incubated for 48hrs at 28°C. Isolated colonies were taken with a loop and inoculated into 10ml of King's B broth containing 1mM magnesium sulfate ($MgSO_4$). The culture was placed into a shaker revolving at 250 rpm at 28°C overnight. The culture was centrifuged for 15 minutes at 3,000 rpm at 10°C. The supernatant was decanted and the bacterial pellet was resuspended in 10 ml of 10 mM of $MgCl_2$. The culture was centrifuged for 15 minutes at 3,000 rpm at 10°C and the bacterial pellet was resuspended in 10 ml of 10 mM $MgCl_2$. The culture was centrifuged for 15 minutes at 3,000 rpm at 10°C and the pellet was resuspended in 20 ml of 10 mM $MgCl_2$. Bacterial concentration was determined by measuring the optical density (OD) at 600nm in a spectrophotometer. The $MgCl_2$ washed *Pst* and *Psp* were diluted to 0.2 OD_{600} (10^8 colony forming units (cfu)/ml). *Psp* was diluted to 10^6 cfu/ml with $MgCl_2$. *Pst* inoculum was diluted to 10^5 cfu/ml. Six weeks old C3 tobacco plants were used for inoculation with bacterial pathogens. The uppermost fully expanded leaves were slowly infiltrated with bacterial inoculum using a needleless 1 ml syringe. Inoculated plants were kept under proper lighted area. Three leaf discs were collected with a #5 cork bore at 0, 3, 6, 9, 12, and 24, and hpi (hours post inoculation) from inoculated leaves of XNN plants. Leaf samples were placed in 1.7ml microcentrifuge tubes and flash frozen in liquid nitrogen. The leaf samples were then stored at -80ºC until used for RNA isolation.

Treatment of Tobacco with Exogenous Salicylic Acid

For SA treatments the uppermost fully expanded leaves were selected for infiltration. A 1ml sterile, needleless syringe was used to slowly infiltrate the 0.1 mM SA throughout the leaf tissue. Three leaf discs were collected with a #5 cork bore at 0, 3, 6, 9, 12, and 24, and hpt (hours post treatment) from inoculated leaves of XNN plants. Samples were placed in 1.7ml microcentrifuge tubes and flash frozen in liquid nitrogen. The leaf samples were then stored at -80ºC until ready for total RNA isolation.

Treatment of Tobacco with NaCl

N. tabacum XNN leaves were treated with 300mM NaCl as described above for SA treatments. Leaf samples were collected and stored at -80C.

RNA Isolation

To isolate total RNA, 3 leaf discs were ground to fine powder using liquid nitrogen. Once sample was powdered, 500µl of TRIzol was added, gently mixed, and 500µl more was added. Sample was kept for 5 minutes followed by the addition of 200µl of chloroform and gently mixed for 15 seconds. After 2 minutes sample was centrifuged at 12000Xg for 10 minutes at 4ºC. The aqueous phase was removed to a new tube and 500µl isopropanol was added to it. Sample was centrifuged again at 12000Xg for 10 minutes in 4ºC. Supernatant was discarded and to the RNA pellet, 1 ml of 75% cold ethanol was added and centrifuged again at 7500Xg for 5 minutes 4ºC. The resultant pellet was air dried in hood for 15 minutes. The pellet was then resuspended in 43µl of DEPC, 5µl of DNAse buffer, and 2ul of DNAse (50µl DNAse mix). All steps starting with TRIzol were repeated with half volumes until the resultant pellet step. The

pellet was resuspended with 20µl of DEPC treated water and heated in a 55 ºC waterbath. RNA was quantified by taking Optical density (OD) using Nano Drop spectrophotometer and stored at -80 °C until cDNA synthesis.

cDNA Synthesis

Total RNA was used to synthesize the first strand complementary DNA (cDNA). To1 µg of total RNA in 8 µl DEPC treated water, 2 µl of oligo-dT_{20} (0.5 µg/ml) was added. The mixture was incubated at 75^{o}C for 10 minutes in the thermocycler (Eppendorf) and cooled to 4^{o}C. The following reaction mixture was added to the 10 µl RNA+Oligo-dT mix for a total volume of 20µl. cDNA synthesis reaction is shown in Table 2.

Table 2: cDNA Synthesis Reaction

Reaction Component	Amount
5X RT Buffer	4 µl
Diethylpyrocarbonate -treated Water (DEPC)	2 µl
Deoxynucleoside triphosphates (dNTP)	1 µl
Dithiotthreitol (DTT)	1 µl
RNAse Inhibitor (RNasin)	1 µl
Reverse Transcriptase (RT)	1 µl

The thermocycler was used to incubate the sample at 42^{o}C for 60 minutes followed by an incubation at 70^{o}C for 10 minutes. After incubation the cDNA sample was stored in -20^{o}C for future analysis.

Polymerase Chain Reaction (PCR)

The cDNA sample was used to amplify gene of interest for cloning and expression analysis. The following 10 µl of PCR mix was used to amplify each gene. PCR reaction is shown in Table 3.

Table 3: PCR Reaction

Reaction Component	Amount
Nuclease free autoclaved water	6.4µl
10X dNTP	1µl
10X Taq Polymerase Buffer	1µl
Taq Polymerase	0.2µl
Forward Primer (10µM)	0.2µl
Reverse Primer (10µM)	0.2µl
cDNA sample	1µl

PCR involves cycles that consist of 3 steps: Denaturation, Annealing, and Extension. Hot-Start PCR was used for heat activation of the DNA polymerase. This requires that there is a one-time initialization step preceding the multiple cycles where samples are heated to 94^{o}C for approximately 2 minutes. The first cycling event is denaturation. The reaction is heated to 94^{o}C for 30 seconds, which melts the DNA template, separating complementary bases, yielding single-stranded DNA. The second cycling event is annealing. The annealing temperature (Ta) is calculated by the melting temperature (Tm) of the synthesized primers for the gene of interest. While the Ta varies (50^{o}-65^{o}C) the time is constant at 30 seconds. The third cycling event is extension. Contrary to the annealing step, the extension temperature is constant at 72^{o}C to optimize polymerase activity and the extension time varies in accordance with the length of the gene of interest (1 minute per 1 kilobase (kb)). Cycle number (28-40) is dependent on the use and presence of the gene of interest. Following the cycling events there is a final elongation at 72^{o}C for 8 minutes and a final hold between 4^{o}-10^{o}C. The forward and reverse primers were used to amplify SBIP-436 are shown in Table 4.

Name	Function	Primer	Direction	Bases	Product Length	Tm	GC %
DK511	Gene Expression	5'TCAGCTTCCCAGAAATTTGGA 3’	Forward	21	200bp	58.7	42.9
DK513	Gene Expression	5' ATACCTGGCCATGTGGATGT 3’	Reverse	20	200bp	60.4	50
DK514	Gene Expressio	5' TGTGGTCAAAGCATCAGGA	Reverse	20	500/700b	58.4	45

	n	A 3’			p		
DK547	Cloning (Gateway)	5' GGGGACAAGTTTGTACAAAAAAGCAGGCTTGATGGCGGATGAGAATTGTGAA 3'	Forward	52	~2400bp	73.4	44.2
DK548	Cloning (Gateway)	5' GGGGACCACTTTGTACAAGAAAGCTGGGTTTCATGTGGTCAAAGCATCAGG 3'	Reverse	51	~2400bp	75.2	49
DK618	Transcript Insert	5' ACCTGCAGCCGCACTTGTTAA 3'	Reverse	21	170bp	62.6	52.4
DK619	Transcript Insert	5' TTCCCTCATTCTCCTTCCCAA 3'	Forward	21	170bp	60.6	47.6
DK602	1000bp segment	5' ATGGCGGATGAGAATTGTGAA 3’	Forward	21	1000bp	58.7	42.9
DK603	1000bp segment	5' GCTTACTGCTTGCATAACGAG 3’	Reverse	21	1000bp	60.6	47.6
DK604	1000bp segment	5' AGCACTCATCCGTCACTTGTG 3’	Forward	21	1126bp	62.6	52.4
DK605	1000bp segment	5' GGGAAGCTGAGATCACATCAC 3’	Reverse	21	1126bp	62.6	52.4
DK515	5' RACE	5' CGTCCACTATCATCCCCTTGGCGTGT 3'	Reverse	26	N/A	69.3	57.7
DK516	5' RACE	5' TCGCCACTTGAGGAAGCAGGGTTTG 3'	Reverse	25	N/A	67.9	56

Table 4: SBIP-436 Primer

CHAPTER 3

RESULTS

Gene Analysis

DNA and Amino Acids Corresponding to SBIP-436

A Yeast-2-hybrid screening, using SABP2 as bait, yielded multiple positive interacting proteins. SBIP-436 was among these interacting proteins. The SBIP-436 sample was sequenced and analyzed. The sequenced cDNA revealed a clone 228bp in length. The sequence, family grouping, and clustal alignment are shown in Figures 4, 5, and 6 respectively. Next, a multiple sequence alignment was performed using the SBIP-436 corresponding 3rd frame amino acid sequence and *A. thaliana* (AtPLD-´) as well as *Ricinus communis* (RcPLD- ´). This showed that SBIP-436 has high homology to phospholipase D-´. These species were chosen due to their high homology as well as their involvement in previous phospholipase D ´ studies. SBIP-436 showed 70% similarity with AtPLD-´ and 71% similarity with RcPLD- ´.

NCBI Multiple Alignment with AtPLD-´ and RcPLD- ´

```
SBIP-436   1    NYFCLMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDD  75
ATPLD- δ   709  NFQRFMIYVHAKGMIVDDEYVLMGSANINQRSMAGTKDTEIAMGAYQPNHTWAHKGRHPRGQVYGYRMSLWAEHLGKTGD  788
RCPLD- δ   689  KHQRFMIYVHAKGMIVDDEYVIMGSANINQRSMAGSRDTEIAMGAYQPNHTWGNKKRHPRGQVYGYRMSLWAEHLGLVDS  768

SBIP-436   76   CFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGART-NLPDALTT  154
ATPLD- δ   789  EFVEPSDLECLKKVNTISEENWKRFIDPKFSELQGHLIKYPLQVDVDGKVSPLPDYETFPDVGGKIIGAHSmALPDTLTT  868
RCPLD- δ   769  LFDEPETLDCVKTVNKIAEDNWRRFTEEDFTPLQGFLLKYPLEVDRNGKVSPLTGQENFPDVGGKVLGARS-TFPDSLTT  847
```

Figure 10: SBIP-436 NCBI Multiple Alignment: Amino acids from SBIP436 (First Row), *Arabidopsis thaliana* PLD ´ – ATPLD (Second Row), and *Ricinus communis* PLD ´ – RCPLD (Third Row) are aligned, displaying their similarity. Highlighted in red is a conserved HxKxxxxD motif representative of the PLD catalytic site.

Further analysis of the alignment showed that an HxKxxxxD Binding Motif associated with PLDs was conserved in the SBIP-436 sequence. This BLAST and alignment analysis demonstrated that the sequenced clone was a partial PLD-´ sequence. In accordance with this a BLAST analysis was performed in the Sol Genomics database revealing that the *N. tabacum* PLD-´ has yet to be annotated or identified.

Therefore it was a goal to identify the full corresponding gene sequence in *N. tabacum*. In order to do this the yeast-2-hybrid SBIP-436 sequence to BLAST in the Sol Genomics database against the *N. tabacum* genome. First, the SBIP-436 sequence was used in a BLAST using the *N. tabacum* Unigene sequence set. The results found a slightly larger sequence SGN-U444527, Figure 11, with a length of 1,726 nt and revealed that the longest ORF (Open Reading Frame) was in the -3 frame consisting of 193 amino acids.

```
NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANK
KKHPHGQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPD
VGGKVLGARTNLPDALTT
```

Figure 11: BLAST in Sol Genomics *N. tabacum* Unigene database using the SBIP-436 Yeast Two-Hybrid clone sequence yielded SGN-U444527 ORF (193aa).

CLUSTAL alignment of the SBIP-436 Yeast-2-hybrid Clone and the *N. tabacum* SGN-U444527 sequence yielded an exact match. Their clustal alignment is shown in Figure 12.

```
   Nt  SGN-           NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 60
SBIP-436 Y2H Clone   ---------------------------SASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 33
                                                 ********************************

   Nt  SGN-           QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPESLD 120
SBIP-436 Y2H Clone   QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQV------------------------- 67
                     *********************************

   Nt  SGN-           CVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGA 180
SBIP-436 Y2H Clone   ------------------------------------------------------------

   Nt  SGN-           RTNLPDALTT 190
SBIP-436 Y2H Clone   ---------- 
```

Figure 12: CLUSTAL 2.1 multiple sequence alignment of SBIP-436 Yeast-2-Hybrid Clone (67aa) and *N. tabacum* SGN-U444527 (190aa).

From what is known from the Arabidopsis PLD ´ sequence, it was expected to find a gene sequence consisting of approximately 3,100 nucleotides (ATPLD ´ 3,125 nt: accession#-NM_179170). The goal was to find a full-length tobacco PLD ´ sequence corresponding to the SBIP-436 Y2H clone as well as the newly found SGN-U444527. No additional full length unigene could be identified. Only small segments were being detected. The SGN-U444527 sequence was further studied, specifically in attaining the amino acid sequence matched with it. As stated earlier the ATPLD ´ consists of 868 amino acids. It is known that the C-terminus of the PLDs is highly conserved. Therefore, the smaller the partial and the closer it lies to the C-terminus results in less probability of positive confirmation of a specific PLD. It was important to understand the variability of the PLDs and use the information to support the identification of SBIP-436. Because neither the *N. tabacum* "Methylation Filtered Genome" nor the Unigene database resulted in any significant findings, we designed and performed a "Genome Crawl".

Genome Crawl

The genome crawl, demonstrated in Figure 13, was designed to take a known sequence like ATPLD´ and attain the sequence corresponding to Nt-PLD´ in an organism closer related phylogenetically. This was repeated in a cascade fashion until the organism of interest *N. tabacum* was searched using preceding sequences.

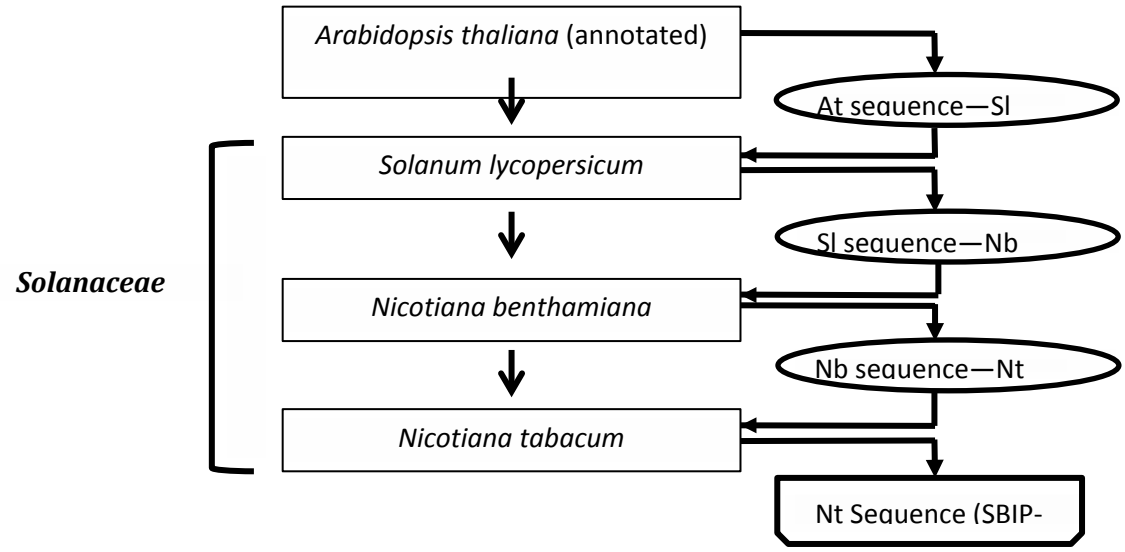


Figure 13: Genome Crawl. The ATPLD´ sequence was used to extract the *S. lycopersicum* PLD´ (S1PLD´). The SlPLD´ was used to attain the *N. benthamiana* PLD´. The NbPLD´ was used search for the *N. tabacum* PLD´ (NtPLD´).

The ATPLD´ sequence (AT4G35790.1) was taken from the TAIR database and subjected to blast in the Sol Genomics dataset - Tomato proteins (ITAG release 2.40). This resulted in a *S. lycopersicum* PLD´ (Solyc02g083340.2.1) construct consisting of 866 aa shown in Figure 14.

AT4G35790.1

ATGGCGGAGAAAGTATCGGAGGACGTTATGCTTCTACACGGTGACCTCGATTTGAAAATTGTTAAGGCTAGGAGGTTACCTAACATGGATATGTTCT
CAGAACATTTGCGCCGTCTCTTCACCGCCTGTAACGCCTGTGCTAGACCCACCGATACCGATGATGTCGATCCCAGAGATAAAGGCGAATTCGGTGA

TAAAAACATCCGTAGCCACCGTAAGGTTATCACCAGCGATCCTTACGTCACCGTCGTCGTTCCTCAAGCGACTCTAGCTCGAACACGTGTTTTGAAA
AACTCACAAGAGCCTCTTTGGGACGAGAAATTCAACATTTCTATAGCGCATCCGTTTGCTTATCTCGAGTTCCAGGTCAAAGACGACGATGTTTTCG
GTGCTCAAATCATTGGCACGGCTAAGATCCCGGTTCGAGACATCGCATCGGGAGAACGCATTTCTGGTTGGTTTCCTGTACTCGGTGCTTCGGGAAA
ACCGCCTAAGGCAGAAACTGCTATTTTCATCGATATGAAATTTACTCCGTTTGACCAGATCCATAGCTACCGATGTGGAATCGCCGGAGATCCGGAG
CGTAGGGGTGTTAGACGGACTTATTTCCCTGTGAGGAAAGGAAGTCAGGTGAGGCTTTACCAGGACGCTCATGTTATGGACGGAACGTTGCCGGCGA
TTGGGTTAGATAACGGGAAAGTTTATGAGCATGGGAAGTGTTGGGAAGATATATGTTATGCTATATCTGAGGCTCACCATATGATTTACATTGTTGG
TTGGTCTATCTTCCATAAGATTAAGCTTGTTAGGGAAACAAAAGTTCCAAGAGATAAGGATATGACGCTTGGGGAATTGCTCAAATACAAATCCCAG
GAAGGTGTTCGAGTTTTGCTACTTGTATGGGATGATAAGACTTCTCATGATAAGTTTGGGATAAAAACGCCTGGAGTTATGGGGACACATGATGAGG
AGACTAGGAAGTTCTTCAAGCATTCTTCTGTGATATGCGTTTTGTCACCCCGGTATGCCAGCAGTAAGCTTGGGTTGTTCAAACAACAGGCAAGTCC
TAGCTCTTCTATATATATCATGACAGTTGTTGGAACTCTCTTCACGCACCATCAAAAGTGTGTTCTTGTAGACACTCAGGCTGTTGGTAATAATCGC
AAGGTCACCGCTTTTATTGGAGGTCTAGATCTTTGTGATGGCCGTTATGACACACCTGAGCATAGGATACTCCACGATCTTGACACTGTATTTAAGG
ATGATTTCCACAATCCTACATTTCCAGCTGGTACCAAGGCCCCAAGACAACCTTGGCACGATTTGCACTGTAGGATAGATGGGCCTGCGGCATATGA
TGTTCTCATAAACTTTGAGCAGCGATGGAGAAAGGCAACACGATGGAAGGAGTTTAGCTTACGTTTAAAGGGGAAAACTCACTGGCAAGATGATGCT
TTGATCCGGATAGGGCGTATATCATGGATTCTGAGTCCAGTGTTTAAATTTCTGAAGGATGGTACTTCGATAATTCCAGAGGACGATCCATGTGTTT
GGGTTTCTAAAGAAGATGATCCAGAGAACTGGCATGTTCAGATATTCCGTTCTATCGACTCAGGATCCGTGAAAGGATTTCCAAAATATGAAGATGA
GGCTGAGGCCCAGCATCTGGAATGTGCCAAGCGTCTTGTTGTAGATAAAAGCATCCAGACTGCATACATCCAGACAATCAGATCTGCTCAGCATTTC
ATATATATCGAGAATCAGTATTTCCTGGGTTCTTCTTATGCTTGGCCTTCTTATAGAGACGCAGGAGCTGACAATCTTATTCCTATGGAGTTGGCAC
TAAAGATTGTTAGCAAAATCAGAGCTAAGGAAAGATTTGCCGTATATGTTGTCATACCATTGTGGCCTGAAGGCGACCCAAAGTCTGGCCCTGTGCA
AGAAATTCTATATTGGCAGAGCCAAACTATGCAGATGATGTATGATGTTATAGCAAAAGAACTGAAAGCGGTGCAATCAGATGCTCATCCTCTCGAT
TACCTTAACTTTTACTGCCTTGGTAAACGAGAGCAGCTTCCAGATGATATGCCAGCCACCAATGGCAGTGTGGTATCAGATTCTTATAATTTCCAGC
GTTTCATGATTTACGTGCACGCAAAAGGGATGATAGTAGATGATGAGTATGTACTCATGGGATCTGCTAATATCAACCAAAGATCTATGGCAGGCAC
CAAAGATACTGAAATCGCCATGGGCGCATACCAACCTAATCATACATGGGCTCACAAGGGAAGACACCCACGTGGCCAGGTGTATGGATACAGAATG
TCACTATGGGCAGAGCATTTAGGCAAAACTGGAGATGAGTTTGTGGAGCCATCAGATCTGGAATGTCTGAAGAAGGTGAACACAATCTCTGAAGAAA
ACTGGAAAAGATTCATAGACCCGAAATTCTCAGAGCTACAAGGTCACTTAATAAAGTATCCTCTACAAGTAGACGTTGATGGTAAAGTAAGCCCTCT
TCCTGATTACGAGACCTTCCCAGATGTTGGTGGTAAGATCATTGGAGCTCATTCCATGGCTCTTCCTGACACTTTAACCACGTAA

Solyc02g083340.2.1

MAENSSQENFICLHGDLELHIIQARHLPNMDLTSERIRRCFTACDVCRKPQTGSTADDGNGELPNVKSTDQKIHHRSIITSDPYVAVCAPHTALART
RVIPNSQNPVWDEHFRIPLAHPMDCLDFRVKDDDVFGAQVMGKVTIPAEKIASGEVVSGWFPVIGASGKSPKPDTALRLWMKFVPYDTNPLYKRGIA
SDPQYLGVRNTYFPLRKGSSVKLYQDAHVSDKFKLPEIQLENNTTFEHNKCWEDICYAITEAHHLIYIVGWSVFHKVKLVREPTRPLPRGGDLTLGE
LLKYKSQEGVRVLLLVWDDKTSHDKFFINTAGVMGTHDEETRKFFKHSSVICVLSPRYASSKLSLIKQQVVGTMFTHHQKCVLVDTQAPGNNRKVTA
FLGGLDLCDGRYDTPEHRLFHDLDTVFKDDVHQPTFPAGTKAPRQPWHDLHCRIDGPAVYDVLINFAQRWRKATKWREFKFFKKTMSHWHDDAMLKI
ERISWILSPAFAVLKDSTAIPEDDPKLHVYGEDHSENWHVQIFRSIDSGSVQGFPKTIDVAQAQNLVCSKNLMVDKSIEAAYIQAIRSAQHFIYIEN
QYFLGSSYAWESYKDAGADHLIPMELALKITSKIRARERFCVYVVVPMWPEGDPKSITMQEILFWQSQTIQMMYQVIATELKSMQILDSHPQDYLNF
YCLGNREEIPGSIAQSSGNGDKVSDSYKFQRFMIYVHAKGMIVDDEYVIMGSANINQRSLAGSKDTEIAMGAYQPHYAWTEKQRRPRGQIYGYRMSL
WAEHLGRIEECFKEPEALTCVRKVNEVAEGNWKSYTAEKFTQLHGHLLKYPIHVGADGKVGPLAEYENFPDVGGRILGNHAPTIPDVLTT

Figure 14: The ATPLD´ sequence (AT4G35790.1) *Top* subjected to BLAST in Sol Genomics Tomato proteins (ITAG release 2.40) and resulted in *S. lycopersicum* PLD´ (Solyc02g083340.2.1) *Bottom*.

