ABSTRACT

Establishing the Foundation of *Impatiens walleriana* as a Nectar Model System

Andrew M. Cox, M.S.

Mentor: Christopher Kearney, Ph.D.

Rapid proliferation of mosquito-vectored viruses require affordable and effective methods are necessary in poor, urbanized tropical regions. Designing a plant-based drug-delivery system would provide this technology. *Impatiens walleriana* is ideal to establish a nectar-model system for testing drug-delivery targeting mosquitoes. Detailed in this thesis, are three building blocks for engineering impatiens to combat mosquito-borne diseases. First, a highly produced nectar protein was identified, iwPHYL21. It is highly expressed, antimicrobial, and may serve as a fusion partner in heterologous protein expression. Second, an impatiens nectar promoter was identified, which may optimize heterologous protein expression in nectar. Finally, promoters from *Arabidopsis* were utilized to express the marker protein GUS in nectaries and nectar, demonstrating the potential for impatiens to deliver toxins to insects. This work will serve to increase the efficiency and utility of the impatiens model-system, bringing us closer to effective, non-pesticide-based control of mosquito-transmitted diseases in the field.

Establishing the Foundations of *Impatiens walleriana* as a Nectar Model System

by

Andrew M. Cox, B.S.

A Thesis

Approved by the Department of Biology

_____

Dwayne D. Simmons, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Master of Science

Approved by the Thesis Committee

_____

Chris Kearney, Ph.D., Chairperson

_____

Cheolho Sim, Ph.D.

_____

Sung Joon Kim, Ph.D.

Accepted by the Graduate School

December 2016

_____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge the guidance and support received from Dr. Chris Kearney while working on this project. His expertise in the skill of communicating complex ideas both clearly and powerfully has made a profound impact on me, and has greatly helped to shape this work. He has a passion for using science to create positive change in the world, and that is the essence of how this project came into existence. In addition, the level of kindness he treats his students and colleagues with will not be forgotten.

I would also like to acknowledge all of my colleagues in the Kearney lab who have truly helped me become a better scientist, and who also made this a very fun experience. Specifically I want to thank Meron for his humor and willingness to listen to thoughts about both this project and life in general. Mishu for his kindness; he was always bringing back souvenirs from his trips for the entire lab. Tommy for his guidance as senior lab member. He taught me how create good work through his efficiency and balanced attitude. Ankan for his wit and ability to make any topic of discussion lively and interesting. I will never forget your recognition of Bonnie M Gold's "Rasputin" from a Bollywood spin-off. Luke, Marissa, Heather, and Thuy were all joys to work with. Thank you for your work on this project and I wish you the best in all you pursue. And last but not least, Grace for your impressive work ethic and ability to treat every day with a fresh attitude. You always looked out for the people in the lab—remembering birthdays and special events, caring for our

cleanliness and safety, and caring for nearly everyone benthamianas. I am hopeful for the future of this work with it in your hands.

I would also like to acknowledge the love and support from friends and family that kept me motivated when progress was slow and celebrated with me with every victory. In particular, I want to thank my parents for their love and encouragement. I know that your nurturing spirits and guidance are gifts that will continue to bless me the rest of my life. Thank you for always being me biggest supporters and helping me discover my potential.

To my wife Emily, thank you for walking with me every step of the way. Through the years we grown so close and I truly cannot imagine my life without you. I am so thankful that the end of this work also brought with it an exciting beginning, our son Jacob. I cannot wait to forge our identity as a family of three. Thank you also for being a steady voice to my sometimes-overzealous enthusiasm. We go together like peanut butter and jelly (an expression I would not appreciate if not for you), and I cannot wait to see what else we can accomplish together!

To my in-laws: Charlie, Sharon, Sarah, Jordan, and Rachel, I thankful for getting to know each of you better through the opportunity of working in the same city as all of you. Your support of Emily and I has been a great gift, and has made my transition into adulthood smoother. I also want to acknowledge Jordan's amazing skills as an editor and writer. I appreciate all of the times you would help take my writing to the next level.

I want to acknowledge Dr. Alejandro Ramirez for his willingness to teach a Biologist mass spectrometry, and his patience as I learned some tricks of the trade. I

want to acknowledge Brett Harper for his openness to providing guidance for mass spectrometry experiments, and for introducing me to Alejandro. I want to acknowledge Dr. Clay Carter at the University of Minnesota for significant support on this project. He provided all the *Arabidopsis* nectar promoter vectors that were used in this study, as well as providing RNA sequencing *en gratis*. I want to acknowledge the collaboration with Dr. Yinghui Dan at Virginia Polytechnic Institute for her work transforming the impatiens, and her care shipping them to us across the country in the heat of summer.