S. lycopersicum PLD´ (Solyc02g083340.2.1) construct was then subjected to BLAST in the Sol genomics dataset *N. benthamiana* proteomics dataset [from Michelle L. Cilia]. The BLAST resulted in several possible *N. benthamiana* PLD´ constructs. The highest homology to the *S. lycopersicum* PLD´ sequence was NbS00010125g0016.1 (890aa). Because there were several possible sequences, the validity of NbS00010125g0016.1 was tested. Although the genome crawl was designed to extract the PLD´ associated with *N. tabacum,* it was important to maintain integrity of the SBIP-436 Y2H clone. SGN-U444527 was subjected to BLAST in the

Sol Genomics *N. benthamiana* Genome v0.4.4 predicted proteins dataset to see if NbS00010125g0016.1 retained highest homology. While NbS00010125g0016.1 was on the list of homologous constructs with an identity score of 73%, the results revealed a predicted protein construct with an identity score of 97% deemed NbS00023265g0007.1 (852aa). In the same fashion the NbS00023265g0007.1 was subjected to BLAST in the Sol Genomics *S. lycopersicum* dataset ITAG Release 1 predicted proteins (SL1.00). A highly similar sequence was found named SL1.00sc02164_456.1.1 (848aa) shown in Figure 15 followed by a clustal alignment in Figure 16.

NbS00023265g0007.1

MADENCENVIHLHGDLDLKILEARRLPNMDLVTERLRRCFTALDVCRKPFTRRRRKGHHRKIITSDPYVTVCLSGATVARTRVISNCQDPVWNEHFK
IPLAHPVSVVEFQVKDNDVFGADYIGVATVPAQKIKSGELIDDWFPIIGPYGKPPKRDCAIRLQMKFTHCNGNPLYNSAISEDYGLKESYFPVRHGG
SVTLYQDAHVPDGMLPEIKLDDNKVFQHSKCWEDICHAILEAHHLVYIVGWSIFHKVKLVREPSKPLPSGGDLTLGELLKYKSEEGVRVLLLVWDDK
TSHSKFFIQTDGLMCTHDEETRKFFKHSSVTCVLSPRYASSKLSIFKQQALLFSCQEKIQLVVGTLYTHHQKCVIVDTQASGNNRKVTAFLGGLDLC
DGRYDTPEHRLFSDLDTVFKDDYHNPTFCTGTKGPRQPWHDLHCKVEGPAAYDVLTNFEQRWRKATKWSELGRRLKRISHWHDDALIKIERISWIIS
PSSSVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVFLAESQNLVCAKNLVIDKSIQMGYIQAIRQAQHFIYIENQYFLGSSYAWHSYKD
AGADNLIPMELALKIASKIREKERFSVYVVIPMWPEGVPTSASVQEILYWQRQTMKMMYGIIAQELKSSQLKDVHPSDYLNFYCLGNREELHGESKT
NPASSNGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWAKKKKHPHGQVYGYRMSLWAEHMGKLDDCFT
KPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPDALTT

SL1.00sc02164_456.1.1

MADENCENVVYLHGDFDLKIIEARRLPNMDLVTERLSRCFTALDICRKPFTRRRRKGHRRKIITSDPYVTVCLTGATVARTRVISNCQDPVWNEHFK
IPLAHPVSVVEFLVKDNDVFGADYIGVATVLAEKIKSGELIDDWFPIIGPYGKPPKPDCAIRLQMRFIHCDGNPSYNGGISEDFGLKASYFPVRHGG
SVTLYQDAHVPDGMLPEIKLDDDKIFEHSKCWEDICHAILEAHHLVYVVGWSIFHKVKLVREPSKPLPSGGDLTLGELLKYKSEEGVRVLLLVWDDK
TSHSKFFIQTDGVMQTHDEETRKFFKHSSVNCVLAPRYASSKLSIFKQQACFTPYQYFVVGTLYTHHQKCVIVDTQASGNNRKVSAFLGGLDLCDGR
YDTPEHRLFRDLDTVFKDDFHNPTFSTGTKAPRQPWHDLHCKIEGPAAYDVLTNFEQRWRKATKWSEFGRRLKKISHWHDDALIKIERISWITSPSS
SVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVLLAESQNLVCAKNLVIDRSIQMAYIQAIRQAQHFIYIENQYFLGSSYAWPSYKEAGA
DNLIPMELALKIASKIRAKERFAVYIVIPMWPEGVPTSASVQEILYWQRQTMKMMYGIIAQELKSSQLQDVHLSDYLNFYCLGNREELHGESKSNYA
SNGDLISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWAMKKRHPHGQVYGYRMSLWAEHMGKLDDIFTKPES
LNCVKHVNEVAEDNWKRFTAHEFKPLQGHLLKYPVQVGTDGQVSSLPGHEYFPDVGGKILGARTNLPDALTT

Figure 15: SGN-U444527 used to BLAST in Sol Genomics *N.benthamiana* Genome v0.4.4 predicted proteins revealed NbS00023265g0007.1, which was then used for BLAST in *S.lycopersicum* dataset ITAG Release 1 predicted proteins (SL1.00) dataset resulting in SL1.00sc02164_456.1.1. Grey highlight represents homology to SGN-U444527 and subsequently the SBIP-436 Yeast-2-Hybrid clone.

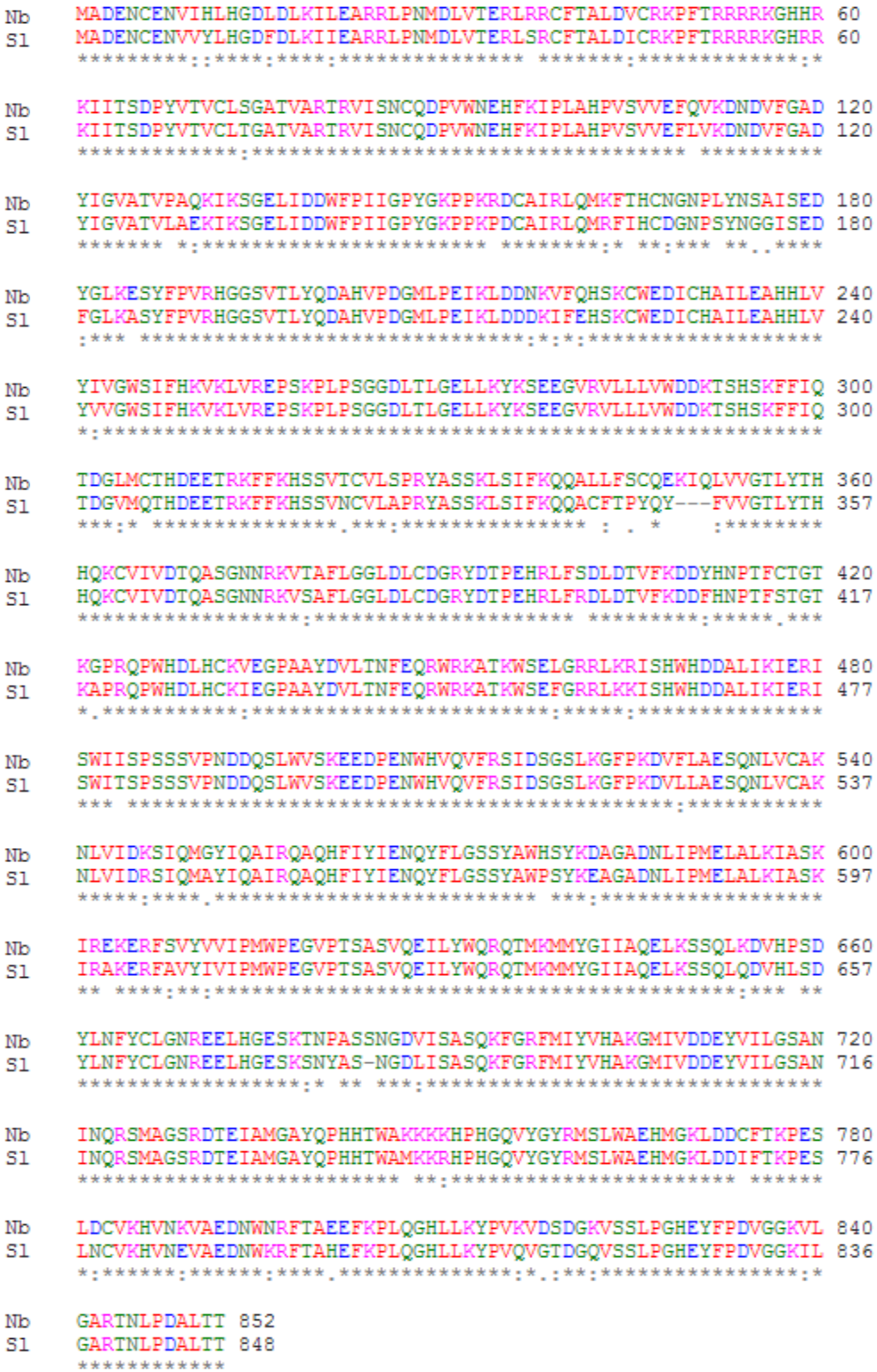


Figure 16: CLUSTAL 2.1 multiple sequence alignment of *N. benthamiana* NbS00023265g0007.1 (852aa) and *S. lycopersicum* SL1.00sc02164_456.1.1 (848aa).

The NbS00023265g0007.1, SL1.00sc02164_456.1.1, SGN-U444527, and the SBIP-436 Y2H clone sequences were aligned via Clustal W2, shown in Figure 17, to verify that the original SBIP436 Y2H clone integrity had not been compromised.

```
Nb              YLNFYCLGNREELHGESKTNPASSNGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSAN 720
U444527         --NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSAN 58
SBIP436         ---------------------------SASQKFGRFMIYVHAKGMIVDDEYVILGSAN 31
Sl              YLNFYCLGNREELHGESKSNYAS-NGDLISASQKFGRFMIYVHAKGMIVDDEYVILGSAN 716
                                           ****************************

Nb              INQRSMAGSRDTEIAMGAYQPHHTWAKKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPES 780
U444527         INQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPES 118
SBIP436         INQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQV----------------------- 67
Sl              INQRSMAGSRDTEIAMGAYQPHHTWAMKKRHPHGQVYGYRMSLWAEHMGKLDDIFTKPES 776
                ************************* **:******

Nb              LDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVL 840
U444527         LDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVL 178
SBIP436         ------------------------------------------------------------
Sl              LNCVKHVNEVAEDNWKRFTAHEFKPLQGHLLKYPVQVGTDGQVSSLPGHEYFPDVGGKIL 836


Nb              GARTNLPDALTT 852
U444527         GARTNLPDALTT 190
SBIP436         ------------
Sl              GARTNLPDALTT 848
```

Figure 17: CLUSTAL 2.1 multiple sequence alignment of *N. benthamiana* NbS00023265g0007.1 (852aa), *S. lycopersicum* SL1.00sc02164_456.1.1(848aa), *N. tabacum* SGN-U444527 (190aa), SBIP-436 Yeast-2-Hybrid Clone (67aa).

NbS00023265g0007.1 sequence was used to BLAST in *N. tabacum* Unigenes dataset without any significant results. NbS00023265g0007.1 was then subjected to blast in the Sol Genomics dataset "*N. tabacum* Methylation Filtered Genome TGI:v.1 Contigs". The results from this BLAST search showed homology in separate sections of NbS00023265g0007.1 gene construct. The sections that resembled portions of the NbS00023265g0007.1 construct were pieced together as overlapping segments to form a possible *N. tabacum* PLD´ gene construct. A *N.tabacum* PLD´ partial was also used as a portion of the construct due to its homology to the SGN-U444527 and the SBIP-436 Y2H Clone. Analyzing from 5' to 3', the first section is

sequence c1562 covers amino acids 1-302. C1562 sequence and clustal alignment with Nb sequence is shown in Figure 18.

c1562

```
MADENCENVIHLHGDLDLKILEARRLPNMDLVTERLRRCFTALDVCRKPFTRRRRKGHHRKIITSDPYVT
VCLSGATVARTRVISNCQDPVWNEHFKIPLAHPVSVVEFQVKDNDVFGADYIGVATVPAQKIKSGELIDD
WFPIIGPYGKPPKPDCAIRLQMKFTHCNGNPVYNSGISEDYGLKESYFPVRHGGSVTLYQDAHVPDGMLP
EIKLDDNKVFEHSKCWEDICHAILEAHHLVYIVGWSIFHKVKLVREPSKPLPSGGDLTLGDLLKYKSEEG
VRVLLLVWDDKTSHSKFFIQTV
```

```
Nb              MADENCENVIHLHGDLDLKILEARRLPNMDLVTERLRRCFTALDVCRKPFTRRRRKGHHR 60
c1562           MADENCENVIHLHGDLDLKILEARRLPNMDLVTERLRRCFTALDVCRKPFTRRRRKGHHR 60
                ************************************************************

Nb              KIITSDPYVTVCLSGATVARTRVISNCQDPVWNEHFKIPLAHPVSVVEFQVKDNDVFGAD 120
c1562           KIITSDPYVTVCLSGATVARTRVISNCQDPVWNEHFKIPLAHPVSVVEFQVKDNDVFGAD 120
                ************************************************************

Nb              YIGVATVPAQKIKSGELIDDWFPIIGPYGKPPKRDCAIRLQMKFTHCNGNPLYNSAISED 180
c1562           YIGVATVPAQKIKSGELIDDWFPIIGPYGKPPKPDCAIRLQMKFTHCNGNPVYNSGISED 180
                ********************************* ******************:***.**** 

Nb              YGLKESYFPVRHGGSVTLYQDAHVPDGMLPEIKLDDNKVFQHSKCWEDICHAILEAHHLV 240
c1562           YGLKESYFPVRHGGSVTLYQDAHVPDGMLPEIKLDDNKVFEHSKCWEDICHAILEAHHLV 240
                ****************************************:******************* 

Nb              YIVGWSIFHKVKLVREPSKPLPSGGDLTLGELLKYKSEEGVRVLLLVWDDKTSHSKFFIQ 300
c1562           YIVGWSIFHKVKLVREPSKPLPSGGDLTLGDLLKYKSEEGVRVLLLVWDDKTSHSKFFIQ 300
                ******************************:***************************** 

Nb              TDGLMCTHDEETRKFFKHSSVTCVLSPRYASSKLSIFKQQALLFSCQEKIQLVVGTLYTH 360
c1562           TV---------------------------------------------------------- 302
                *
```

Figure 18: CLUSTAL 2.1 multiple sequence alignment of *N. benthamiana* NbS00023265g0007.1 (852aa) and *N. tabacum* processed tobacco genome sequences c1562 ORF (302aa).

The second section is a partial sequence taken from the NCBI database. The GQ904710.1 (381aa) spans amino acids 358-738 of the NbS00023265g0007.1 gene construct shown in Figure 19.

GQ904710.1

```
FTHHQKCVIVDSQASGNNRKITAFLGGLDLCDGRYDTPEHRLFRDLDTVFKDDYHNPTFGAGTKGPRQPWHDLHCKVEGPAAYDVLTNFEQRWRKAT
KWSELGRRLKKISHWHDDALIKIERISWIISPSSSVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVFLAESQNLVCAKNLVIDKSIQMG
```

```
YIQAIRQAQHFIYIENQYFLGSSYAWHSYKDAGADNLIPMELALKIASKIREKERFSVYVVIPMWPEGVPTSASVQEILYWQRQTMKMMYGIIAQEL
KSSQLKDVHPSDYLNFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGA

Nb              TDGLMCTHDEETRKFFKHSSVTCVLSPRYASSKLSIFKQQALLFSCQEKIQLVVGTLYTH 360
GQ904710.1      ------------------------------------------------------------FTH 3
                                                                             :**

Nb              HQKCVIVDTQASGNNRKVTAFLGGLDLCDGRYDTPEHRLFSDLDTVFKDDYHNPTFCTGT 420
GQ904710.1      HQKCVIVDSQASGNNRKITAFLGGLDLCDGRYDTPEHRLFRDLDTVFKDDYHNPTFGAGT 63
                ********:********:********************** *************** :**

Nb              KGPRQPWHDLHCKVEGPAAYDVLTNFEQRWRKATKWSELGRRLKRISHWHDDALIKIERI 480
GQ904710.1      KGPRQPWHDLHCKVEGPAAYDVLTNFEQRWRKATKWSELGRRLKKISHWHDDALIKIERI 123
                ******************************************:***************

Nb              SWIISPSSSVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVFLAESQNLVCAK 540
GQ904710.1      SWIISPSSSVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVFLAESQNLVCAK 183
                ************************************************************

Nb              NLVIDKSIQMGYIQAIRQAQHFIYIENQYFLGSSYAWHSYKDAGADNLIPMELALKIASK 600
GQ904710.1      NLVIDKSIQMGYIQAIRQAQHFIYIENQYFLGSSYAWHSYKDAGADNLIPMELALKIASK 243
                ************************************************************

Nb              IREKERFSVYVVIPMWPEGVPTSASVQEILYWQRQTMKMMYGIIAQELKSSQLKDVHPSD 660
GQ904710.1      IREKERFSVYVVIPMWPEGVPTSASVQEILYWQRQTMKMMYGIIAQELKSSQLKDVHPSD 303
                ************************************************************

Nb              YLNFYCLGNREELHGESKTNPASSNGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSAN 720
GQ904710.1      YLNFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSAN 363
                *************** **.*****.************************************

Nb              INQRSMAGSRDTEIAMGAYQPHHTWAKKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPES 780
GQ904710.1      INQRSMAGSRDTEIAMGA------------------------------------------ 381
                ******************

Nb              LDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVL 840
GQ904710.1      ------------------------------------------------------------

Nb              GARTNLPDALTT 852
GQ904710.1      ------------
```

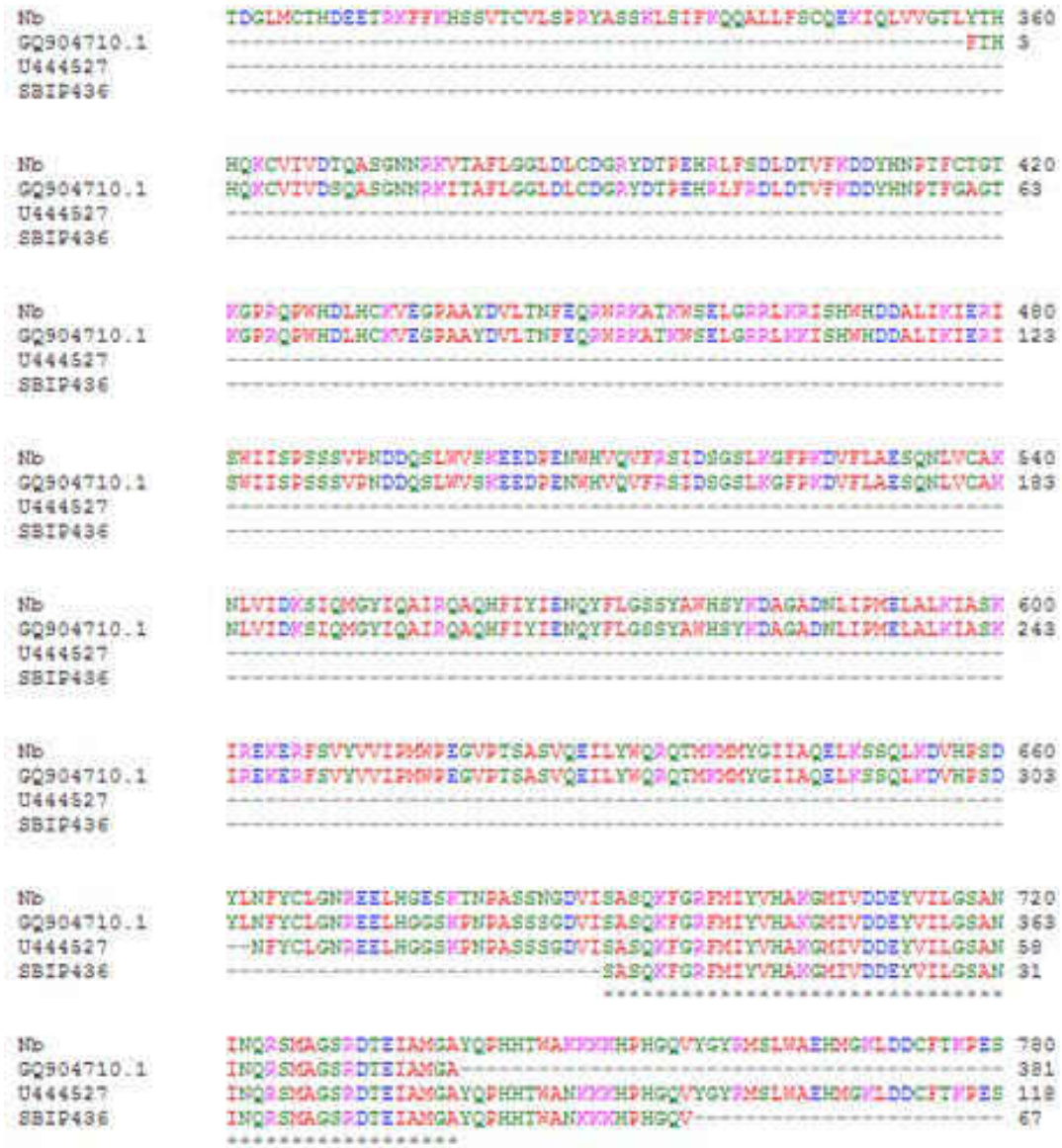
Figure 19: CLUSTAL 2.1 multiple sequence alignment of *N. benthamiana* NbS00023265g0007.1 (852aa) and *N. tabacum* PLD´ partial sequence GQ904710.1 (381aa).

The third section is the original SGN-U444527 that was found using the SBIP-436 Y2H Clone. This section can also be found from the processed tobacco genome sequences c6690 in the Sol Genomics dataset "*N. tabacum* Methylation Filtered Genome TGI:v.1 Contigs". This

section also overlaps with the preceding GQ904710.1 section. The SGN-U444527 sequence covers amino acids 663-852 of the NbS00023265g0007.1 gene construct shown in Figure 20.

SGN-U444527 (c6690)

NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGY
RMSLWAEHMGKLDDCFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPDALTT



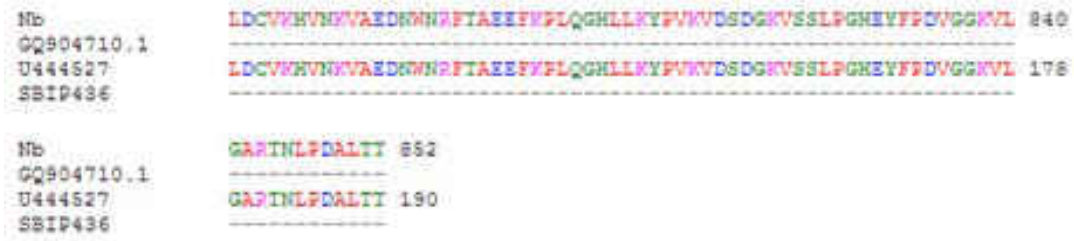


Figure 20: CLUSTAL 2.1 multiple sequence alignment of *N. benthamiana* NbS00023265g0007.1 (852aa), GQ904710.1 (381aa), *N. tabacum* SGN-U444527 (190aa), SBIP-436 Yeast Two-Hybrid Clone (67aa).

The 3 sequences were trimmed of overlapping portions, consolidated, and aligned to the *N.benthamiana* NbS00023265g0007.1. This is demonstrated in Figure 21. There is a missing section between amino acid 303-356. There were sequences found that demonstrated some homology to this segment, but the segments were represented with weak homology. Therefore, the sequences did not overlap the surrounding sequences well enough to be confident in their validity. Regions of the *N. tabacum* PLD´ chosen for amplification are shown in Figure 22.

N. tabacum PLD´ (797aa) construct (putative)

MADENCENVIHLHGDLDLKILEARRLPNMDLVTERLRRCFTALDVCRKPFTRRRRKGHHRKIITSDPYVTVCLSGATVARTRVISNCQDPVWNEHFK
IPLAHPVSVVEFQVKDNDVFGADYIGVATVPAQKIKSGELIDDWFPIIGPYGKPPKPDCAIRLQMKFTHCNGNPVYNSGISEDYGLKESYFPVRHGG
SVTLYQDAHVPDGMLPEIKLDDNKVFEHSKCWEDICHAILEAHHLVYIVGWSIFHKVKLVREPSKPLPSGGDLTLGDLLKYKSEEGVRVLLLVWDDK
TSHSKFFIQTVFTHHQKCVIVDSQASGNNRKITAFLGGLDLCDGRYDTPEHRLFRDLDTVFKDDYHNPTFGAGTKGPRQPWHDLHCKVEGPAAYDVL
TNFEQRWRKATKWSELGRRLKKISHWHDDALIKIERISWIISPSSSVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVFLAESQNLVCAK
NLVIDKSIQMGYIQAIRQAQHFIYIENQYFLGSSYAWHSYKDAGADNLIPMELALKIASKIREKERFSVYVVIPMWPEGVPTSASVQEILYWQRQTM
KMMYGIIAQELKSSQLKDVHPSDYLNFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEI
AMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEY
FPDVGGKVLGARTNLPDALTT

```
Nb              MADENCENVIHLHGDLDLKILEARRLPNMDLVTERLRRCFTALDVCRKPFTRRRRKGHHR 60
NtPLDdelta      MADENCENVIHLHGDLDLKILEARRLPNMDLVTERLRRCFTALDVCRKPFTRRRRKGHHR 60
SBIP436         ------------------------------------------------------------

Nb              KIITSDPYVTVCLSGATVARTRVISNCQDPVWNEHFKIPLAHPVSVVEFQVKDNDVFGAD 120
NtPLDdelta      KIITSDPYVTVCLSGATVARTRVISNCQDPVWNEHFKIPLAHPVSVVEFQVKDNDVFGAD 120
SBIP436         ------------------------------------------------------------

Nb              YIGVATVPAQKIKSGELIDDWFPIIGPYGKPPKRDCAIRLQMKFTHCNGNPLYNSAISED 180
NtPLDdelta      YIGVATVPAQKIKSGELIDDWFPIIGPYGKPPKPDCAIRLQMKFTHCNGNPVYNSGISED 180
SBIP436         ------------------------------------------------------------

Nb              YGLKESYFPVRHGGSVTLYQDAHVPDGMLPEIKLDDNKVFQHSKCWEDICHAILEAHHLV 240
NtPLDdelta      YGLKESYFPVRHGGSVTLYQDAHVPDGMLPEIKLDDNKVFEHSKCWEDICHAILEAHHLV 240
SBIP436         ------------------------------------------------------------

Nb              YIVGWSIFHKVKLVREPSKPLPSGGDLTLGELLKYKSEEGVRVLLLVWDDKTSHSKFFIQ 300
NtPLDdelta      YIVGWSIFHKVKLVREPSKPLPSGGDLTLGDLLKYKSEEGVRVLLLVWDDKTSHSKFFIQ 300
SBIP436         ------------------------------------------------------------

Nb              TDGLMCTHDEETRKFFKHSSVTCVLSPRYASSKLSIFKQQALLFSCQEKIQLVVGTLYTH 360
NtPLDdelta      TV------------------------------------------------------FTH 305
SBIP436         ------------------------------------------------------------

Nb              HQKCVIVDTQASGNNRKVTAFLGGLDLCDGRYDTPEHRLFSDLDTVFKDDYHNPTFCTGT 420
NtPLDdelta      HQKCVIVDSQASGNNRKITAFLGGLDLCDGRYDTPEHRLFRDLDTVFKDDYHNPTFGAGT 365
SBIP436         ------------------------------------------------------------

Nb              KGPRQPWHDLHCKVEGPAAYDVLTNFEQRWRKATKWSELGRRLKRISHWHDDALIKIERI 480
NtPLDdelta      KGPRQPWHDLHCKVEGPAAYDVLTNFEQRWRKATKWSELGRRLKKISHWHDDALIKIERI 425
SBIP436         ------------------------------------------------------------

Nb              SWIISPSSSVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVFLAESQNLVCAK 540
NtPLDdelta      SWIISPSSSVPNDDQSLWVSKEEDPENWHVQVFRSIDSGSLKGFPKDVFLAESQNLVCAK 485
SBIP436         ------------------------------------------------------------

Nb              NLVIDKSIQMGYIQAIRQAQHFIYIENQYFLGSSYAWHSYKDAGADNLIPMELALKIASK 600
NtPLDdelta      NLVIDKSIQMGYIQAIRQAQHFIYIENQYFLGSSYAWHSYKDAGADNLIPMELALKIASK 545
SBIP436         ------------------------------------------------------------

Nb              IREKERFSVYVVIPMWPEGVPTSASVQEILYWQRQTMKMMYGIIAQELKSSQLKDVHPSD 660
NtPLDdelta      IREKERFSVYVVIPMWPEGVPTSASVQEILYWQRQTMKMMYGIIAQELKSSQLKDVHPSD 605
SBIP436         ------------------------------------------------------------
```

Figure 21: CLUSTAL 2.1 multiple sequence alignment of *N.benthamiana* NbS00023265g0007.1 (852aa), *N. tabacum* PLD´ (797aa) construct (putative), SBIP-436 Yeast-2-Hybrid Clone (67aa).

SBIP-436 Amplification Regions

Figure 22: SBIP-436 Amplification Regions. 1. In Purple is SBIP-436 1kb-A (602-603), 2. In Blue is SBIP-436 1kb-B (604-605), 3. In Green is SBIP-436-210 (DK511-DK513), 4. Highlighted in Green and Yellow is SBIP-436-500/700 (DK511-DK514) which is a continuation of the SBIP-436 -A. Grey highlights indicate overlapping portions.

Primers to amplify full length SBIP-436 were synthesized. These are shown in Figure 23 SBIP-436 Full-1 (DK547-DK548) were constructed as gateway cloning primers containing the appropriate attb1 and attb2 sites for the forward and reverse primers respectively. SBIP-436 Full-2 (DK602-DK616) were constructed using only the SBIP-436 full gene sequence, lacking the gateway attb sites.

ATGGCGGATGAGAATTGTGAAAATGTGATACATCTACATGGAGACCTTGATTTGAAGATTCTAGAAGCTCGACGATTACC
TAATATGGATTTGGTTACCGAACGTCTACGTCGTTGCTTTACAGCATTAGATGTTTGCCGGAAACCTTTCACACGCCGCC
GCCGGAAAGGTCACCACCGGAAAATCATAACTAGCGATCCGTATGTTACTGTTTGCTTATCCGGCGCTACTGTGGCGCGT
ACCCGCGTGATTTCGAACTGTCAGGATCCTGTTTGGAACGAGCATTTCAAAATCCCGCTAGCTCATCCTGTTTCCGTTGT
TGAATTCCAGGTTAAGGATAACGACGTTTTCGGCGCGGATTACATCGGCGTTGCTACCGTTCCAGCTCAGAAGATCAAGT
CCGGCGAGCTCATTGATGACTGGTTCCCTATAATTGGACCTTACGGTAAACCTCCAAAGCGTGATTGTGCTATTAGACTC
CAAATGAAATTCACACATTGCAACGGTAATCCGTTATATAACAGCGCCATATCAGAGGATTACGGCTTAAAAGAGAGTTA

TTTTCCGGTGAGGCACGGAGGATCGGTTACTTTATATCAGGACGCTCACGTGCCAGATGGAATGTTGCCGGAGATTAAAT
TGGATGATAACAAGGTGTTCCAGCATAGTAAGTGTTGGGAAGATATATGCCATGCGATATTGGAGGCACATCATTTGGTG
TATATAGTTGGGTGGTCAATATTTCATAAGGTGAAGCTGGTTAGGGAGCCAAGTAAGCCGTTGCCGAGCGGTGGGGATTT
GACACTTGGGGAGTTGCTTAAGTATAAATCGGAGGAAGGAGTGAGGGTATTGTTGTTGGTTTGGGATGATAAGACTTCTC
ACAGCAAATTCTTCATTCAAACGGATGGACTGATGTGTACTCATGATGAAGAAACTCGGAAATTTTTCAAGCACTCATCC
GTCACTTGTGTGCTGTCACCTCGTTATGCAAGCAGTAAGCTAAGCATTTTCAAGCAACAGGCATTGCTTTTCTCCTGCCA
AGAGAAAATTCAACTGGTGGTTGGAACCCTTTATACGCACCATCAGAAGTGCGTGATTGTGGACACACAAGCCAGTGGCA
ACAATCGGAAGGTCACAGCTTTCCTTGGTGGTCTAGACCTCTGTGATGGGCGTTATGATACACCTGAGCATAGATTATTC
TCTGATCTTGACACAGTCTTCAAGGATGATTATCATAATCCAACATTTTGTACAGGAACCAAGGGACCTAGACAGCCATG
GCATGACTTGCATTGCAAGGTTGAAGGACCTGCTGCATATGATGTGCTCACAAACTTTGAGCAGAGGTGGAGAAAAGCCA
CAAAATGGTCAGAGTTGGGAAGACGTTTAAAAAGGATATCTCATTGGCATGACGATGCTTTGATCAAAATAGAAAGAATT
TCTTGGATTATTAGTCCTTCCTCTTCTGTTCCGAATGATGACCAATCTTTGTGGGTTTCCAAGGAAGAAGATCCTGAAAA
CTGGCATGTTCAGGTATTCCGATCTATTGATTCAGGGTCTTTGAAAGGATTTCCTAAAGATGTTTTTTTGGCTGAATCAC
AGAACCTTGTCTGTGCAAAAAATTTGGTGATCGATAAGAGCATCCAAATGGGATATATTCAGGCAATAAGACAGGCACAA
CATTTTATCTATATTGAGAATCAATATTTCCTTGGGTCATCATATGCTTGGCATTCATACAAAGATGCAGGTGCTGATAA
TCTAATTCCCATGGAGCTTGCATTAAAGATAGCCAGTAAAATTCGGGAAAAAGAGCGATTTTCTGTTTATGTTGTCATTC
CAATGTGGCCTGAGGGAGTCCCCACTTCTGCTTCAGTGCAAGAAATTCTGTATTGGCAGAGGCAAACCATGAAAATGATG
TATGGAATCATTGCTCAAGAGCTAAAATCTTCTCAACTTAAGGATGTACATCCTTCTGACTATCTAAACTTCTATTGTCT
GGGTAATCGAGAAGAATTACATGGAGAATCAAAGACAAACCCTGCTTCCTCAAATGGTGATGTGATCTCAGCTTCCCAGA
AATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTTATTTTAGGATCCGCTAAT
ATTAACCAAAGGTCTATGGCTGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTCATCACACTTGGGCTAA
CAAGAAAAAACATCCACATGGCCAGGTATATGGTTATAGAATGTCTCTGTGGGCAGAGCATATGGGCAAGTTAGACGATT
GCTTCACAAAGCCAGAAAGTTTGGACTGTGTGAAGCATGTGAATAAGGTTGCTGAAGATAATTGGAACAGATTCACTGCT
GAGGAGTTCAAACCTTTACAAGGTCATCTTCTCAAGTACCCAGTCAAAGTAGATTCTGATGGGAAAGTAAGTTCCTTACC
TGGTCATGAATATTTTCCTGATGTTGGTGGTAAAGTACTAGGAGCTCGAACCAACCTTCCTGATGCTTTGACCACATGAC
CACATGACTCTTCCAGC

Figure 23: SBIP-436 Full Gene Primers. Highlighted in green is the sequence portion of the DK547 (attb1 site is added- not shown) forward primer and the DK602 (Lacking attb site). Highlighted in Yellow plus the grey portion represents the DK548 (attb2 site is added-not shown) reverse primer. Highlighted in Blue represents plus the grey portion represents the DK616 (Lacking attb site) reverse primer.

PCR Amplification of SBIP-436 Expression Segments

The first SBIP-436 expression primer set, SBIP-436 -A (DK511-DK513), was designed to amplify a product of 210bp from 3' end. PCR conditions were optimized in regards to gene length and primer annealing temperatures. Optimal annealing temperature was set at 60^{o}C and the optimal extension time was set at 45 seconds. For best amplification for visualization of the 210bp segment 32 PCR cycles was used. PCR product was then analyzed by agarose gel electrophoresis on a 1.2% gel, visualized under UV, and photographed. The PCR product can be viewed in Figure 24.

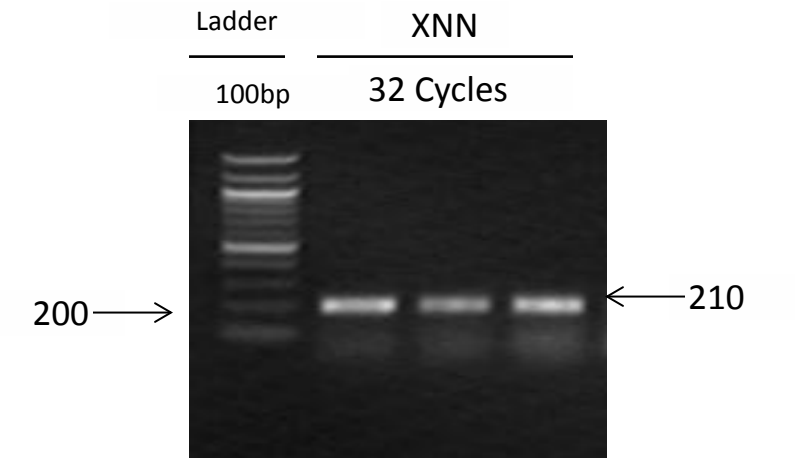


Figure 24: PCR amplification of SBIP-436-A. 1.2% agarose gel showing ~210bp fragment. 100bp ladder is used as a size marker and the 200bp band is labeled as a reference.

The second SBIP436 expression primer set SBIP-436-B1 (DK511-DK514) was designed to amplify a product of 490bp. PCR conditions were optimized in regards to gene length and primer annealing temperatures. Optimal annealing temperature was set at 60^{o}C and the optimal extension time was set at 45 seconds. For best amplification for visualization of the 490bp segment 32 PCR cycles was used. PCR product was then subjected to agarose gel electrophoresis as described above. Agarose gel electrophoresis revealed a second unexpected amplicon of approximately 700bp. The PCR product can be viewed in Figure 25.

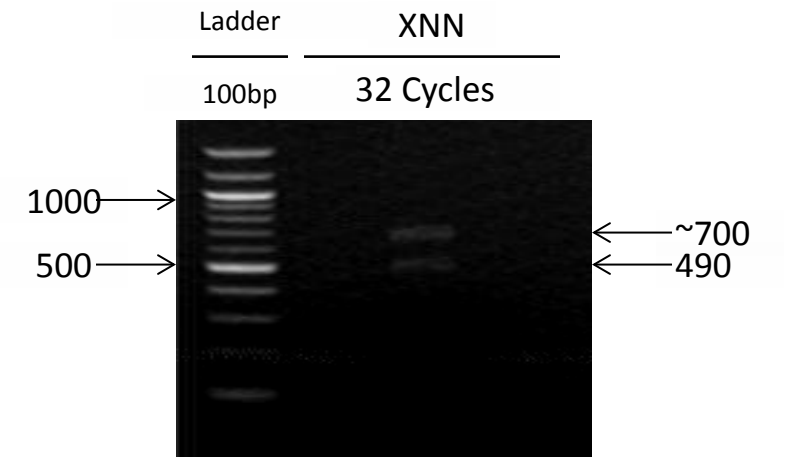


Figure 25: PCR amplification of SBIP-436 -B1 (DK511-DK514). 1.2% agarose gel showing amplification (32 cycles) of a 490bp and a 700bp. PCR was run for 32 cycles. The 500 and 1000bp bands are labeled as a reference.

The 490bp amplicon from SBIP-436 -B1 contains the 210bp amplicon from SBIP-436-A because both primer sets share the same forward primer (DK511). The unexpected ~700bp amplicon resulting from the SBIP-436 -B1 set suggested that the DK514 reverse primer is either nonspecific or is amplifying a gene segment representing either, a) SBIP-436 isoform or, b) existence of a alternatively spliced transcript of SBIP-436. Because the SBIP-436 -A did not yield an unexpected band it was ruled unlikely to be nonspecific because of the common DK511 forward primer. In addition, the PLD family is known to maintain multiple isoforms, highly conserved domains, and implications of alternative splicing. It was found prudent to reveal the identity of the unexpected ~700bp amplicon and establish the relationship of these three segments to each other. This was achieved by cloning and sequencing the amplicons.

Cloning the segments yielded from SBIP-436 -B1 PCR called for separation of the 2 amplified products. To achieve this separation a QIAquick Gel Extraction Kit was used. The gel extraction kit allowed us to cut the 2 products from the gel to separate them. PCR were separated

on a 1.2% agarose gel. Following, electrophoresis the gel was visualized and the amplicons excised using a scalpel. The PCR product can be viewed in Figure 26.

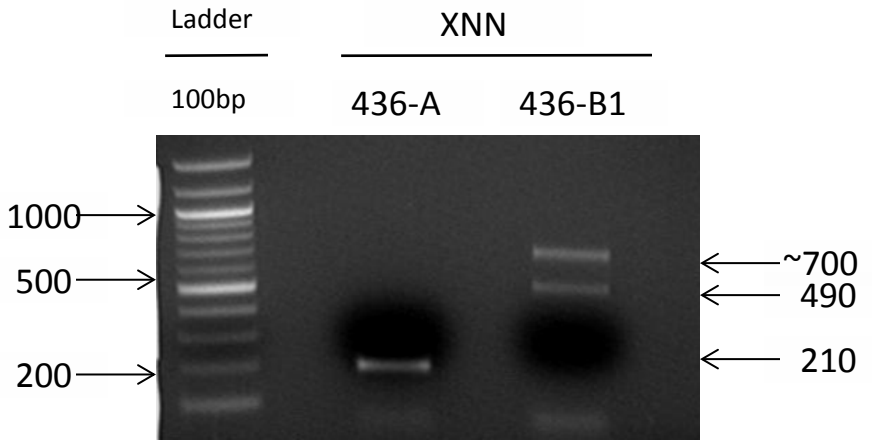


Figure 26: Gel Extraction-PCR amplification of SBIP-436 -A (DK511-DK513) and SBIP-436 -B1 (DK511-DK514). 1.2% Agarose gel showing expression of the 210bp segment of SBIP-436, expression of the 490bp segment of SBIP-436, and the unexpected band located in the approximate 700bp range. PCR was run for 32 cycles.