DEDICATION

To the millions of people impacted by mosquito-transmitted diseases

CHAPTER ONE

Introduction and Background

Nectar Background, Genetic Engineering of Nectar

Nectar Introduction

Composition and purpose of nectar. Nectar is an aqueous solution, comprised primarily of carbohydrates, in the form of monosaccharides and disaccharides mostly, as well as other solutes including proteins, amino acids, terpenes, sugar esters, polyphenols (Bentley and Elias 1983; Kevan and Baker 1983). The type of sugar present in nectar varies; sucrose, glucose and fructose are predominant (Bentley and Elias 1983; Roshchina and Roshchina 2012). The type of sugars present in nectar can influence interactions with arthropods, as in the case of the pollinators that are attracted and with symbiotic insects(Perret 2001; Wolff 2006). For example, post-secretory hydrolysis of sucrose in *Acacia* discourages non-mutualistic ants from feeding on extrafloral nectar (Heil et al. 2005). Insects can visualize hexose sugars, like sucrose and glucose, through fluorescence from ultraviolet light (Silberglied 1979; Brewster 1994), thus potentially explaining the selective mechanism for effective pollinator attraction. Sucrose-rich nectar plants are more often pollinated by bees, moths, butterflies and hummingbirds as opposed to hexose-rich nectar plants which are more often pollinated by flies, small bees, passerine birds and neotropical bats (Kevan and Baker 1983; Baker and Baker 1990; Wunnachit et al. 1992). Originally, nectar was thought to serve a sole purpose

of enhancing a plant species' fitness through attracting pollinators to increase genetic diversity (Goulson 1999; Rudgers and Gardener 2004). Nectar does indeed serve this purpose, but less obviously extrafloral nectar is known to encourage mutualistic insect species to visit or inhabit the plant (Rudgers and Gardener 2004; González-Teuber et al. 2009a; Escalante-Pérez et al. 2012). It has been observed that extrafloral nectar can facilitate a mutualistic relationship with insects that protect the plant, as in myrmecophytes (Bronstein 1998). Amino acid composition varies in type and concentration, which is thought to affect the type of pollinator attracted to the nectar as well (Gardener and Gillman 2001).

Drawing from the limited but converging evidence of nectar proteomes, there is very limited variety in the nectar proteins occupying nectar of a single species when compared to protein variety in non-vascularized tissue, and they usually range from 10 kDa-70 kDa on an SDS-PAGE gel (Peumans et al. 1997a; Shepherd 2005; Kram et al. 2008; González-Teuber et al. 2009b; Hillwig et al. 2010; Hillwig et al. 2011; Nepi et al. 2012; Zha et al. 2012; Seo et al. 2013; Zha et al. 2013; Chen and Kearney 2015a). These studies include plant species from the genera *Nicotiana, Acacia, Petunia, Jacaranda, Allium, Mucana, Acacia, Impatiens, Ricinus, Campsis, Cucurbita,* and *Passiflora.* The range of concentration of proteins in nectar varies, although there have only been limited studies analyzing nectar protein content. The collective work of Carter, Graham and Thornburg characterized the floral nectarome of *Nicotiana spp.* and found total floral nectar protein concentration to be 240 μg/ml (Carter et al. 1999; Carter and Thornburg 2003; C.J. Carter and Thornburg 2004). The first study published covering nectar protein

analysis determined the concentration of floral nectar leek proteins to be 220 μg/ml (Peumans et al. 1997b). The woody vine *Mucana sempervirens* has a calculated nectar protein concentration of 500 μg/mL(Zha et al. 2013). Chen and Kearney (2015) analyzed extrafloral nectar protein via SDS-PAGE for *Impatiens walleriana, Ricinus communis, Campsis radicans, Passiflora edulis* and *Nicotiana tabacum* and calculated ranges of total protein concentration in nectar to be between 29 μg/ml–4.67 mg/ml. This study revealed staggering differences in extrafloral nectar protein amounts where *I. walleriana* had over 8x greater total nectar protein and individual protein concentration for the 21 kDa protein than the plant with the second most.