The gel extraction method failed to yield sufficient amounts of amplified PCR product to be useful for cloning. A direct PCR product purification method was used to increase the amount of purified PCR product. This process does not individually separate the 490 and 700 bp amplified fragment. It was expected that both amplified fragments will get cloned separately and individual clones containing either the 490bp or ~700bp amplicon could be identified. Plasmid DNA from individual clones were purified and screened using agarose gel electrophoresis shown in Figures 27, 28, and 29.

SBIP-436 -A (210bp) pGEMT Vector Plasmid DNA

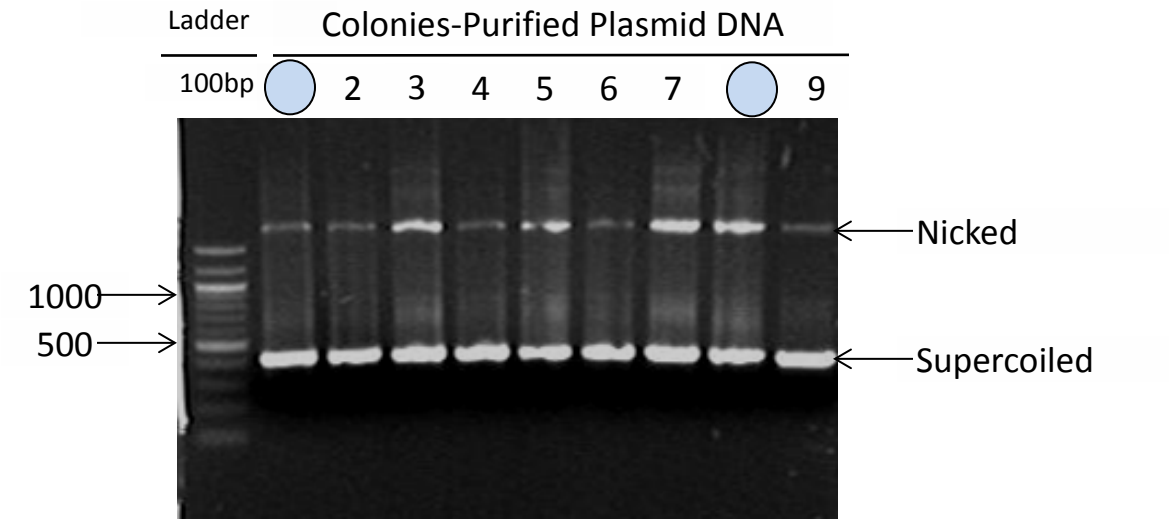


Figure 27: Plasmid DNA of SBIP-436 -A in pGEMT Vector. 1.0% Agarose gel showing viability of the plasmid DNA. Each numbered well represents a white colony that has been selected and purified. 100bp ladder is utilized on the left hand side to be used as a marker. The 500bp and 1000bp band is labeled as a reference. 2µl of purified plasmid DNA was loaded in each well. Circled colony number represents colony plasmid sample chosen for sequencing.

SBIP-436 -B1 (490bp and ~700bp) pGEMT Vector Plasmid DNA

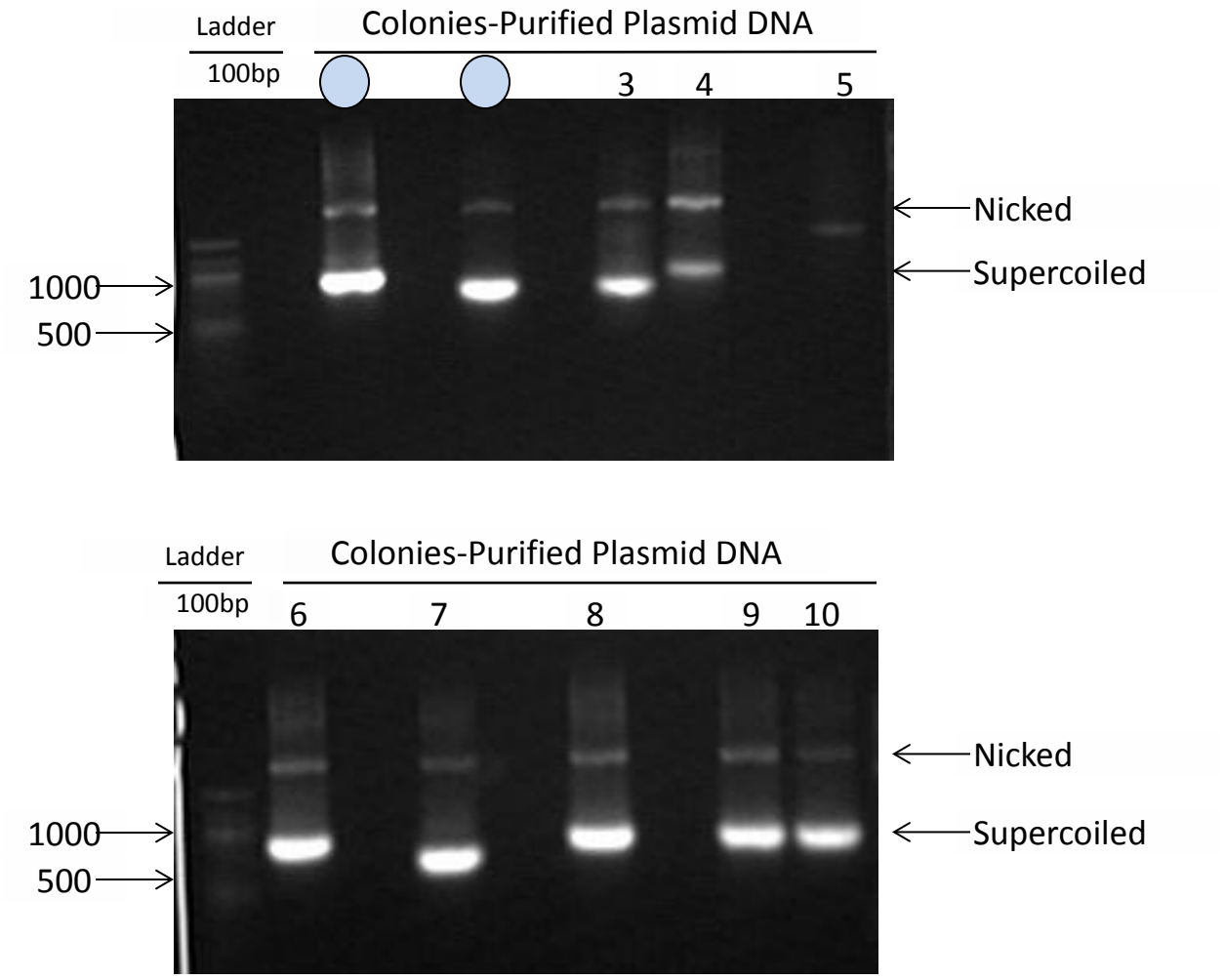


Figure 28: Plasmid DNA of SBIP-436 -B1 in pGEMT Vector. 1.0% Agarose gel showing viability of the plasmid DNA. Each numbered well represents a white colony that has been selected and purified. Circled colony number represents colony plasmid sample chosen for sequencing.

SBIP-436 -B1 (490bp and ~700bp) TOPO Vector Plasmid DNA

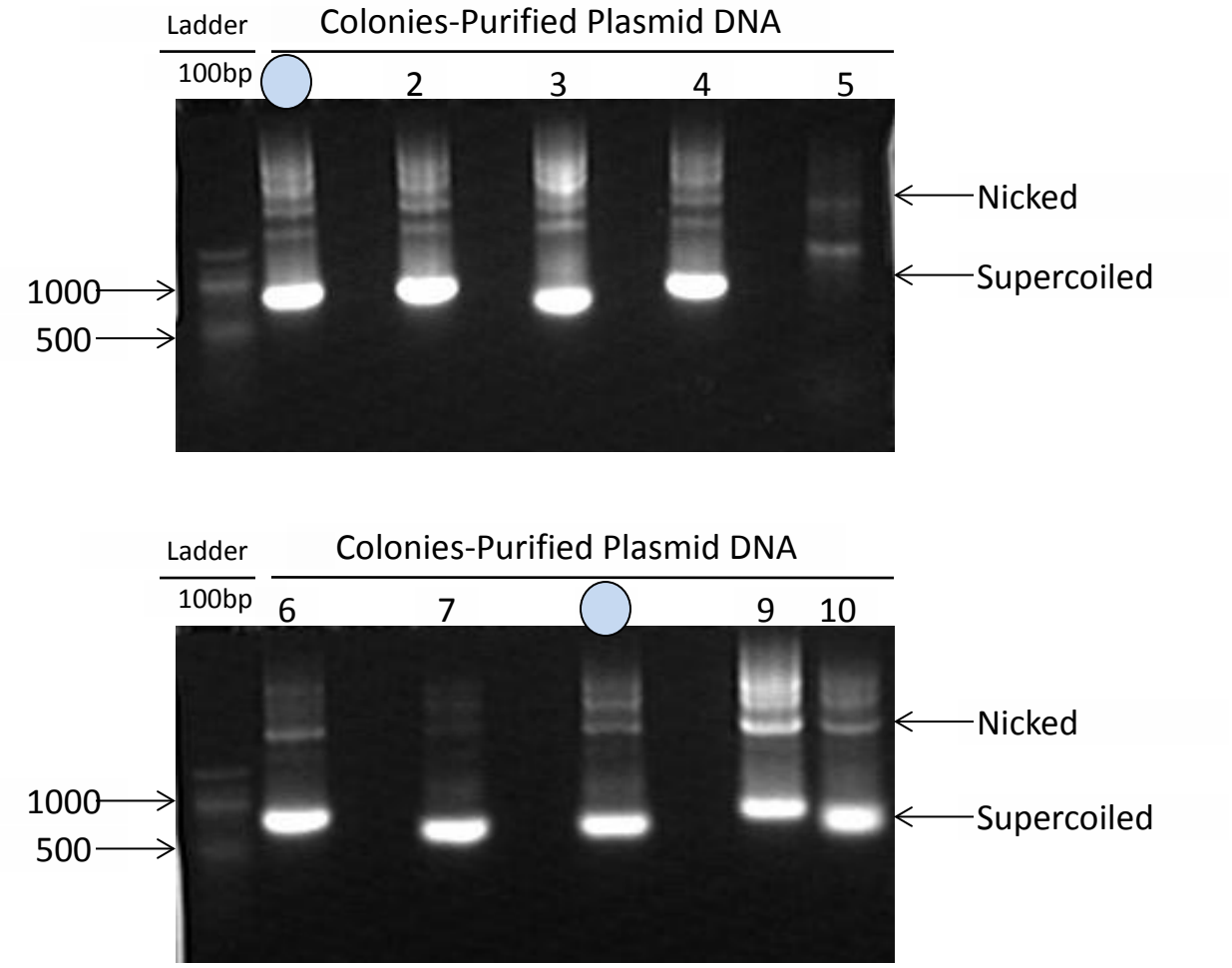


Figure 29: Plasmid DNA of SBIP-436 -B1 in TOPO Vector. 1.0% Agarose gel showing viability of the plasmid DNA. Each numbered well represents a white colony that has been selected and purified. Circled colony number represents colony plasmid sample chosen for sequencing.

Clones 1 and 2 from SBIP-436 -A in pGEMT vector were selected for sequencing to represent the 210bp segment. Clones 1 and 2 from SBIP-436 -B1 in pGEMT Vector as well as clones 1 and 8 from SBIP-436 -B1 in TOPO Vector were selected to represent the 490bp/~700bp segments. The purified plasmid DNA was sequenced using M13 forward and reverse primers The sequenced clones are shown in Figure 30.

SBIP-436 -A (DK511-513)- Sequence Clone 1 pGEMT Sample 1– 210bp

Nucleotide sequence

GATTTCAGCTTCCCAGAAATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTAATTTTAGGATCCGCTAAT
ATTAACCAAAGATCTATGGCGGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTTATCACACTTGGGCCAAGAAGAAAAAACATCCAC
ATGGCCAGGTATAA

Translated amino acid sequence

ISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPYHTWAKKKKHPHGQV

SBIP-436 -A (DK511-513)- Sequence Clone 2 pGEMT Sample 2– 210bp

Nucleotide sequence

GATTTCAGCTTCCCAGAAATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTAATTTTAGGATCCGCTAAT
ATTAACCAAAGATCTATGGCGGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTCATCACACTTGGGCCAAGAAGAAAAAACATCCAC
ATGGCCAGGTATAA

Translated amino acid sequence

ISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWAKKKKHPHGQV

SBIP-436 -B1 (DK511-514)- Sequence Clone 1 pGEMT Sample 3 – 664bp

Nucleotide sequence

GATTCAGCTTCCCAGAAATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTTATTTTAGGATCCGCTAATA
TTAACCAAAGATCTATGGCTGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTCATCACACTTGGGCTAACAAGAAAAAACATCCACA
TGGCCAGGTCAGTGAGTTTTTATTTCCCTCATTCTCCTTCCCAAATAAGGAAACTAGAACATTTTTGTGATGATCAAAACTAGCAAAAATAATAGCA
GTACACTATTTGTTATCCATTCTAGCCAGTCTACAAAAGGATCTCCTTGACATATTTGATTAAAGTTAACAAGTGCGGCTGCAGGTATATGGTTATA
GAATGTCTCTGTGGGCAGAGCATATGGGCAAGTTAGACGATTGCTTCACAAAGCCAGAAAGTTTGGACTGTGTGAAGCATGTGAATAAGGTTGCTGA
AGATAATTGGAACAGATTCACTGCTGAGGAGTTCAAACCTTTACAAGGTCATCTTCTCAAGTACCCAGTCAAAGTAGATTCTGATGGGAAAGTAAGT
TCCTTACCTGGTCATGAATATTTTCCTGATGTTGGTGGTAAAGTACTAGGAGCTCGAACCAATCTTCCTGATGCTTTGACCACAA

Translated amino acid sequence

DSASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKIIAVH
YLLSILASLQKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPG
HEYFPDVGGKVLGARTNLPDALTT

SBIP-436 -B1 (DK511-514)- Sequence Clone 2 pGEMT Sample 4 – 490bp

Nucleotide sequence

GATTTCAGCTTCCCAGAAATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTTATTTTAGGATCCGCTAAT
ATTAACCAAAGATCTATGGCTGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTCATCACACTTGGGCTAACAAGAAAAAACATCCAC
ATGGCTAGGTACATGGTTATAGAATGTCTCTGTGGGCAGAGCATATGGGCAAGTTAGACGATTGCTTCACAAAGCCAGAAAGTTTGGACTGTGTGAA
GCATGTGAATAAGGTTGCTGAAGATAATTGGAACAGATTCACTGCTGAGGAGTTCAAACCTTTACAAGGTCATCTTCTCAAGTACCCAGTCAAAGTA
GATTCTGGTGGGAAAGTAAGTTCCTTACCTGGTCATGAATATTTTCCTGATGTTGGTGGTAAAGTACTAGGAGCTCGAACCAATCTTCCTGATGCTT
TGACCACA

Translated amino acid sequence

ISASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGVHGYRMSLWAEHMGKLDDCFTKPESLDCVKH
VNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSGGKVSSLPGHEYFPDVGGKVLGARTNLPDALTT

SBIP-436 -B1 (DK511-514)- Sequence Clone 1 TOPO Sample 5 – 664bp

Nucleotide sequence

CCTTTCAGCTTCCCAGAAATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTTATTTTAGGATCCGCTAAT
ATTAACCAAAGATCTATGGCTGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTCATCACACTTGGGCTAACAAGAAAAAACATCCAC
ATGGCCAGGTCAGTGAGTTTTTATTTCCCTCATTCTCCTTCCCAAATAAGGAAACTAGAACATTTTTGTGATGATCAAAACTAGCAAAAATAATAGC
AGTACACTATTTGTTATCCATTCTAGCCAGTCTACGAAAGGATCTCCTTGACATATTTGATTAAAGTTAACAAGTGCGGCTGCAGGTATATGGTTAT
AGAATGTCTCTGTGGGCAGAGCATATGGGCAAGTTAGACGATTGCTTCACAAAGCCAGAAAGTTTGGACTGTGTGAAGCATGTGAATAAGGTTGCTG
AAGATAATTGGAACAGATTCACTGCTGAGGAGTTCAAACCTTTACAAGGTCATCTTCTCAAGTACCCAGTCAAAGTAGATTCTGATGGGAAAGTAAG
TTCCTTACCTGGTCATGAATATTTTCCTGATGTTGGTGGTAAAGTACTAGGAGCTCGAACCAATCTTCCTGATGCTTTGACCACAAAG

Translated amino acid sequence

LSASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKIIAVH
YLLSILASLRKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPG
HEYFPDVGGKVLGARTNLPDALTT

SBIP-436 -B1 (DK511-514)- Sequence Clone 8 TOPO Sample 6 – 490bp

Nucleotide sequence

CTTTCAGCTTCCCAGAAATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTTATTTTAGGATCCGCTAATATTAACCAAAGGTCT
ATGGCTGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTCATCACACTTGGGCTAACAAGAAAAAACATCCACATGGCCAGGTATATGGTTATAGAATGTCT
CTGTGGGCAGAGCATATGGGCAAGTTAGACGATTGCTTCACAAAGCCAGAAAGTTTGGACTGTGTGAAGCATGTGAATAAGGTTGCTGAAGATAATTGGAACAGATTCACT
GCTGAGGAGTTCAAACCTTTACAAGGTCATCTTCTCAAGTACCCAGTCAAAGTAGATTCTGATGGGAAAGTAAGTTCCTTACCTGGTCATGAATATTTTCCTGATGTTGGT
GGTAAAGTACTAGGAGCTCGAACCAACCTTCCTGATGCTTTGACCACAAAG

Translated amino acid sequence

LSASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPESLDCVK
HVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPDALTT

Figure 30: SBIP-436 -A and -B1 sequenced gene segments. Sequences were trimmed of vector components and translated using EXPASY bioinformatics translate tool. Sample 1- Colony 1SBIP-436 -A pGEMT Vector, Sample 2- Colony 2 SBIP-436 -A pGEMT Vector, Sample 3- Colony 1SBIP-436 -B1 pGEMT Vector, Sample 4- Colony2 SBIP-436 -B1 pGEMT Vector, Sample 5- Colony 1SBIP-436 -B1 TOPO Vector, Sample 5- Colony 8 SBIP-436 -B1 TOPO Vector.

Clones 1 and 2 from SBIP-436 -A in pGEMT Vector were selected for sequencing to represent the 210bp segment. Clones 1 and 2 from SBIP-436 -B1 in pGEMT Vector as well as clones 1 and 8 from SBIP-436 -B1 in TOPO Vector were selected to represent the 490bp/~700bp segments.

Sequencing revealed which gene segments each sample represented. Samples 1 and 2 represent 210bp segment shown in Figure 31. Samples 3 and 5 represent the 664bp amplicon show in Figure 32. Samples 4 and 6 represent 490bp amplicon shown in Figure 33. Figure 34 shows clustal alignment of all clones and Figure 35 shows clustal alignment of samples 5 and 6.

Clustal Alignment of Sample 1 and Sample 2

```
EXP1S1              --------------------------ISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP1S2              --------------------------ISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
NtSGN-U444527       NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 60
                                               **********************************

EXP1S1              QRSMAGSRDTEIAMGAYQPYHTWAKKKKHPHGQV---------------------------- 68
EXP1S2              QRSMAGSRDTEIAMGAYQPHHTWAKKKKHPHGQV---------------------------- 68
NtSGN-U444527       QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPESLD 120
                    ********************-****-*********

EXP1S1              ------------------------------------------------------------
EXP1S2              ------------------------------------------------------------
NtSGN-U444527       CVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGA 180


EXP1S1              ----------
EXP1S2              ----------
NtSGN-U444527       RTNLPDALTT 190
```

Figure 31: CLUSTAL 2.1 multiple sequence alignment of *N. tabacum* SGN-U444527 (190aa), SBIP-436 -A (DK511-513)- Sequence Clone 1 pGEMT Sample 1 (68aa), SBIP-436 -A (DK511-513)- Sequence Clone 2 pGEMT Sample 2 (68aa).

Clustal Alignment of Sample 3 and Sample 5

```
EXP2S3              ---------------------------DSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP2S5              ---------------------------LSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
NtSGN-U444527       NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 60
                                                 ********************************

EXP2S3              QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKII 94
EXP2S5              QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKII 94
NtSGN-U444527       QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHG---------------------------- 92
                    ********************************

EXP2S3              AVHYLLSILASLQKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 154
EXP2S5              AVHYLLSILASLRKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 154
NtSGN-U444527       ---------------------------QVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 126
                                                 ********************************

EXP2S3              KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 214
EXP2S5              KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 214
NtSGN-U444527       KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 186
                    ************************************************************

EXP2S3              ALTT 218
EXP2S5              ALTT 218
NtSGN-U444527       ALTT 190
                    ****
```

Figure 32: CLUSTAL 2.1 multiple sequence alignment of *N. tabacum* SGN-U444527 (190aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 1 pGEMT Sample 3 (218aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 2 pGEMT Sample 5 (218aa).

Clustal Alignment of Sample 4 and Sample 6

```
NtSGN-U444527       NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 60
EXP2S6              ---------------------------LSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP2S4              ---------------------------ISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
                                                 :********************************

NtSGN-U444527       QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPESLD 120
EXP2S6              QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVYGYRMSLWAEHMGKLDDCFTKPESLD 94
EXP2S4              QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHG-VHGYRMSLWAEHMGKLDDCFTKPESLD 93
                    ******************************** *:**************************

NtSGN-U444527       CVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGA 180
EXP2S6              CVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGA 154
EXP2S4              CVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSGGKVSSLPGHEYFPDVGGKVLGA 153
                    ************************************.***********************

NtSGN-U444527       RTNLPDALTT 190
EXP2S6              RTNLPDALTT 164
EXP2S4              RTNLPDALTT 163
                    **********
```

Figure 33: CLUSTAL 2.1 multiple sequence alignment of *N. tabacum* SGN-U444527 (190aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 1 pGEMT Sample 4 (163aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 8 pGEMT Sample 6 (164aa).