Nectar Protein Function. Research on the function of nectar proteins is relatively limited. Table 1-1 lists nectar proteins with experimentally verified function or activity. The results and conclusions of the studies involving the nectar proteins shown in Table 1.1, suggest that nectar proteins either function in defending the plant from the invading pathogens, bacterial or fungal, or regulate the sugar environment of the nectar for the end purpose of influencing insect-plant interactions or possibly regulating nectar secretion/resorption. For defense, some species appear to enlist a redox cycle in their nectar to protect from infection (C. Carter and Thornburg 2004). Plants like tobacco, leek and *Acacia* may defend against microbial intruders through generating and degrading the highly reactive superoxide free radical. Other nectar proteins like chitinases, glucanases, and phylloplanin simply inhibit fungal infections directly(Shepherd 2005; González-Teuber et al. 2009b).

Table 1.1. Experimentally Verified Nectar Proteins and Their Function

Plant	Protein Name	Protein Activity	Location of Protein
Tobacco	Nectarin I	-Superoxide dismutase	Floral nectar (Carter et al. 1999; Carter and Thornburg 2000)
	Nectarin II	-Possible breakdown product of Nec3	Floral nectar (Park and Thornburg 2009)
	Nectarin III	-Carbonate anhydrase -Monodehydroasacorbate reductase	Floral nectar (Carter and Thornburg)
	Nectarin IV	-Endoglucanase inhibitor	Floral nectar (Naqvi 2005)
	Nectarin V	-Glucose oxidase	Floral nectar (C.J. Carter and Thornburg 2004)
	NTα-Gal	-Acidic α-galactosidase	Floral nectar (Zha et al. 2012)
	T-phylloplanin	-Anti-fungal	Trichome exudate (Shepherd 2005; King 2011)
Leek	No specific proteins assigned activity	-Agglutination of mannose -Aliinase	Floral nectar (Peumans et al. 1997b)
Petunia	Multiple unnamed proteins (Peroxidase and ribonuclease may be same protein)	-Peroxidase -Ribonuclease -Endochitinase -Fructokinase	Floral nectar (Hillwig et al. 2011)
Acacia	Pathogenesis-related Proteins	-Chitinase -β-1,3-glucanase -Peroxidase	Extrafloral nectar (González-Teuber et al. 2009b)
Mucana	MS-desi	-Citrate synthase inhibitor	Floral nectar (Zha et al. 2013)
Pumpkin	Xylosidase	-β-d-xylosidase	Floral nectar (Nepi et al. 2011)
Jacaranda	JNP1	-GDSL lipase/esterase	Floral nectar (Kram et al. 2008)

With such limited amount of nectar proteins classified it is difficult to know what type of method a plant may employ to defend itself. As mentioned previously, sugar composition in nectar has large effects on the type of insects attracted. Therefore, plants may have evolved particular methods of attracting specific insects through modulating nectar metabolite composition and concentration with secreted proteins. There does appear to exist a distinct difference in nectar proteins between monocots and eudicots. A recent hypothesis proposed (Lin et al. 2014) describes this difference spawning from the fact that monocots are highly-specialized for wind pollination whereas eudicots utilize pollination more from insects.

Current Art Foundational for Utilizing Transgenic Nectar-Nectary Promoter Elements and Nectar Protein Signal Peptides

Transgenesis of Nectar. The production of transgenic nectar proteins has been demonstrated as a possibility by one published study (Helsper et al. 2011) in which human epidermal growth factor (hEGF) was produced in tobacco nectar. Presently, the author is not aware of any other instance of the production of transgenic nectar proteins in the literature, although Dr. Robert Thornburg at Iowa State University has unpublished data of GFP expression in tobacco nectar.

Fundamentally, transgenesis of nectar proteins requires knowledge of a nectary-specific promoter sequence and a signal peptide that induces secretion of the transgene into nectar. While in theory this approach should result in successful production and localization of the transgene into nectar, there exist some caveats. Very few genes involved in the formation and secretion of nectar proteins have been identified and characterized. Thus, unknown protein-protein interactions could

prove a significant limitation in the secretion of foreign proteins highly dissimilar from native nectary proteins. In addition, what limited research there is on nectar proteins does not compare possible differences between floral and extrafloral nectar proteins. Different molecular chaperones for nectar could be highly species specific, dependent on the co-evolved symbiotic organisms. Until there is more clarity the production of transgenic nectar proteins may encounter unforeseen obstacles. In spite of the limited research on nectar proteins, there has been significant progress in the past decade towards understanding both the physiology of nectar production and secretion.