Clustal Alignment of All Sequenced Samples

```
NtSGN-U444527      NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 60
SBIP_436_Y2H       ----------------------------SASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 33
EXP1_S1_210        ----------------------------ISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP1_S2_210        ----------------------------ISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP2_S4_490        ----------------------------ISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP2_S6_490        ----------------------------LSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP2_S3_664        ----------------------------DSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP2_S5_664        ----------------------------LSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
                                                 **********************************

NtSGN-U444527      QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQ---------------------------- 93
SBIP_436_Y2H       QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQV--------------------------- 67
EXP1_S1_210        QRSMAGSRDTEIAMGAYQPYHTWAKKKKHPHGQV--------------------------- 68
EXP1_S2_210        QRSMAGSRDTEIAMGAYQPHHTWAKKKKHPHGQV--------------------------- 68
EXP2_S4_490        QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHG----------------------------- 66
EXP2_S6_490        QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHG----------------------------- 66
EXP2_S3_664        QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKII 94
EXP2_S5_664        QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKII 94
                   ******************.****.*******

NtSGN-U444527      ----------------------------VYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 126
SBIP_436_Y2H       ------------------------------------------------------------
EXP1_S1_210        ------------------------------------------------------------
EXP1_S2_210        ------------------------------------------------------------
EXP2_S4_490        ----------------------------VHGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 99
EXP2_S6_490        ----------------------------QVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 100
EXP2_S3_664        AVHYLLSILASLQKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 154
EXP2_S5_664        AVHYLLSILASLRKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 154

NtSGN-U444527      KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 186
SBIP_436_Y2H       ------------------------------------------------------------
EXP1_S1_210        ------------------------------------------------------------
EXP1_S2_210        ------------------------------------------------------------
EXP2_S4_490        KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSGGKVSSLPGHEYFPDVGGKVLGARTNLPD 159
EXP2_S6_490        KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 160
EXP2_S3_664        KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 214
EXP2_S5_664        KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 214

NtSGN-U444527      ALTT 190
SBIP_436_Y2H       ----
EXP1_S1_210        ----
EXP1_S2_210        ----
EXP2_S4_490        ALTT 163
EXP2_S6_490        ALTT 164
EXP2_S3_664        ALTT 218
EXP2_S5_664        ALTT 218
```

Figure 34: CLUSTAL 2.1 multiple sequence alignment of *N. tabacum* SGN-U444527 (190aa), SBIP-436 Yeast Two-Hybrid Clone (67aa), SBIP-436 -A (DK511-513)- Sequence Clone 1 pGEMT Sample 1 (68aa), SBIP-436 -A (DK511-513)- Sequence Clone 2 pGEMT Sample 2 (68aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 1 pGEMT Sample 4 (163aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 8 pGEMT Sample 6 (164aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 1 pGEMT Sample 3 (218aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 1 pGEMT Sample 5 (218aa)

Clustal Alignment of Sample 5 and Sample 6

```
NtSGN-U444527       NFYCLGNREELHGGSKPNPASSSGDVISASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 60
EXP2_S5_664         ---------------------------LSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
EXP2_S6_490         ---------------------------LSASQKFGRFMIYVHAKGMIVDDEYVILGSANIN 34
                                               :**********************************

NtSGN-U444527       QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHG----------------------------- 92
EXP2_S5_664         QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKII 94
EXP2_S6_490         QRSMAGSRDTEIAMGAYQPHHTWANKKKHPHG----------------------------- 66
                    ********************************

NtSGN-U444527       ---------------------------QVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 126
EXP2_S5_664         AVHYLLSILASLRKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 154
EXP2_S6_490         ---------------------------QVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVN 100
                                               **********************************

NtSGN-U444527       KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 186
EXP2_S5_664         KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 214
EXP2_S6_490         KVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPGHEYFPDVGGKVLGARTNLPD 160
                    ************************************************************

NtSGN-U444527       ALTT 190
EXP2_S5_664         ALTT 218
EXP2_S6_490         ALTT 164
                    ****
```

Figure 35: CLUSTAL 2.1 multiple sequence alignment of *N. tabacum* SGN-U444527 (190aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 1 pGEMT Sample 5 (218aa), SBIP-436 -B1 (DK511-514)- Sequence Clone 8 pGEMT Sample 6 (164aa).

Analysis of the sequenced clones of the SBIP-436 gene reveal multiple insights into its genetic nature. As expected, -A samples (1 and 2) align well with the *N. tabacum* SGN-U444527. The same can be said for the -B1 samples (3-6). The -B1 unexpected amplicon that yielded an approximate 700bp band was revealed to be a 174bp insert that gave the segment a 664bp size. In regards to expression, it was found that the -A amplifies a segment that is common to both the 664bp

amplicon and the 490 bp amplicon in the -B1 set. Therefore, hypothetically if there was a 20% decrease in expression of the 664bp transcript and a 20% increase in the 490bp transcript, then the -A 210bp transcript would register no change in expression via visualization. This was important to keep in mind when performing expression analysis. Given that there are multiple isoforms of PLDs and implications of alternative splicing, it was logical to analyze the 664bp transcript due to its 174bp insert. A splice analysis was performed to identify the 664bp transcript as a separate isoform or an alternatively spliced transcript.

Splice Analysis of SBIP-436 -B1 – 664bp Transcript

Wang Computings ASSP (Alternative Splice Site Predictor) (http://wangcomputing.com/assp/evaluation.html) was used to search for acceptor and donor sites as well as possible branch points within the 664bp transcript. The Sample 5 sequence containing the gene insert was used for splice analysis shown in Figure 36.

SBIP-436 -B1 (DK511-514)- Sequence Clone 1 TOPO Sample 5 – 664bp

Nucleotide sequence

```
CCTTTCAGCTTCCCAGAAATTTGGACGGTTTATGATTTATGTACACGCCAAGGGGATGATAGTGGACGATGAGTATGTTATTTTAGGATCCGCTAAT
ATTAACCAAAGATCTATGGCTGGTTCAAGAGACACAGAGATAGCTATGGGAGCATATCAGCCTCATCACACTTGGGCTAACAAGAAAAAACATCCAC
ATGGCCAGGTCAGTGAGTTTTTATTTCCCTCATTCTCCTTCCCAAATAAGGAAACTAGAACATTTTTGTGATGATCAAAACTAGCAAAAATAATAGC
AGTACACTATTTGTTATCCATTCTAGCCAGTCTACGAAAGGATCTCCTTGACATATTTGATTAAAGTTAACAAGTGCGGCTGCAGGTATATGGTTAT
AGAATGTCTCTGTGGGCAGAGCATATGGGCAAGTTAGACGATTGCTTCACAAAGCCAGAAAGTTTGGACTGTGTGAAGCATGTGAATAAGGTTGCTG
AAGATAATTGGAACAGATTCACTGCTGAGGAGTTCAAACCTTTACAAGGTCATCTTCTCAAGTACCCAGTCAAAGTAGATTCTGATGGGAAAGTAAG
TTCCTTACCTGGTCATGAATATTTTCCTGATGTTGGTGGTAAAGTACTAGGAGCTCGAACCAATCTTCCTGATGCTTTGACCACAAAG
```

Translated amino acid sequence

```
LSASQKFGRFMIYVHAKGMIVDDEYVILGSANINQRSMAGSRDTEIAMGAYQPHHTWANKKKHPHGQVSEFLFPSFSFPNKETRTFLSKLAKIIAVH
YLLSILASLRKDLLDIFDSQVRLQVYGYRMSLWAEHMGKLDDCFTKPESLDCVKHVNKVAEDNWNRFTAEEFKPLQGHLLKYPVKVDSDGKVSSLPG
HEYFPDVGGKVLGARTNLPDALTT
```

Figure 36: SBIP-436 - Sample 5- Colony 1 SBIP-436 -B1 TOPO Vector sequenced gene segment. Sequences were trimmed of vector components and translated using EXPASY bioinformatics translate tool.

The ASSP analysis revealed multiple acceptor and donor sites. The results revealed a Constitituve Donor site and an acceptor site that aligned well with the 174bp insert mapped from the Clustal alignments as well as maintained sufficient confidence scores. The ASSP acceptor and donor site list are shown in Table 5. The outlining boxes show the constitutive donor site at 198bp region and the Alternative isoform/cryptic Acceptor at 373 region.

Table 5: ASSP- SBIP-436Acceptor and Donor Sites

				Activations**			
Position (bp)	Putative splice site	Sequence	Score*	Intron GC*	Alt./Cryptic	Constitutive	Confidence**
68	Alt. isoform/cryptic donor	TGGACGATGAgtat gttatt	6.975	0.371	0.828	0.119	0.857
83	Constitutive acceptor	gttatttttagGATC CGCTAA	4.353	0.371	0.337	0.638	0.472
198	Constitutive donor	ACATGGCCAGgtca gtgagt	13.037	0.343	0.168	0.776	0.783
202	unclassified donor	GGCCAGGTCAgtga gttttt	10.666	0.329	0.420	0.497	0.000
241	Alt. isoform/cryptic acceptor	cccaaataagGAAA CTAGAA	2.655	0.386	0.666	0.318	0.523
314	Alt. isoform/cryptic acceptor	tccattctagCCAG TCTACG	5.520	0.286	0.877	0.116	0.868
372	Alt. isoform/cryptic donor	GCGGCTGCAGgtat atggtt	9.055	0.443	0.755	0.180	0.762
373	Alt. isoform/cryptic acceptor	gcggctgcagGTAT ATGGTT	2.280	0.414	0.701	0.288	0.590
404	Alt. isoform/cryptic acceptor	ctgtgggcagGCA TATGGG	3.693	0.400	0.766	0.223	0.708
454	Alt. isoform/cryptic donor	AGTTTGGACTgtgt gaagca	4.801	0.400	0.927	0.052	0.944
474	Alt. isoform/cryptic donor	TGTGAATAAGgttg ctgaag	4.632	0.400	0.942	0.040	0.958
529	Alt. isoform/cryptic donor	CCTTTACAAGgtca tcttct	7.818	0.400	0.848	0.112	0.868
542	Alt. isoform/cryptic donor	ATCTTCTCAAgtac ccagtc	7.787	0.386	0.602	0.323	0.463
544	Alt. isoform/cryptic acceptor	tcttctcaagTACC CAGTCA	2.693	0.414	0.931	0.065	0.930
551	Alt. isoform/cryptic acceptor	aagtacccagTCAA AGTAGA	2.383	0.414	0.927	0.070	0.924
555	Alt. isoform/cryptic donor	CCCAGTCAAAgtag attctg	6.759	0.386	0.803	0.143	0.822
573	Alt. isoform/cryptic donor	TGATGGGAAAgtaa gttcct	13.024	0.414	0.659	0.258	0.609
589	Alt. isoform/cryptic donor	TCCTTACCTGgtca tgaata	4.822	0.414	0.967	0.023	0.976
616	Alt. isoform/cryptic donor	GATGTTGGTGgtaa agtact	7.448	0.443	0.886	0.081	0.908

A splice analysis map (Figure 37) shows the donor site at 198bp mark and the acceptor site at the 373 bp mark (Table 5). The map also shows a decrease in codon usage between these 2 sites suggesting that this section is indeed a splice variant.

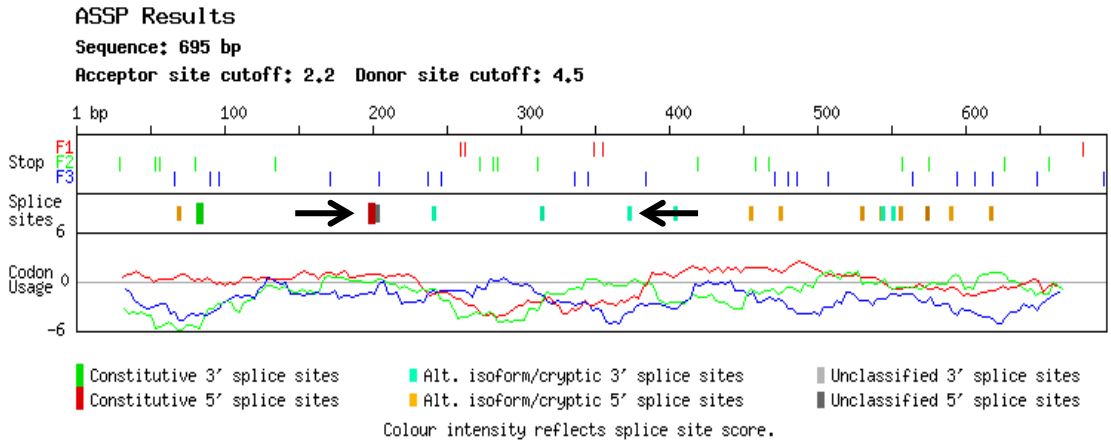


Figure 37: ASSP Slice Site Map. Splice sites found in middle track. The "Donor" site in red indicated by left arrow. The "Acceptor" site in green indicated by right arrow. Codon usage is mapped in bottom track.

The sequence of Sample 5 was then searched for the actual donor and acceptor sites. The donor site was identified by the CAG –GT sequence and the acceptor site was identified by the AG—G sequence. A possible branch point was also found as a TTTGAT sequence. These components can be seen in sequence in Figure 38.

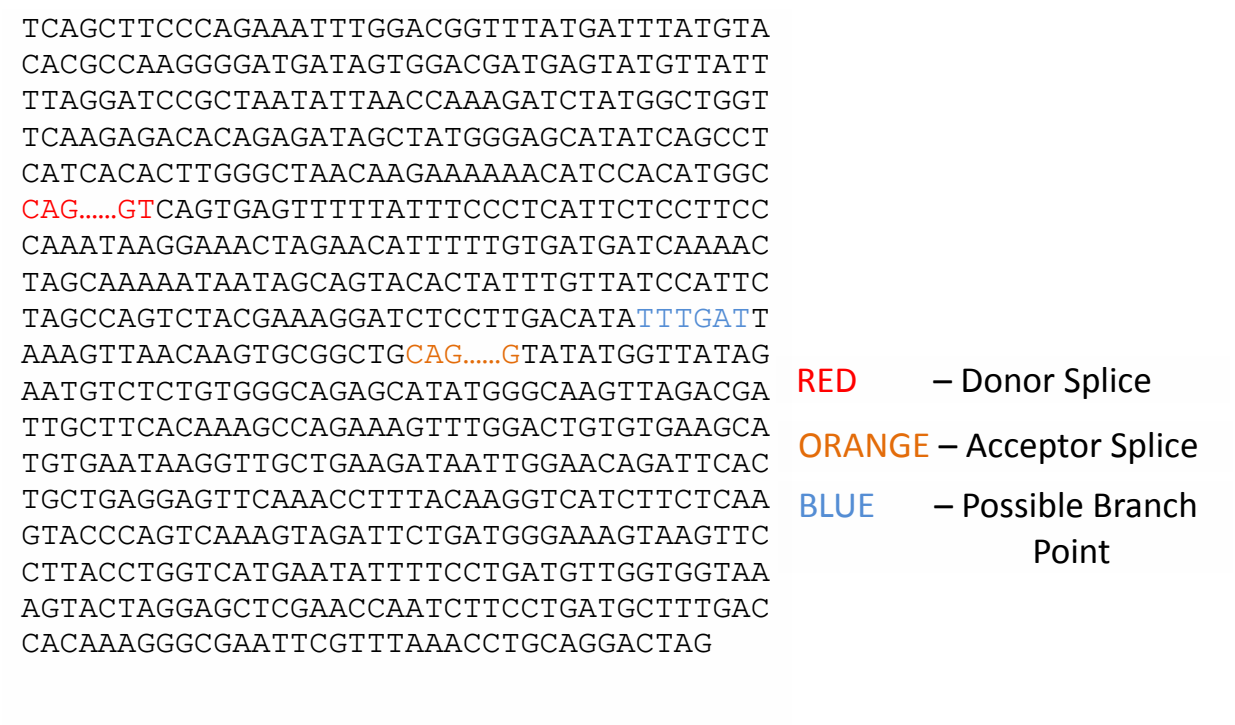


Figure 38: Splice sites in SBIP-436 -B1 Sample 5 Sequence. Donor site is in red. Acceptor Site is in orange. Possible branch point is in blue.

PCR Amplification of 174 bp Gene Insert

Primers were made to amplify the 174bp gene insert from the SBIP-436 -B1 664bp transcript so that the transcript expression could be viewed separately. The primer set was DK619 forward primer 5' TTCCCTCATTCTCCTTCCCAA 3' and DK618 reverse primer 5' ACCTGCAGCCGCACTTGTTAA 3'. SBIP-436 -B2 (DK619-DK618), was designed to amplify a product of 174bp. PCR conditions were optimized in regards to gene length and primer annealing temperatures. Optimal annealing temperature was set at 55^{o}C and the optimal extension time was set at 45 seconds. For best amplification for visualization of the 174bp segment 32 PCR cycles were used. PCR product was then subjected to agarose gel electrophoresis shown in Figure 39.

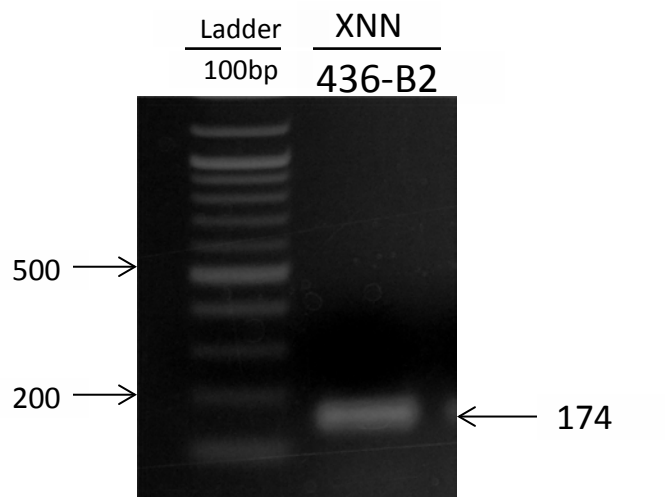


Figure 39: PCR amplification of SBIP-436 -B2. 1.2% Agarose gel showing expression of 174bp fragment of 664bp transcript segment of SBIP-436.

PCR Amplification of Full SBIP-436

The SBIP-436 full gene primer sets SBIP-436 Full-1 (DK547-548) and SBIP-436 Full-2 (DK602-616) were designed to amplify a product of approximately 2577bp according to SBIP-436 gene construct). PCR conditions were optimized in regards to gene length and primer annealing temperatures. Optimal annealing temperature was set at 55^{o}C and the optimal extension time was set at 2 minutes 45 seconds. PCR amplification was allowed until 34 PCR cycles was reached. PCR product was then subjected agarose gel electrophoresis in a 1.0% gel, visualized over UV, and photographed. The PCR product can be viewed in Figures 40 and 41.

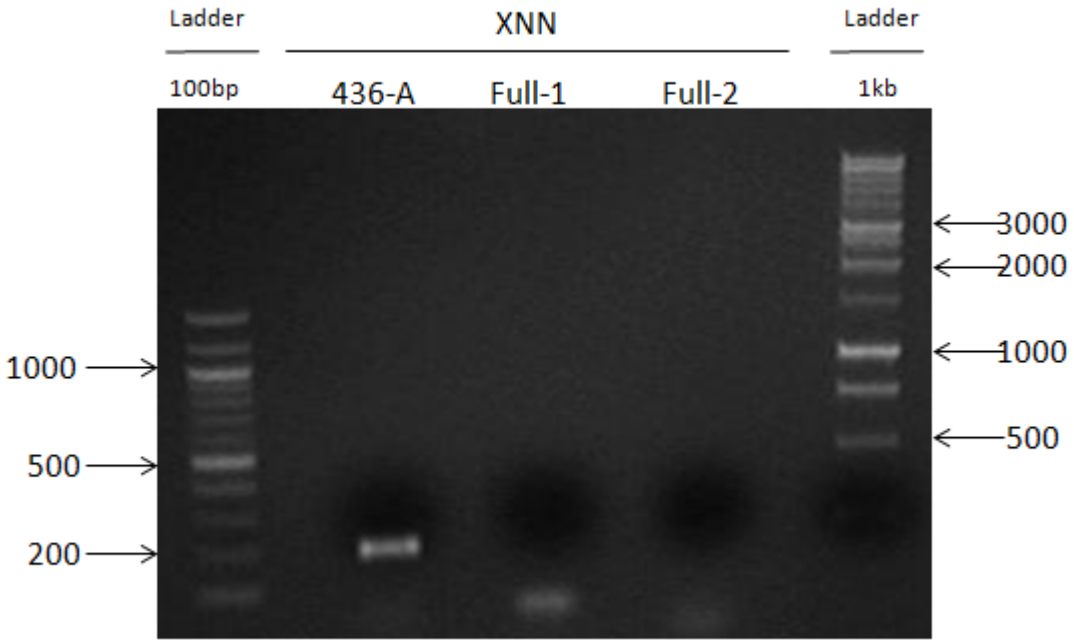


Figure 40: PCR amplification of SBIP-436 Full-1 (DK547-548) and SBIP-436 Full-2 (DK602-616) using *N. tabacum* XNN cDNA. SBIP-436 -A (DK511-DK513) used as control.