Nectary formation and nectar production and secretion. One of the earliest known genes necessary for nectary formation is CRABS CLAW (CRC). CRC is a nectary-specific transcription factor that evolved with the formation of the core eudicots, rosids and asterids, around 120 million years ago (Lee et al. 2005; Bell et al. 2010). Knockdown of CRC in *Arabidopsis* leads to the absence of both median and lateral floral nectary formation (Lee et al. 2005). A pivotal study for understanding the biology of nectary formation and the genes involved in nectar production and secretion, led by Dr. Clay Carter, identified 270 upregulated genes in floral nectary tissue in *Arabidopsis thaliana,* as well as two consensus promoter element sequences specific for upregulated nectary genes (Kram et al. 2009). Further studies by Carter regarding several of the upregulated nectary genes advanced insight into the formation and secretion of nectar in nectaries. CELL WALL INVERTASE 4 (atCWINV4) is a protein in *Arabidopsis* that hydrolyzes sucrose to fructose and glucose and is highly upregulated in floral nectary tissue. Without

atCWINV4 there is very little invertase activity, and the necessary hexose sugars for nectar formation and secretion are not produced (Ruhlmann et al. 2010). To elaborate, the hexose sugars produced from atCWINV4 create a high osmotic concentration gradient that pulls nectar from parenchymal cells (Lin et al. 2014). Without this sink, nectar secretion is significantly inhibited (Ruhlmann et al. 2010). In addition, the transmembrane sucrose transporter SWEET9, in *Arabipdopsis* (atSWEET9), has been identified as a gene upregulated heavily in nectary tissue and functions in plasma membranes as a sucrose transporter (Lin et al. 2014). It is currently understood that atSWEET9 is present in nectaries of both rosids and asterids, and is thought to have co-evolved somewhere alongside CRC in core eudicots (Lee et al. 2005; Lin et al. 2014: 9). Utilizing the promoter elements for atCRC, atCWINV4, and atSWEET9, several independent studies have produced the transgenes eGFP and GUS in floral nectary tissue of both rosids and asterids (Lee et al. 2005; Ruhlmann et al. 2010; Lin et al. 2014). Specifically, GUS and eGFP were produced in nectaries of the rosid species *Arabidopsis thaliana, Brassica rapa* and asterid species *Nicotiana attenuata* using native SWEET9 promoters for each (Lin et al. 2014). GUS was produced in the nectaries of the rosid species *A. thaliana* using a native CRC promoter and in an independent experiment a CRC promoter from *Lepidium africanum*, which is also a rosid (Lee et al. 2005). Additionally, GUS was produced in the nectary of the asterid species *Nicotiana tabacum* using the CRC promoter from *A. thaliana.* Lastly, GFP was produced in nectaries of the *A. thaliana* using the native promoter element for CWINV4 (Ruhlmann et al. 2010). These

experiments demonstrate a flexibility for floral nectary promoters to induce nectary-specific transcription between rosids and asterids.

Nectar protein signal peptides. Quite apparently, nectar proteins possess a trafficking signal that acts to facilitate transport from the translation in the endoplasmic reticulum to exterior of the nectary cells and into nectar. In plants, as well as animals and yeast, it is understood that secretory signals exist as amino-terminus peptide sequences, translated in the open reading frame with the rest of the protein (Hadlington and Denecke 2000; von Heijne 2001). These signal peptides are short length (~5-30) amino acid sequences that recruit chaperone proteins which function to facilitate transport to the plant cell vacuole or plasma membrane (Hadlington and Denecke 2000; von Heijne 2001). At some point in the process, the signal peptide is cleaved by a protease interacting with a recognition sequence near the carboxy-terminus of the signal peptide. What is pertinent, is the fact that for nectar proteins, an N-terminal peptide sequence induces a secretory process (Helsper et al. 2011). Whether a nectar protein signal peptide is necessary for secretion, or if any secretory signal peptide can induce localization of proteins into nectar, is not yet known.