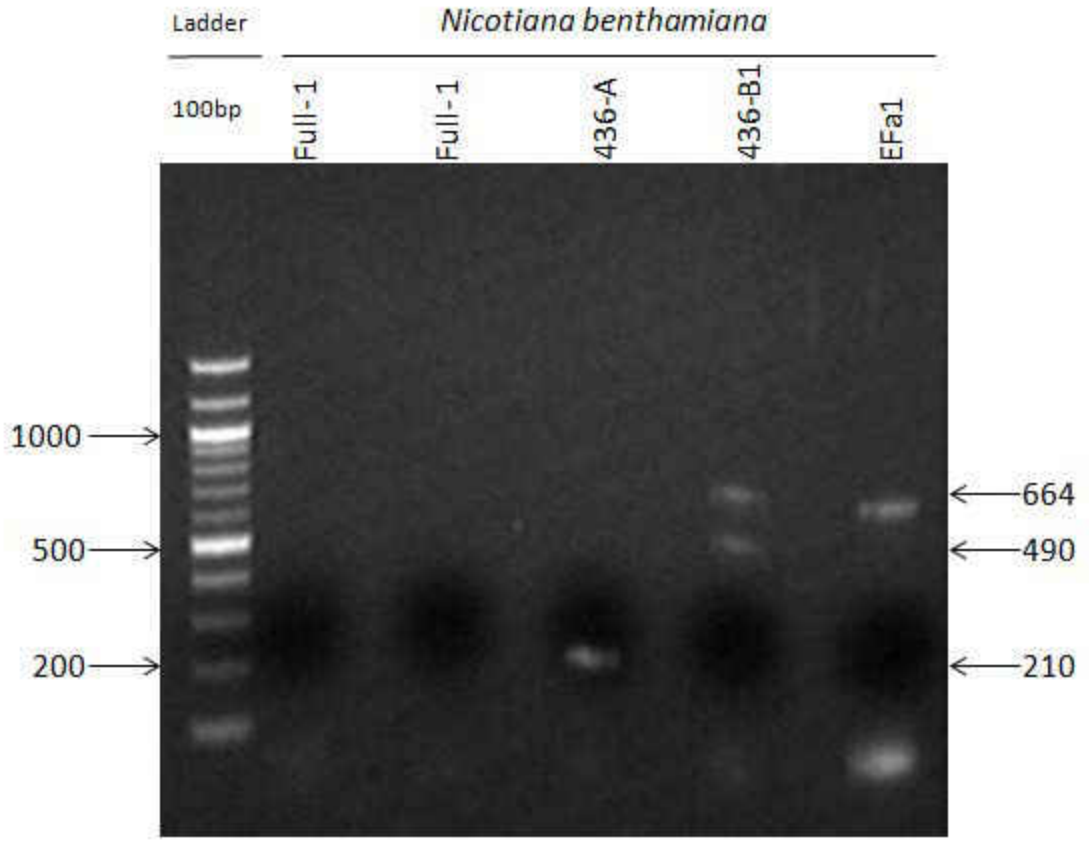


Figure 41: PCR amplification of SBIP-436 Full-1 (DK547-548) and SBIP-436 Full-2 (DK602-616) using *N. benthamiana* cDNA. SBIP-436 -A (DK511-DK513) and SBIP-436 -B1 (DK511-DK513) were amplified. *Efa1* was used as control.

The SBIP-436 1000bp gene primer sets SBIP-436 1kb-A (DK602-DK603) and SBIP-436 1kb-B (DK604-605) were designed to amplify a product of approximately 1000bp According to SBIP-436 gene construct). The 1000bp segments were intended to amplify smaller portions of the SBIP-436 construct. PCR conditions were optimized in regards to gene length and primer annealing temperatures. Optimal annealing temperature was set at 55^{o}C and the optimal extension time was set at 1 minute 15 seconds. PCR amplification was allowed until 34 PCR

cycles was reached. PCR product was then subjected agarose gel electrophoresis in a 1.0% gel, visualized over UV, and photographed. The PCR products can be viewed in Figure 42 and 43.

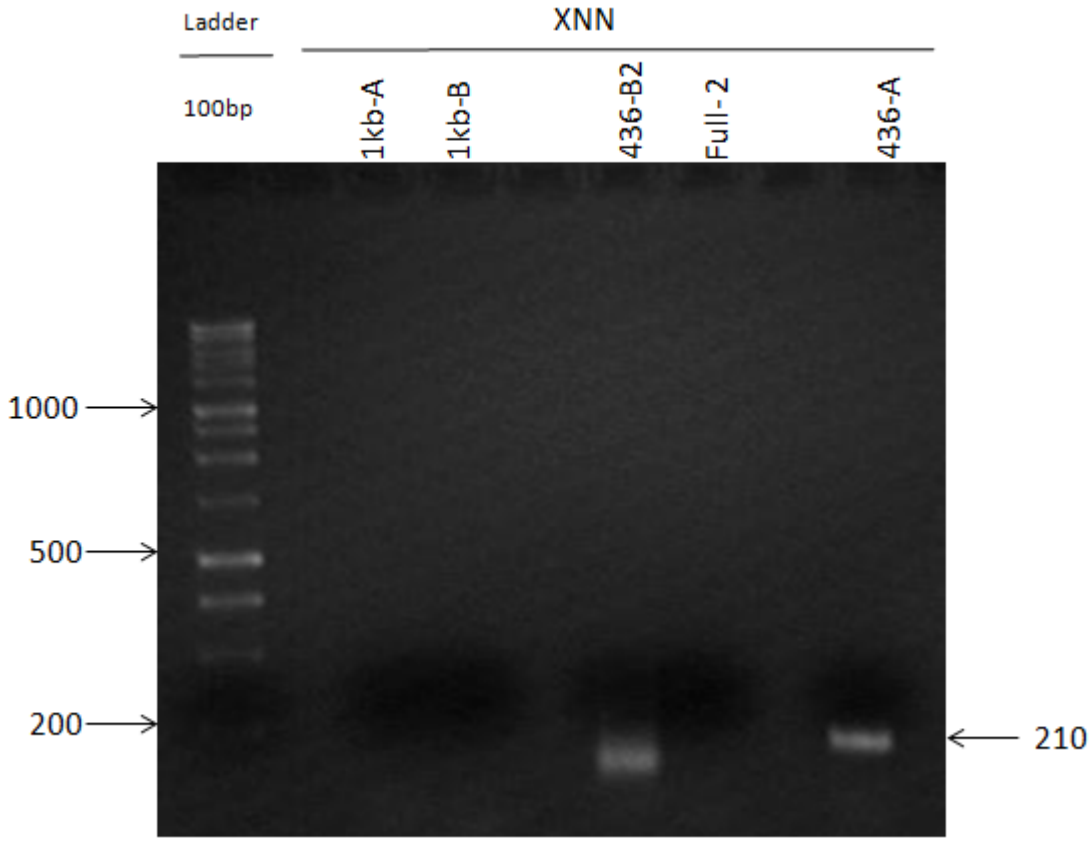


Figure 42: PCR amplification of SBIP-436 1kb-A (DK602-DK603) and SBIP-436 1kb-B (DK604-605) using *N. tabacum* XNN cDNA. SBIP-436 -A (DK511-DK513) and SBIP-436 -B2 (DK618-DK619) used as control.

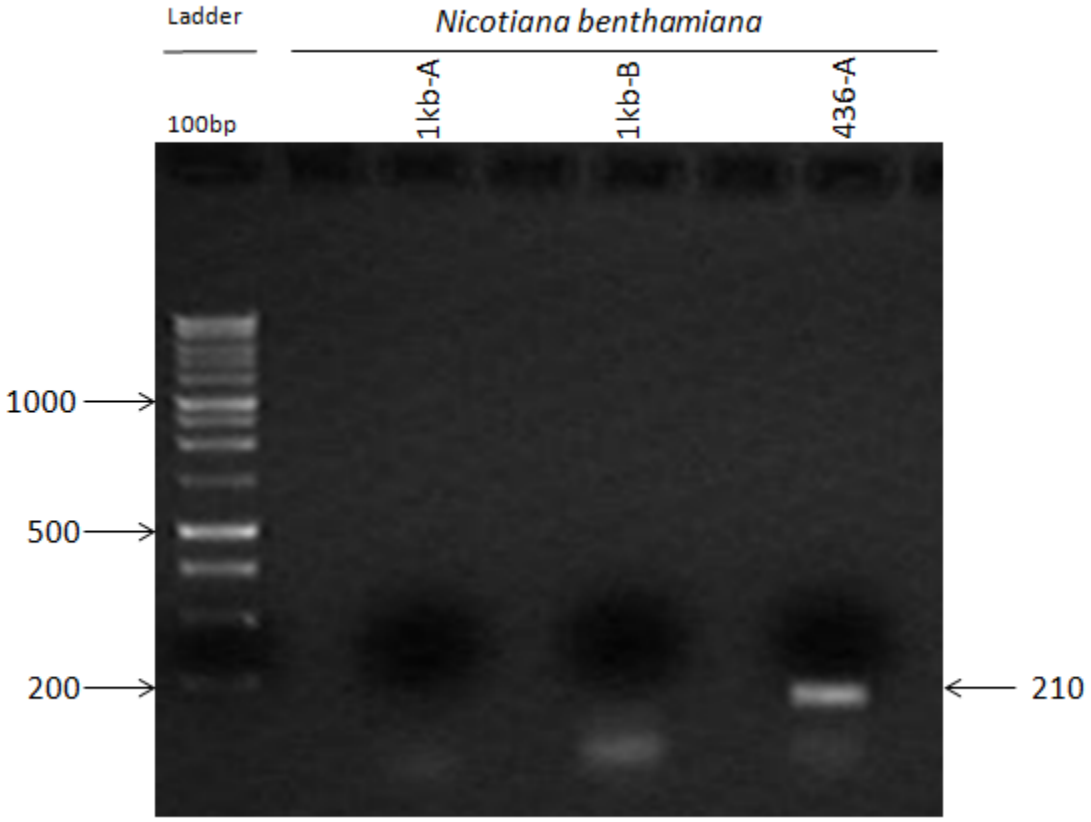


Figure 43: PCR amplification of SBIP-436 1kb-A (DK602-DK603) and SBIP-436 1kb-B (DK604-605) using *N. benthamiana* cDNA. SBIP-436 -A (DK511-DK513) used as control.

Touchdown PCR was performed using the SBIP-436 full gene primer sets SBIP-436 Full-1 (DK547-548) to amplify 2577bp SBIP-436. Touchdown PCR allows multiple annealing temperatures to better ensure an amplified product. The lower the annealing temperature the less specific amplified products will result. PCR conditions were optimized in regards to gene length and primer annealing temperatures. Optimal annealing temperature was first set at 55^{o}C and the optimal extension time was set at 2 minutes 45 seconds PCR amplification was allowed till 25 cycles of decreasing temperatures, beginning at 55^{o} in increments of 0.5^{o} followed by 20 cycles

of 42.5^{o} was reached. PCR product was then subjected agarose gel electrophoresis in a 1.0% gel, visualized over UV, and photographed. The PCR product can be viewed in Figure 44.

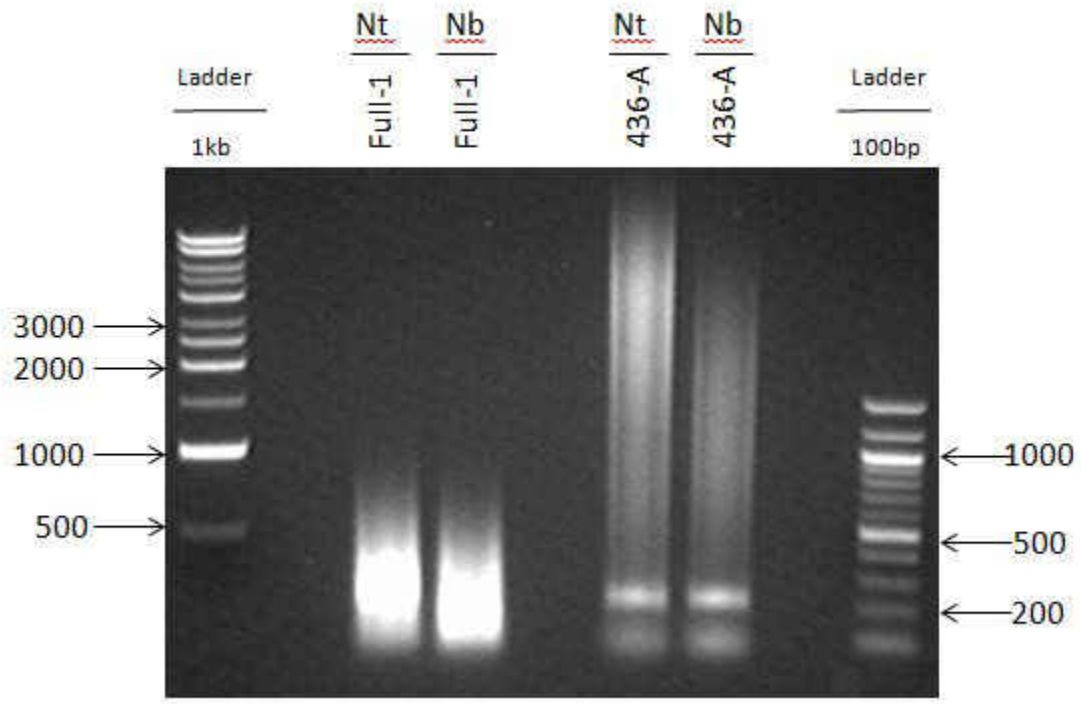


Figure 44: Touchdown PCR amplification of SBIP-436 Full-1 (DK547-548) and SBIP-436 Full-2 (DK602-616) using *N. tabacum* XNN and *N. benthamiana* cDNA. PCR amplification was allowed till 25 cycles of decreasing temperatures in increments of 0.5^{o} followed by 20 cycles of 42.5^{o} was reached. SBIP-436 -A (DK511-DK513) used as control.

Despite the multiple avenues of approaching amplification of the predicted SBIP-436 full gene, there were no successful amplifications. The attempts to amplify 1000bp segments instead of full length 436 gene also yielded negative results. Although these larger segments could not be

amplified in either *N. tabacum* or *N. benthamiana*, Figure 41 demonstrates that SBIP-436 -A and SBIP-436 -B1 were both amplified using *N. benthamiana* cDNA.

Expression Analysis of SBIP-436

An expression analysis was performed to reveal the role of SBIP-436 in plant defense. Expression of the SBIP-436 gene was analyzed in regards to SA mediated defense, predominantly associated with defense against biotic stressors, as well as defense against abiotic stressors. Tobacco plants were either infected with TMV or treated with 0.1mM salicylic acid to analyze the effect on expression of SBIP-436. Tobacco plants were treated with 300mM NaCl to analyze salt stress and mock treated plants were analyzed to represent the effects of wounding (abiotic stressors). SBIP-436 expression was also examined in each of the *N. tabacum* mutant lines (XNN, NahG, C3, and 1-2).

SBIP-436 Expression in *N. tabacum* Mutant Plants

Expression of SBIP-436 was examined in the tobacco mutant lines to test if decreased SA (NahG) or the absence of SABP2 (1-2) would have any effect on SBIP-436 expression. Expression of SBIP-436 in *N. tabacum* mutant plants is shown in Figure 45.

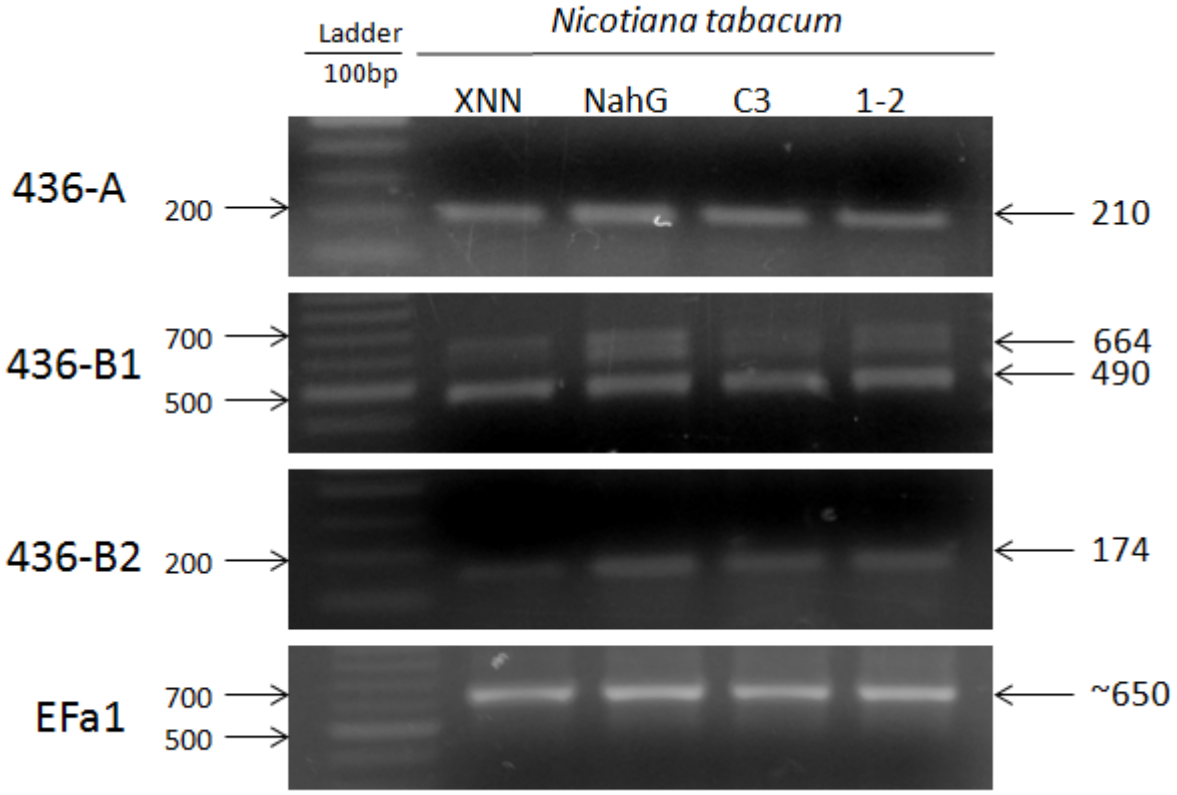


Figure 45: Expression of SBIP-436 in *N. tabacum* mutant lines. Expression agarose gels include: -A (DK511-DK513), -B1 (DK511-DK514), -B2 (DK619-DK618), and EF±1(Control). Actual amplicon size is labeled on the right hand side. 1µl of cDNA synthesized from total RNA from XNN, NahG, C3, and 1-2 was used for PCR amplifications.

SBIP-436 -A expression does not vary across various plants but slight variation does occur in SBIP-436 -B1 and SBIP-436 -B2. There is a slight increase in expression of the 664bp band in the *NahG* transgenic plants. This is further supported by SBIP-436 -B2 which is representative of the 174bp gene insert that separates the 490bp from the 664bp of the SBIP-436 -B1.

SBIP-436 Expression in TMV Infected Tissue

Six-week-old *N. tabacum* XNN plants were infected with TMV and samples taken at 0, 24, 48, and 72 hours postinoculation. Expression of SBIP-436 was analyzed via RT-PCR amplification and visualized via agarose gel electrophoresis. Expression of SBIP-436 upon TMV infection was performed to reveal the possible involvement of SBIP-436 in defense against biotic stressors and the SA mediated defense pathway. Expression of SBIP-436 after treatment with TMV is shown in Figure 46.

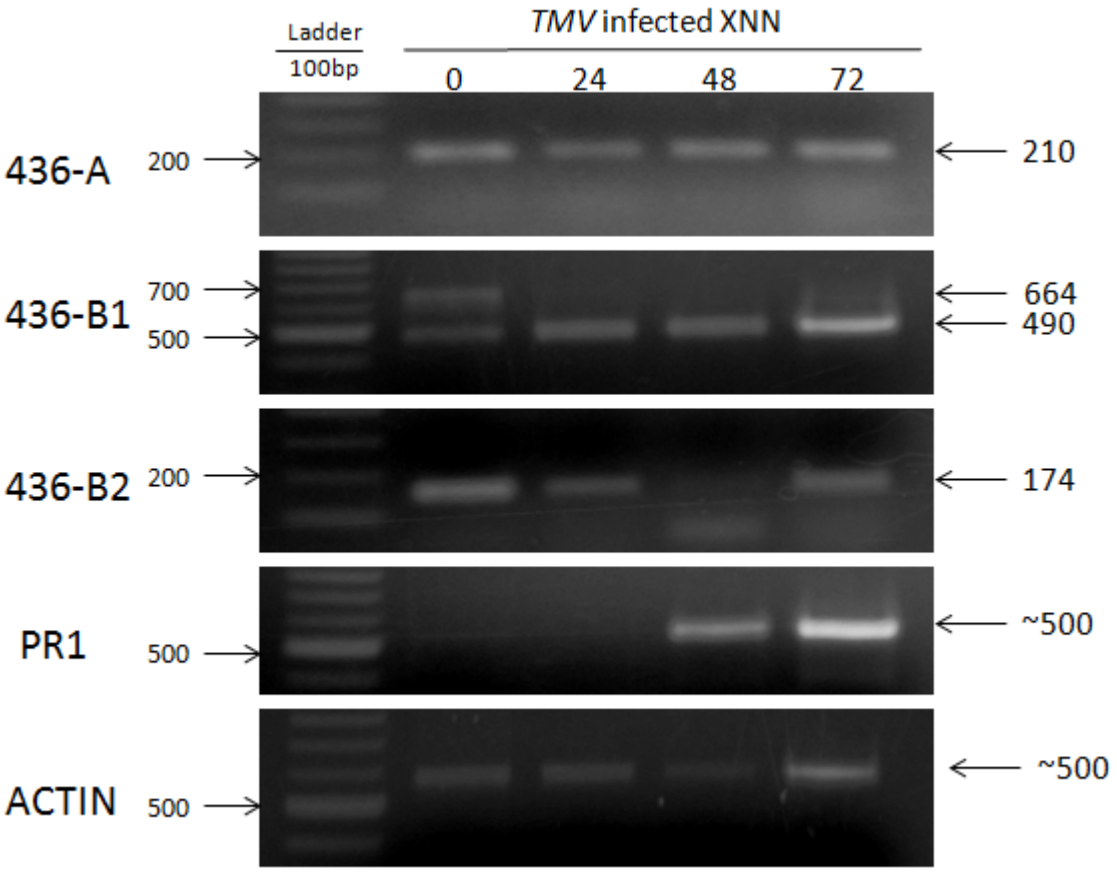


Figure 46: Expression of SBIP-436 in TMV infected *N. tabacum* XNN. Expression agarose gels include: -A (DK511-DK513), -B1 (DK511-DK514), -B2 (DK619-DK618), *PR1* (control for activation of SA pathway), and Actin (loading control). 1.2% Agarose gel showing expression of

gene segments of interest. One microliter of cDNA synthesized from total RNA from *TMV* infected XNN was used for PCR amplifications.

SBIP-436 -A primer set showed no visible change in expression across all time points. SBIP-436 -B1 primer set however demonstrated variability. There was a dominant switch in expression from the 664bp transcript to the 490bp transcript from 0hr to 24hr. The 490bp transcript remained dominant through the remaining 48hr and 72hr samples. The SBIP-436 -B2 expression yields a clearer picture of the modulation of expression of SBIP-436 transcripts. SBIP-436 -B2 primer set represents the 664bp transcript by amplifying only the 174bp insert that separates the 664bp transcript from the 490bp transcript found from sequencing (Figure 32; Figure 39). -B2 showed a significant decrease in expression from 0hr to 24hr, a continual decrease from 24hr to 48hr, followed by an increase at the 72hr mark. *Actin* gene amplification was used as a loading control and *PR1* was used as a positive control for activation of SA pathway upon TMV infection. TMV induced *PR1* expression at 48hr and a stronger expression at 72hr mark. This assured that the TMV infection was present and the SA mediated defense cascade was activated.

SBIP-436 Expression in Salicylic Acid Treated Tissue

Six-week-old *N. tabacum* XNN plants were treated with 0.1mM SA and samples taken at 0, 3, 6, 9, 12, and 24 hours post inoculation. Expression of SBIP-436 was analyzed via agarose gel electrophoresis. Expression of SBIP-436 under treatment with SA was performed to reveal the possible involvement of SBIP-436 in the SA mediated defense pathway. Expression of SBIP-436 after treatment with SA is shown in Figure 47.

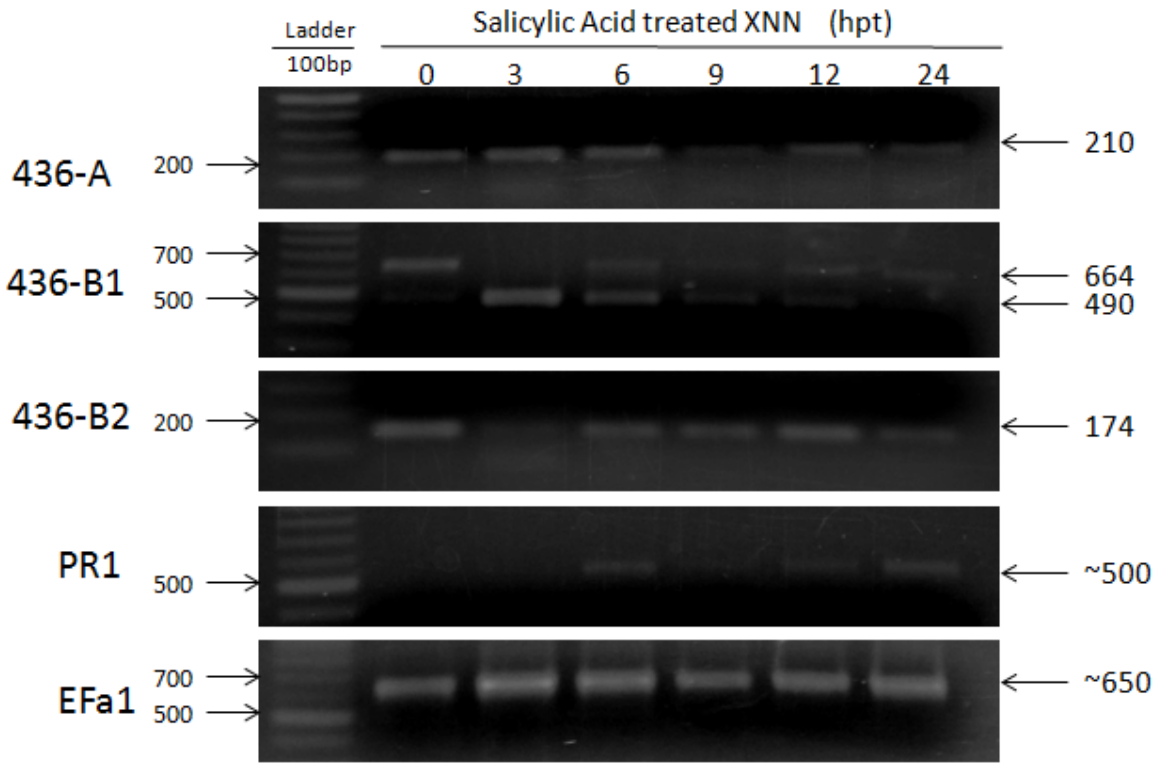


Figure 47: Expression of SBIP-436 in .1mM SA treated *N. tabacum* XNN. Expression agarose gels include: -A (DK511-DK513), -B1 (DK511-DK514), -B2 (DK619-DK618), *PR1* (Activation of SA pathway), and *EF±1* (Control). PCR was run for 30 cycles. Actual amplicon size is labeled on the right hand side. One microliter of cDNA synthesized from total RNA from SA treated XNN was used for PCR amplifications.

SBIP-436 -A primer set showed a drop in expression at the 9hr time point and an increase back normal level at the 12hr time point. SBIP-436 -B1 primer set demonstrated variability between the 2 transcripts. There was a dominant switch from the 664bp transcript to the 490bp transcript from 0hr to3hr. The expression of both transcripts seems to equalize at 6hr, 9hr, and 12hr time points. The 664bp transcript then becomes dominant at the 24hr time point. The SBIP-436 -B2 expression yields a clearer picture of the expression modulation of SBIP-436 transcripts by showing expression of the 174bp insert associated with the 664bp transcript. -B2 follows the

same trend as the 664bp transcript by decreases substantially at 3hr, increasing at the 6hr, 9hr, and 12hr time points, and decreasing slightly at the 24hr mark. *EF±1* was used as a control and *PR1* was used as a positive defense control. Exogenous SA treatment induced *PR1* expression at 6hr, a decrease at 9hr, an increase at 12hr, and the strongest expression at 24hr mark. This assured that the SA mediated defense cascade was active.

SBIP-436 Expression in NaCl Treated Tissue

Six-week-old *N. tabacum* XNN plants were treated with 300mM NaCl and samples taken at 0, 3, 6, 9, 12, and 24 hours postinoculation. Expression of SBIP-436 was analyzed via agarose gel electrophoresis. Expression of SBIP-436 under treatment with NaCl was performed to reveal the possible involvement of SBIP-436 in the defense against salt stress and ABA mediated defense. Expression of SBIP-436 after treatment with NaCl is shown in Figure 48.

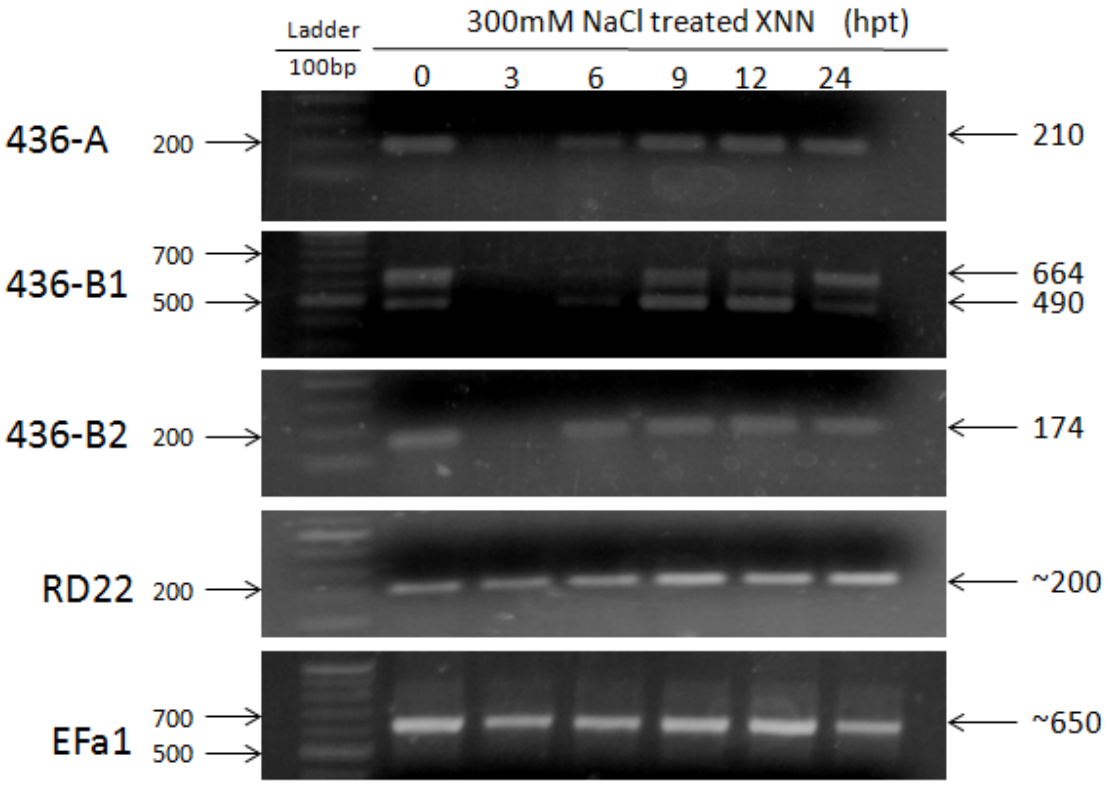


Figure 48: Expression of SBIP-436 in 300mM NaCl treated *N. tabacum* XNN. Expression agarose gels include: -A (DK511-DK513), -B1 (DK511-DK514), -B2 (DK619-DK618), *RD22* (Activation of abiotic stress response), and *EF±1* (Control). PCR was run for 30 cycles. Actual amplicon size is labeled on the right hand side. 1µl of cDNA synthesized from total RNA from NaCl treated XNN was used for PCR amplifications.

SBIP-436 -A primer set showed a significant drop in expression at the 3hr time point, a slight increase at the 6hr time point, and an increase to basal level that sustained through the 9hr, 12hr, and 24hr time points. SBIP-436 -B1 primer set demonstrated variability between the 2 transcripts. There was a significant decrease in expression from 0hr to 3hr. The expression of both transcripts increases slightly at the 6 hr mark and increase to basal levels with sustain

through the 9hr and 12hr time points. The 664bp transcript then becomes dominant at the 24hr time point. The SBIP-436 -B2 expression yields a clearer picture of the expression modulation of SBIP-436 transcripts by showing expression of the 174bp insert associated with the 664bp transcript. -B2 follows the same trend as the 664bp transcript by decreases substantially at 3hr, increasing and leveling at the 6hr, 9hr, 12hr, and 24hr mark. *EF±1* was used as a control and *RD22* was used as a positive defense control. NaCl treatment increased *RD22* expression at 9hr, and maintained through 12hr and 24hr time points. This assured that the NaCl treatment was active.

SBIP-436 Expression in Mock Treated Tissue

Mock treatment was used alongside *Pseudomonas* and *TMV* infection to rule out any effect of medium or wounding on overall expression results. This gave insight into the effect of wounding as an abiotic stressor. Wounding can be caused during sample taking leading to activation of defense mechanisms within the plant. Variation in SBIP-436 expression in mock treatment samples could implicate its involvement in defense against wounding. Expression of SBIP-436 after subjected to mock treatment/wounding is shown in Figures 49 and 50.

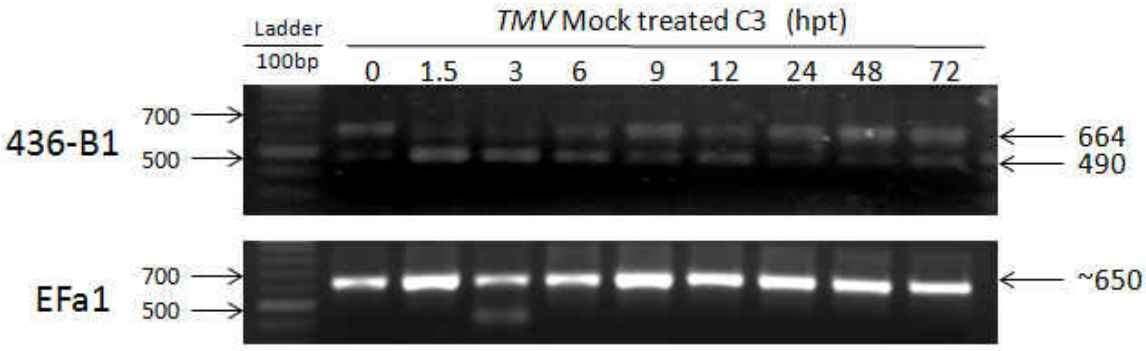


Figure 49: Expression of SBIP-436 in buffer (0.05 M Na_2HPO_4) treated *N. tabacum* C3 plants. Primers used for expression include: -B1 (DK511-DK514), and EF±1 (Control). 1.2% agarose gel showing expression of gene segments of interest.

Modulation between the two transcripts amplified from the -B1 primer set seems to be occurring during the Mock treatment. There was a dominant switch from the 664bp transcript to the 490bp transcript from 0hr to 1.5hr. The 490bp transcript remained dominant through the 1.5hr and 3 hr samples. The expression of both transcripts seems to level out at 6hr, 9hr, and 12hr time points. The 664bp transcript then becomes dominant at the 24hr, 48hr, and 72hr time points. It is important to note that the 24hr, 48hr, and 72hr expression resembles the expression of the 0hr sample, which stands contrary to the SBIP-436 expression in the TMV infected samples (Figure 46).

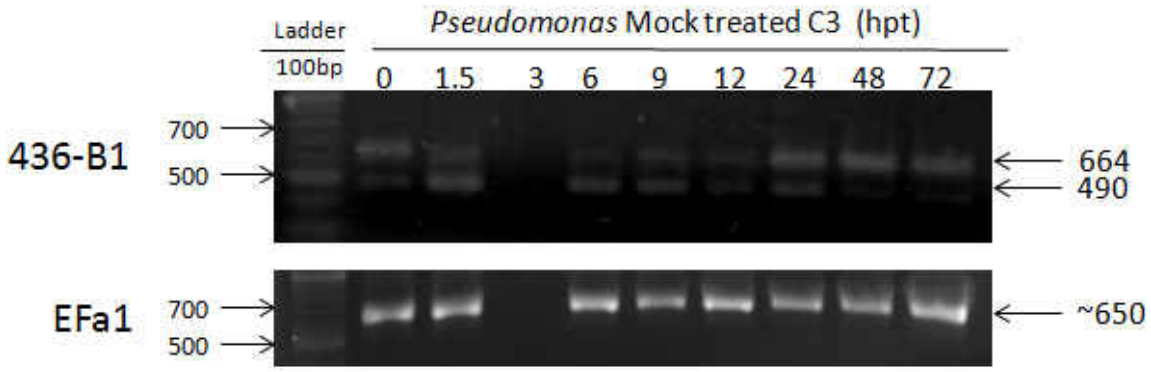


Figure 50: Expression of SBIP-436 in Mock (10mM $MgCl_2$) treated *N. tabacum* C3 plants. Expression primers include: -B1 (DK511-DK514), and *EF±1* (Control). 1.2% Agarose gel showing expression of gene segments of interest.

Expression of 436 in $MgCl_2$ (mock) treated revealed similar results as the mock treatment for TMV inoculations. There was a dominant switch from the 664bp transcript to the 490bp transcript from 0hr to 1.5hr. The expression of both transcripts seems to level out at 6hr, 9hr, and 12hr time points. The 664bp transcript then becomes dominant at the 24hr, 48hr, and 72hr time points. It is important to note that the 24hr, 48hr, and 72hr expression resembles the expression of the 0hr sample.

CHAPTER 4

DISCUSSION

Plants are in a perpetual evolutionary arms race with their environment. Their sessile nature has allowed them to evolve a complex innate immune system that can recognize biotic and abiotic stressors at the cellular level. Upon stress recognition, signal transduction pathways are able to express a tailored defense for a specific threat (Anderson et al., 2011). The SA signal is a robust activator of defense responses in response to pathogen attackers and is integral for the establishment of SAR in uninfected distal tissues of the plant. MeSA, the mobile ester form of SA, is converted back to SA by SABP2. (Kumar and Klessig 2003; Forouhar et al. 2005). In SABP2 silenced 1-2 plants *PR-1*, the defense marker gene representing SAR establishment, was not fully expressed (Kumar and Klessig 2003). Furthermore, when SABP2 silenced plants are treated with acibenzolar-*S*-methyl (ASM), a functional analog of SA, *PR-1* fails to express, while treatment with acibenzolar yields *PR-1* expression comparable to control plants (Tripathi et al. 2010). This makes SABP2 a critical component of the downstream expression of SA mediated defense genes supporting the establishment of SAR. SABP2's central role in the SA signaling pathway presented an opportunity to reveal potential protein members that dictate the SA mediated defense pathway. A yeast-2-hybrid screening was performed to identify possible protein interactions with SABP2. Proteins with positive interaction were identified as SABP2 Interacting Proteins (SBIPs) and among these was SBIP-436. The primary focus of this study was to characterize SBIP-436 and outline its role in the SA mediated defense pathway. Preliminary bioinformatics analysis resulted in 2 important findings. First, that SBIP-436 belonged to the phospholipase D family with high homology to PLD´ and second, that the SBIP-436 yeast-2-hybrid clone sequence was presented as a partial of the full gene. Phospholipase D

proteins are phosphodiesterases that cleave the head group from phospholipids. PLDs are associated with lipid signaling cascades that support plant defense and much like varying phytohormones are predominantly associated with a specific stressor, the separate PLD isoform classes are predominantly associated with a specific phytohormone signaling cascade. Therefore, crosstalk between PLD isoforms may be a revealing layer beneath phytohormone crosstalk. PLD´ is unique to other isoforms in that it is activated by oleic acid (18:1) and is exclusively bound to the plasma membrane (Zhang et al. 2003). Research attempted to establish if SBIP-436 was indeed the *N. tabacum* PLD´. An expression study was performed to determine if SBIP-436 was expressed differentially between biotic and abiotic stressors. This held the possibility of simultaneously implicating its role in the SA signaling pathway and providing further information on the identity of SBIP-436.

SBIP-436 Gene Analysis

The SBIP-436 Y2H clone sequence was subjected to BLAST in the NCBI database and the results revealed that the sequence belong to the PLD family of proteins and represented homology to PLD´ (Figures 5, 6, and 11). SBIP-436 showed 70% similarity with AtPLD´ and 71% similarity with RcPLD´. The alignment also revealed that SBIP-436 was a small partial maintaining the c-terminus of comparable PLD´ sequences. Additionally, it was discovered that the *N. tabacum* PLD´ has not yet been identified nor annotated. Research efforts were then focused on identifying the full length *N. tabacum* PLD´. Using the SBIP-436 Y2H nucleotide sequence, a BLAST was performed in Sol genomics *N. tabacum* Unigene sequence set. This yielded a slightly larger fragment with an exact match to the SBIP-436 sequence known as SGN-U444527 (Figure 12). Because the SGN-U444527 sequence maintained SBIP-436 sequence exactly and was a larger fragment, it was used to search for the full *N. tabacum* PLD´ from that

point on. When no larger segments could be found, the genome crawl was used to walk the AtPLD´ sequence through the (Solanaceae) *S. lycopersicum* database, the SlPLD´ was used to find the *N. benthamiana* PLD´, and the NbPLD´ was to be used to find the full *N. tabacum* PLD´ (NtPLD´). The NtPLD´ results were scattered partials corresponding to various PLDs. It became apparent that the multiple PLD isoforms had highly conserved segment portions that blurred the lines of interpreting their actual identities in another organism. In addition, *N. benthamiana* is diploid while *N. tabacum* is tetraploid. Therefore, sequence hits from the *N. tabacum* genome were higher in number, blurring the lines even further. So, instead of using the genome crawl from the top-down, it was used from the base-up. The main focus being maintain the integrity of the SBIP-436 Y2H clone sequence. If a sequence could be found in *N. benthamiana* corresponding to the SBIP-436 Y2H clone then that resulting sequence could be used to match the true SBIP-436 *N. tabacum* PLD´. The SGN-U444527 BLAST in the Sol Genomics *N.benthamiana* Genome v0.4.4 predicted proteins dataset and revealed a full protein with an identity score of 97% named NbS00023265g0007.1. Again, in the base-up fashion, the NbS00023265g0007.1 (Nb PLD´) was used to BLAST in *S. lycopersicum* dataset ITAG Release 1 predicted proteins (SL1.00) dataset. This revealed a highly similar sequence named SL1.00sc02164_456.1.1 (SlPLD´). Both of these sequences maintained homology with the SBIP-436 Y2H clone sequence (Figures 15, 16, and 17). Nb PLD´ was used to BLAST in *N. tabacum* Unigenes dataset and revealed no significant results. NbPLD´ was then subjected to BLAST in the *N. tabacum* Methylation Filtered Genome TGI: v.1 Contigs. dataset. The ORFs of the genomic results were then analyzed and from these the *N. tabacum* PLD´ (NtPLD´) construct was generated. The c1562 ORF was a good match to the first 302aa of the NbPLD´ sequence (Figure 19). One partial putative NtPLD´ had been annotated in the NCBI database

named GQ904710.1. It was discovered as a 90kD PLD that binds to microtubules (Gardiner et al. 2001). GQ904710.1 overlaps the SGN-U444527 and matches well the NbPLD´ sequence (Figure 19 and Figure 20). The 3 segments were spliced together and trimmed of overlapping portions. These sections when put together resulted in NtPLD´ gene construct (Figure 21). There was a gap from amino acids 303-356. The results from searching for this section showed weak homology and less than desirable overlapping sections. Overall, a putative *N. tabacum* PLD´ gene construct was formed and a putative *N. benthamiana* PLD´ was identified (Figure 15 and Figure 20).