The only prior published example of transgenic nectar. While experiments involving expression of transgenes in the nectary are becoming more numerous, heterologous protein secretion into nectar has only been demonstrated in published literature once (Helsper et al. 2011). This study used the N-terminal signal peptide and promoter corresponding to a nectar protein, CARN2, from carnation. This signal

peptide was fused to human epidermal growth factor (hEGF), which resulted in hEGF being secreted into to the nectar of *Nicotiana langsdorffii* X *N. sanderae* hybrid plants (Picard-Nizou et al. 1997). Carnation is phylogenetically classified as a core eudicot, and evolved soon after the formation of eudicots. The tobacco plant that expressed hEGF was not a core eudicot, but rather a more evolved eudicot, belonging to the asterid clade. Interestingly, this experiment demonstrated that secretion of a transgene into nectar was possible using promoter elements and signal peptides from different clades in eudicots.

The second part of the study mentioned above, involved offering a floral nectar solution including hEGF to bees. The bees could physiologically not reach the floral nectar in the long narrow flower petals. Thus a floral nectar solution, including the same amount of hEGF expressed in the nectar of the tobacco plants, was fed to bees. Expression of hEGF was sufficient enough that bees feeding on the transgenic nectar produced honey with detectable amounts of hEGF, as demonstrated by immunochemistry (Helsper et al. 2011). Equally as important, is the fact that the bees were seemingly unperturbed by the presence of hEGF in a nectar solution and drank enough nectar to produce honey. This suggests that foreign proteins in nectar do not automatically deter insect pollinators in the field.

Impatiens walleriana *as a Model Nectar Drug Delivery Plant*

Mosquitoes and Genetically Modified Nectar

Mosquitoes and their impact on human health. The animal that is most harmful to human life is the mosquito. The diseases that mosquitoes carry kill more people

each year than any other animal, including humans themselves. Estimates from the World Health Organization in 1996 reported mosquito-caused deaths to range from one million to several million per year, with many millions more suffering extreme pain, job loss and incapacitation due to acute and chronic illnesses (http://www-.who.int). Since that time, efforts from governments and organization like the Gates Foundation have helped to drastically reduce the number of mosquito-caused deaths and the cases of mosquito-transmitted diseases like malaria (Organization and others 2016). Even with this significant progress, deaths from mosquito-caused diseases still are estimated to be between 725,000 and 1 million each year (Gates; Mosquito.org). Nearly half of the population of Earth live in an area the puts them at risk of catching a mosquito-borne illness (Organization and others 2016). Vector-spread diseases have been especially difficult to eradicate due to the large population of insect reservoirs, as demonstrated by the thwarted efforts to eliminate malaria in Brazil (Deane 1988). Malaria has been eliminated in many areas of the world, particularly the temperate and Europe; however the success of eradication is largely credited to the widespread use of insecticide, namely DDT, a by-product of a healthy economy to control the large insect populations in urban areas (Gallup and Sachs 2001).

Case study of malaria eradication. Malaria was endemic in the early to mid-20th century in the southern United States. However, after only a decade of nation-wide mobilization against mosquitoes, malaria was declared eradicated from the United States in 1951(https://www.cdc.gov/malaria/about/history/elimination-_us.html). *Anopheles,* the genus of mosquito containing species, such as *A.*

quadrimaculatus, that transmits *Plasmodium vivax,* are still present in the United States. According to the CDC's website, the reason malaria became eradicated was due to the fact that the population of *Anopheles* was reduced in areas near human dwelling to such an extent that *Plasmodium* could not enter humans to undergo massive proliferation. The funding and organization the CDC had from the United States' government was crucial in the elimination of urban or semi-urban mosquito populations. They manufactured and distributed 4.5 million home insecticide sprays, routinely cleared standing water and instructed citizens to do the same, and even dropped insecticide from aircraft over areas with very high density of *Anopheles* (www.cdc.gov). The highly organized and well-funded efforts of the CDC and the United States population made ending malaria transmission possible without a vaccine. This example demonstrates that elimination of mosquito-borne diseases using insecticide requires a level of funding and organization not present in many areas plagued by mosquito-borne disease presently. It is difficult to tabulate the fiscal investment the of manufacturing nearly 5 million home DDT sprays in 1950 in the United States due to the ambiguous original manufacturing cost, inflation and scarce data on the topic. However, today the United States spends approximately $10 billion, one-fourth of the amount spent by the entire world, for about 500 million kg of insecticide, mainly for use in agriculture and prevention of insects in urbanized areas (Pimentel 2005). Hence, the cost of controlling mosquitoes in third-world countries via liberal insecticide application is not realistic. Additionally, the negative public health and environmental impact from loosely-restricted use of insecticides are well-documented for causing poisonings in

humans, domestic animals and livestock, and inducing carcinogenic effects in animals (Pimentel 2005). Taken altogether, even though insecticides have eliminated mosquito-borne disease from affluent areas of the world, the mounting evidence of the negative impact to ecosystems, and the tremendous human and fiscal cost of sustaining such efforts, are not well-suited options for combating mosquitoes in the highly-biodiverse, economically-poor areas of the world today.