Amplification of SBIP-436

To amplify full length SBIP-436, gene primers and expression primers were constructed using the NtPLD´ construct (Figure 22 and Figure 23). SBIP-436 -A and -B1 were designed to perform an expression analysis of SBIP-436. SBIP-436 -A amplified a segment of 210 bp and -B1 was intended to amplify a segment of 490 bp. -A did amplified its intended 210 bp product (Figure 24). While -B1 did amplify its intended 490bp product there was also an unintended amplification of ~700bp product (Figure 25). Given the highly conserved nature of the PLD c-terminus, were the primers were designed, it was prudent to sequence both amplified products for comparison. The amplified products of -A and -B1 were subjected to gel extraction and purification (Figure 26). The -B1 products failed to yield sufficient amounts of DNA so a direct PCR purification method was implemented. The -A product was cloned into pGEMT vector and the -B1 products were cloned into both pGEMT and TOPO vectors. The plasmid DNA was purified and selected for sequencing (Figure 27, 28, and 29). Figure 30 shows the sequences from all 6 samples. Samples 1 and 2 represent the -A, Samples 3 and 5 represent the 664bp product of -B1, and Samples 4 and 6 represent the 490bp product of -B1. All the sequences

matched well to the SGN-U444527. Samples 3 and 5 reveal a mid-sequence insert of 174bp when aligned to SGN-U444527 (Figure 32 and Figure 35). Samples 4 and 6 show a match to SGN-U444527. Therefore, when all are aligned the 664bp product maintains the 490bp segment with an additional mid-sequence 174bp insert (Figure 34 and Figure 35). Given PLDs nature of multiple isoforms and alternative splicing it was important to identify which was being presented. To analyze this portion an ASSP (Alternative Splice Site Predictor) tool was used courtesy of Wang Computings (http://wangcomputing.com/assp/evaluation.html). The 664bp product represented by Sample 5 was and showed positive results of alternative splicing. Donor and acceptor splice sites were found directly surrounding the 174bp insert (Table 5). The ASSP splice map displays the splice sites along with a decrease in codon usage at splice junctions (Figure 37). Primers were constructed for the 174bp insert (SBIP-436 -B2) to separately view the expression pattern of the single transcript. SBIP-436 -B2 successfully amplified the 174 bp segment (Figure 39). Overall, the analysis of the relationship between the 664bp segment and the 490 bp segment revealed that the separation two transcripts is a product of alternative splicing.

Two sets of primers were constructed to amplify the full NtPLD´. The first, SBIP-436 Full-1 included cloning attb sites while SBIP-436 Full-2 only contained segment nucleotides. In addition, primers were constructed in an attempt to amplify larger portions (1000 bp) of the NtPLD´ construct. They are known as SBIP-436 1kb-A (First 1000bp) and SBIP-436 1kb-B (Second 1000bp). All attempts to amplify the full gene and 1000bp segments yielded negative results. Because there was high homology between the NtPLD´ and NbPLD´ sequences, attempts to amplify the full gene were also attempted using *N. benthamiana* cDNA. Although attempts to amplify the full gene in *N. benthamiana* did not succeed, the previously mentioned expression primers were successful. SBIP-436 -A and -B1 were successfully amplified using *N.*

benthamiana cDNA (Figure 41). It is important to note that both amplified products of -B1 were amplified using *N. benthamiana* cDNA.

Expression Analysis of SBIP-436

An expression analysis was performed to reveal the role of SBIP-436 in plant defense. Expression of the SBIP-436 gene was analyzed in regards to SA mediated defense, predominantly associated with defense against biotic stressors, as well as defense against abiotic stressors. SBIP-436 expression was also examined in each of the *N. tabacum* mutant lines (XNN, NahG, C3, and 1-2). It was important to maintain the relationship of the SBIP-436 EXP sets to each other. The SBIP-436 -B1, as previously stated, amplify 2 transcripts implicated in alternative splicing. The expression of these 2 transcripts was found to be differentially modulated across all stressors. The SBIP-436 -A segment is common to both of the SBIP-436 -B1 transcripts because they share the same forward primer (DK511) and the -A reverse DK513 primer ends prior to the proposed donor splice site (Figure 34). Therefore, any differentiation between -B1 transcripts that presented as an increase in one and a decrease in the other would register no change in the -A expression level. This is quite evident across this expression study. The SBIP-436 -B2 product isolates the 664bp transcript from -B1 by only amplifying the 174bp gene insert.

SBIP-436 expression across the various mutant lines only showed slight deviation in NahG plant. The -B1 664bp transcript is expressed slightly more in NahG plants than in others (Figure 45). Because NahG plants express as SA hydroxylase that converts SA to catechol, SA levels are decreased in these plants. This increase in expression may be due to a change in regulation of 18:1 levels, which activate PLD´. SSI2 is fatty acid desaturase that converts 18:0 to 18:1. Mutants lacking the SSI2 exhibit a drastic increase in SA levels (Kachroo et al. 2005). The

relationship between 18:1 and SA may be the reason for this slight change in expression in the NahG plant.

SBIP-436 expression in the TMV infected XNN demonstrated the effect of SA- mediated defense pathway and stood as a marker for SBIP-436 expression in the presence of a biotic stressor (Figure 46). There was definite modulation between the -B1 transcripts. *PR-1*, marker for activation of the SA pathway, was present at 48hr and increases at the 72 hr. The 664bp transcript expression decrease as it approaches increase in expression of the *PR-1* while the 490bp transcript expression increases. -B2 shows a slight return of the 664bp transcript. Compared to the TMV mock treated expression study from 24hr to 72hr the 664bp transcript is dominant in the mock treated and the 490bp transcript is the dominant transcript in the TMV infected tissue. This demonstrates that the 490 bp transcript is being actively used, meaning that the 664 may be differentially spliced during pathogen attack and may play a role in the activation of the SA pathway. The SA treated XNN plants allows activation of the SA pathway without the pathogen variable. Figure 47 demonstrates similar pattern of expression to the TMV infected expression. There is an immediate decrease in the 664bp transcript and an increase in the 490bp transcript approaching the expression of *PR-1* at 6hr. additionally for 6hr to 24 hr both transcripts tend to balance, but only after the initiation of *PR-1*.

SBIP-436 expression in 300mM NaCl treated XNN shows a drastic decrease in both -B1 transcripts from 0hr to 3 hr (Figure 48). At 6hr the expression increases continuing through 9hr and both transcripts seems to be balanced. At 24 hrs there is seems to be a decrease in 490bp transcript and an increase in the 664bp transcript. Results from treatment with this abiotic stressor reveal a definite decrease in expression between 0hr and 3hr. This is supported by the

matched decrease in expression of -A that maintained regular expression in both TMV and SA treated samples.

Both mock treated expression studies for TMV and *Pseudomonas*, were presented as a source of wounding abiotic stress, which is associated with JA. Both demonstrate modulation of the -B1 transcripts. There was a decrease in expression of the 664bp transcript and an increase in the 490bp transcript from 0hr to 1.5 hr. This could explain the same occurrence in the biotic stressors expression studies. Because PLD´ is involved in membrane restructuring and microtubule reorganization it is possible that wounding is causing the sudden differential expression between the two -B1 transcripts. However, this does explain the drastic decrease in expression of both transcripts in the 300mM NaCl treated XNN nor the dominant presence of the 490bp transcript in the TMV infected XNN. Overall, the2 transcripts are being differentially expressed between abiotic and biotic stressors. These findings may also have implications in crosstalk between SA and JA.

This research was conducted to characterize SBIP-436 and better understand its role in the SA mediated defense pathway in *N. tabacum.* The true *N. tabacum* PLD´ is yet to be identified nor annotated. This study has pieced together a putative *N. tabacum* PLD´ construct and identified a putative *N. benthamiana* PLD´. SBIP-436 sequence analysis revealed that the sequence is undergoes alternative splicing. The expression studies demonstrated that the SBIP-436 is differentially expressed between biotic and abiotic factors and may play a role in crosstalk between SA and JA.

Future Direction

The research presented in this study is only the beginning of this SBIP-436 project. For finding the full corresponding SBIP-436 sequence, a 5' RACE could be performed. To better understand the role of SBIP-436 in the SA mediated defense pathway, it will be practical to generate transgenic tobacco silenced in SBIP-436 expression using the pHANNIBAL (or similar) silencing vector. One could vary sequences to silence various isoforms i.e. isoforms containing 664bp segment or both 490bp and the 664bp segment. Continuing research into SBIP-436 could give insight into lipid signaling in the SA pathway, crosstalk between SA and other hormones, and identification of other *N. tabacum* PLD isoforms.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman, DJ. 1990. "Basic local alignment search tool." J. Mol. Biol. 215:403-410.

An C., and Zhonglin M. 2011. Salicylic Acid and Its Function in Plant Immunity. Journal of Integrative Plant Biology 53(6): 412-28.

Anderson JP, Gleason C, Foley R, Thrall P, Burdon J, Singh K. 2010. Plants versus Pathogens: An Evolutionary Arms Race. Functional Plant Biology 37(6): 499.

Thomma B, Nürnberger T, Joosten M, 2011. Of PAMPs and effectors: The blurred PTI-ETI dichotomy. *Plant Cell.* 23(1): 4-15.

Broekaert, Delaur´e, De Bolle, Cammue1. 2006. The Role of Ethylene in Host-Pathogen Interactions. The Annual Review of Phytopathology. 44:393-416.

Choi J, Choi D, Lee S, Min Ryu C, Hwang I. 2011. Cytokinins and plant immunity: old foes or new friends? Trends in Plant Science, 1360-1385.

Dempsey D, Vlot A, Wildermuth M, and Klessig D. 2011. Salicylic Acid Biosynthesis and Metabolism. The Arabidopsis Book.

Durrant W and Dong X. 2004. Systemic acquired resistance. Annual Review of Phytopathology 42:185-209.

Erb M, Meldau S, Howe G. 2012. Role of phytohormones in insect-specific plant reactions. Trends in Plant Science, 17:5, 1360-1385.

Forouhar F, Yang Y, Kumar D, Chen Y, Fridman E, Park SW, Chiang Y, Acton TB, Montelione GT, Pichersky E, Klessig DF, Tong L. 2005. Crystal structure and biochemical studies identify tobacco SABP2 as methyl salicylate esterase and further implicate it in plant innate immunity. Proceedings of the National Academy of Sciences of the United States of America 102:1773-1778.

Gardiner J, Harper J, Weerakoon N, Collings D, Ritchie S, Gilroy S, Cyr R, Marc J. 2001. A 90-kD Phospholipase D from Tobacco Binds to Microtubules and the Plasma Membrane. The Plant Cell 13(9): 2143.

Grant M and Loake G. 2007. Salicyclic Acid in Plant Defence-the Players and Protagonists. Current Opinion in Plant Biology. 10(5): 466-472.

Guo L, Devaiah S, Narasimhan R, Pan X, Zhang Y, Zhang W, Wang X. 2012. Cytosolic Glyceraldehyde-3-Phosphate Dehydrogenases Interact with Phospholipase D to Transduce Hydrogen Peroxide Signals in the Arabidopsis Response to Stress. The Plant Cell 24(5): 2200-212.

Javid M, Sorooshzadeh A, Moradi F, Sanavy S, Allahdadi I. 2011. The Role of Phytohormones in Alleviating Salt Stress in Crop Plants." Australian Journal of Crop Science 5(6): 726-34.

Jones J and Dangl J. 2006. The Plant Immune System. Nature. 444(16): 323-329.

Kachroo P, Venugopal S, Navarre D, Lapchyk L, Kachroo A. 2005. Role of Salicylic Acid and Fatty Acid Desaturation Pathways in Ssi2-Mediated Signaling. Plant Physiology 139(4): 1717-735.

Kalachova T, Iakovenko O, Kretinin S, Kravets V. 2013. Involvement of phospholipase d and nadph-oxidase in salicylic acid signaling cascade. Plant Physiology and Biochemistry 66: 127-133.

Katagiri T, Takahashi S, Shinozaki K. 2001. Involvement of a novel Arabidopsis phospholipase D, AtPLD´, in Dehydration-inducible Accumulation of Phosphatidic Acid in Stress Signalling. The Plant Journal 26(6): 595-605.

Kobayashi Y, Kobayashi I. 2007. Depolymerization of the actin cytoskeleton induces defense responses in tobacco plants. J Gen Plant Pathol 73: 360–364

Kumar D and Klessig DF. 2003. High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity. PNAS 100(26):16101-6.

Kunkel B and Brooks D. 2002. Cross talk between signaling pathways in Pathogen Defense. Plant Biology Journal 5: 325-331.

Li W, Li M, Zhang W, Welti R, Wang X. 2004.The plasma membrane–bound phospholipase D´ enhances freezing tolerance in Arabidopsis Thaliana. Nature Biotechnology 22(4): 427-33.

Loake G, Grant M. 2007. Salicylic acid in plant defence- the players and protagonists. Current Opinion in Plant Biology 10:366-472

Métraux, Jean-Pierre. 2002. Recent Breakthroughs in the Study of Salicylic Acid Biosynthesis. Trends in Plant Science 7(8): 332-34.

Miura K and Tada Y. 2014. Regulation of water, salinity, and cold stress responses by salicylic acid. Frontiers in Plant Science 54:1-11.

Munnik T, Dhonukshe P, Laxalt A, Goedhart J, Gadella, T. 2003. Phospholipase d activation correlates with microtubule reorganization in living plant cells. The Plant Cell 15: 2666–2679.

Punting CP and Kerr ID. 1996. A novel family of phospholipase D homologues that includes phospholipid synthases and putative endonucleases: identification of duplicated repeats and potential active site residues. Protein Science 5:91

Qin C. 2002. Kinetic analysis of Arabidopsis phospholipase D ´, substrate preference and mechanism of activation by ca2 and phosphatidylinositol 4,5-bisphosphate. Journal of Biological Chemistry 277: 49685-9690.

Qin C and Wang X. 2002. The Arabidopsis phospholipase D family. Characterization of a calcium-independent and phosphatidylcholine-selective PLD zeta 1 with distinct regulatory domains. Plant Physiology 128(3):1057-068.

Rainteau D, Humbert L, Delage E, Vergnolle C, Cantrel C, Maubert MA, Lanfranchi S, Maldiney R, Collin S, Wolf C, Zachowski A, Ruelland E. 2012. Acyl chains of phospholipase D transphosphatidylation products in Arabidopsis cells: A study using multiple reaction monitoring mass spectrometry. Ed. Gustavo Bonaventure. Public Library of Science ONE 7(7): E41985.

Scholthof KBG. 2007. The disease triangle: pathogens, the environment and society Nature Reviews Microbiology 5: 152-156

Shimizu-Sato S, Tanaka M, Mori H. 2009. Auxin–cytokinin interactions in the control of shoot branching. Plant Molecular Biology 69(4): 429-35.

Spoel S and Dong X. 2008. Making sense of hormone crosstalk during plant immune responses. Cell Host & Microbe 3(6): 348-51.

Strawn MA, Marr SK, Inoue K, Inada N, Zubieta C. Wildermuth MC. 2007. Arabidopsis isochorismate synthase functional in pathogen-induced salicylate biosynthesis exhibits properties consistent with a role in diverse stress responses. Journal of Biological Chemistry 282(8):5919-33.

Thomma, BPHJ, Nurnberger T, Joosten MHAJ. 2011. Of PAMPs and effectors: The blurred PTI-ETI dichotomy. The Plant Cell Online 23(1): 4-15.

Triphathi D, Jiang YL, Kumar D. 2010. SABP2, a methyl salicylate esterase is required for the systemic acquired resistance induced by acibenzolar-S-methyl in plants. FEBS Letters 584(15):3456-63.

Tsuda K and Katagiri F. 2010. Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. Current Opinion in Plant Biology 13(4):459-65.

Uraji M, Katagiri T, Okuma E, Ye W, Hossain MA, Masuda C, Miura A, Nakamura Y, Mori IC, Shinozaki K, Murata Y. 2012. Cooperative function of PLD and PLD 1 in abscisic acid-induced stomatal closure in Arabidopsis. Plant Physiology 159(1): 450-60.

Vance DE. 2008. Biochemistry of Lipids, Lipoproteins and Membranes. Amsterdam: Elsevier.

Volt CA, Dempsey DA, Klessig DF. 2009. Salicylic acid, a multifaceted hormone to combat disease. Annual Review of Phytopathology 47:177-206.

Wang C, Zien C, Afitlhile M, Welti R, Hildebrand DF, Wang X. 2000. Involvement of phospholipase D in wound-induced accumulation of jasmonic acid in Arabidopsis. Plant Cell 12:2237–2246.

Wang, X. 2002. Kinetic analysis of arabidopsis phospholipase d delta. The Journal of Biological Chemistry, *277*(51), 49685–49690

Wang, X. 2002. Phospholipase D in hormonal and stress signaling. *Current Opinion in Plant Biology* 5(5): 408-14.

Wick JY. 2012. Aspirin: A history, a love story. The Consultant Pharmacist 27(5): 322-29.

Wildermuth MC. Dewney J., Wu G. Ausubel, Frederick M. 2001. Isochorismate synthase is required to synthesize salicylic acid for plant defence. Nature 414(6863):562-5.

Yang SF, Freer S, Benson AA. 1967. Transphosphatidylation by phospholipase D. The Journal of Biological Chemistry. 242:477-484.

Zhang W, Wang C, Qin C, Wood T, Olafsdottir G, Welti R, Wang X. 2003. The oleate-stimulated phospholipase D, PLD´, and phosphatidic acid decrease H_2O_2-induced cell death in Arabidopsis. Plant Cell 15: 2285–2295

APPENDECIES

APENDIX A – Abbreviations

1-2 - SABP2 - silenced plant (transgenic *N.t*. cv Xanthi nc in which *SABP2* gene expression is silenced by RNA interference- pHANNIBAL.

18:1 – Oleic Acid

aa – Amino Acid

ABA- Abscisic Acid

AD - Activating Domain

ASSP - Alternative Splice Site Predictor

At- *Arabidopsis thaliana*

BA2H - Benzoic-2-hydroxylase

BD - Binding Domain

BLAST - Basic Local Alignment Search Tool

bp – Base Pair

C3 - Control plant (*Nicotiana tabacum* cv Xanthi nc, and contains empty pHANNIBAL silencing vector.

EFalpha1 - Elongation Factor alpha 1

GA- Gibberellic Acid

HR - Hypersensitive response

IAA- Indole-3-Acetic Acid

ICS - Isochorismate synthase

IPL - Isopyruvate lyase

ISR - Induced systemic resistance

JA - Jasmonic acid

kb - Kilobase

KBM - King's B Medium

KDa - Kilo Dalton

KO – Knockout

MeSA - Methyl salicylate

ml - milli litre

mM - milli Molar

NahG - Plant expressing salicylate hydroxylase which converts SA to catechol.

NO - Nitric oxide

NPR1 - Non-expresser of pathogenesis-related protein 1

nt – Nucleotide

OD - Optical Density

OE - Overexpressed

ORF - Open Reading Frame

PAL ICS IPL

PAMPs - Pathogen-associated molecular patterns

PA-Phosphatidic Acid

PC - Phosphatidylcholine

PCD - Program cell death

PE - Phosphatidylethanolamine

PG - Phosphatidylglycerol

PIP_2 - Phosphatidylinositol 4,5-bisphosphate

PLD- Phospholipase D

PR - Pathogenesis-related

PRR PTI ETI ETS HR

PRRs - Pattern recognition receptors

PS - Phosphatidyl-Serine

Psp – Pseudomonas syringae phaseolicola

Pst - Pseudomonas syringae tabaci

R protein - Resistance protein

Rc-*Ricinus communis*

ROI - Reactive oxygen intermediates

Rpm – Revolutions per Minute

RT-PCR – Reverse Transcriptase Polymerase Chain Reaction

SA - Salicylic acid

SABP2 - Salicylic acid binding protein 2

SAMT - Salicylic acid methyl transferase

SAR - Systemic acquired resistance

SBIP-436– Salicylic Acid Binding Protein 2 Interacting Protein-436

Ta – Annealing Temperature

TAE - Tris-Acetate EDTA

Tm – Melting Temperature

TMV - Tobacco mosaic virus

UV - Ultra violet

Y2H- Yeast Two-Hybrid

µg - micro gram

µl - micro litre

APPENDIX B – Media and Other Chemicals

0.1% Diethyl Pyrocarbonate (DEPC) Treated Water

Diethyl pyrocarbonate = 0.1 ml

Add to 100 ml distilled water

Incubate for overnight at 37°C

Autoclave for 20 minutes

Rifampicin (14 mg/ml)

Rifampicin (powder) = 0.14 g

Add to 10 ml of Methanol

Add to King's B media - 25 µg/ml

1M Magnesium Sulphate

MgSO4 = 246.48 g

Adjust volume to 1 liter with distilled water

10 mM Magnesium Chloride

$MgCl_2$ = 0.952 g

Adjust volume to 1 liter with distilled water

King's B Medium

Protease peptone # 3 = 20 g

Potassium phosphate dibasic = 1.50 g

Magnesium sulfate = 1.50 g

Glycerol = 10 ml

Adjust volume to 1 liter with distilled water

Adjust pH to 7.0

Agar = 17.50 g

Autoclave for 30 minutes

Place in 55^{o} water bath until ready for use

King's B Liquid Broth

Protease peptone # 3 = 20 g

Potassium phosphate dibasic = 1.50 g

Glycerol = 10 ml

Adjust volume to 1 liter with distilled water

Adjust pH to 7.0

Autoclave for 30 minutes before use

VITA
PHILLIP DEAN

Education:	Master of Science in Biology, 2014 East Tennessee State University, Johnson City, TN
	Bachelor of Science in Biology & Environmental Science, 2010 Tusculum College, Greeneville, TN
Professional Experience:	Graduate Assistant, East Tennessee State University, Tennessee, Department of Biological Sciences, 2012-2014.
Award:	2010 Curtis Owens Literary Prize for Poetry
Conferences:	Phillip Dean and Dhirendra Kumar. Characterization of a Putative Phospholipase D ´ Like Gene as a Lipid Signaling Modulator and Its Role in Salicylic Acid Mediated Defense Pathway in *Nicotiana tabacum*. Oral presentation at *Appalachian Student Research Forum* 2014, East Tennessee State University, Johnson City, Tennessee.
	Phillip Dean and Dhirendra Kumar. Characterization of a Putative Phospholipase D ´ Like Gene as a Lipid Signaling Modulator and Its Role in Salicylic Acid Mediated Defense Pathway in *Nicotiana tabacum*. Poster Presentation at *Appalachian Student Research Forum* 2013, East Tennessee State University, Johnson City, Tennessee.