Vaccines and other methods of mosquito control. Evidence from the elimination of other diseases shows that vaccines are effective and affordable for providing protection against diseases. Approved vaccines for humans exist for only a limited number of mosquito-borne diseases, such as Japanese encephalitis (Oya 1988; Hennessy et al. 1996) and Yellow fever (Gaucher et al. 2008). Many other mosquito-borne diseases still lack a viable vaccine in humans, such as malaria, West Nile virus, St. Louis encephalitis, Zika virus, Dengue virus, and Chikungunya. There are hopeful outcomes for the development of vaccines for West Nile virus (Monath et al. 2006; Wiwanitkit 2007) and St. Louis encephalitis (Phillpotts et al.), however pathogen-induced mosquito-transmitted diseases like malaria and Chikungunya do not have viable vaccines at present. As a result, different approaches have been researched and implemented to end devastating mosquito-borne diseases, including: bed nets, toxic sugar bait traps (Müller et al. 2008; Müller and Schlein 2008; Müller et al. 2010; Beier et al. 2012), transgenic mosquitoes (Ito et al. 2002; Kang et al. 2016), insecticides, mosquito-repellent clothing, and even mosquito-tracking lasers (Keller et al. 2016).

As mentioned above, Müller and colleagues recently investigated the possibility of using attractive toxic sugar bait (ATSB) in the control of mosquitoes. In these studies ATSB is mixed using a solution of over mostly-ripe fruit juice (75 %), red wine (5 %), brown sugar (5 %-20 %), BaitStab ™ (1 % of antifungal and antibacterial additives), and boric acid (1 %). ATSB has shown to have a significant impact on *Anopheles* populations, even when they are presented with other commonly-visited food sources, possessing a knockdown rate of nearly 100 % (Beier et al. 2012)

Controlling mosquito populations with nectar. The data presented in this thesis pertains to a different approach for controlling mosquito-borne diseases. Specifically, this approach utilizes the conserved behavioral trait of mosquitoes imbibing nectar meals, a trait known to be present in all mosquitoes (Foster 1995), to attract and kill local mosquito populations. Male mosquitoes never feed on human blood but rather rely energy from nectar and other sugar sources such as honeydew (Wa and Rg 1994; Foster 1995). Female mosquitoes of the genera, *Culex, Aedes* and *Anopheles* all have been recorded as to feeding on nectar, although geographical, seasonal and special variations affect nectar-feeding frequency (Müller and Schlein 2005).

Using plant attraction to lure mosquitoes to a toxic bait is not a novel idea. Müller has demonstrated the attraction of *Anopheles sergenttii*, a carrier of *Plasmodium falciparum*, to *Acacia raddiana* nectar, and, through the manual application of insecticide to the nectar, demonstrated knockdown of nearly the entire local population (Müller and Schlein 2006). In a similar experiment, *Culex*

pipiens' attraction to 26 common plant species in the Mediterranean was recorded (Schlein and Müller 2008). Nearly 60 % of mosquitoes were attracted to flowers of *Tamarix jordanis* Boiss, a tree native to Jordan River area. Knockdown of nearly 90% of *Culex* the surrounding population was achieved through insecticide nectar application to all 26 plant species.

These experiments demonstrate that manual application of insecticide to nectar sources can control mosquito populations, however there might exist less labor-intensive, more effective application of the biological relationship between mosquitoes and nectar. As demonstrated through the uptake of hEGF-rich nectar by bees, it is unlikely that nectar proteins influence pollinators' attraction to feeding on nectar. Therefore, changing the specific nectar proteins through constructing and growing genetically modified plants could serve as a drug-delivery system to insects that feed on nectar, namely the mosquito. While there has been extensive observation of mosquitoes preferential attraction to the nectar of varying plant species (Hocking 1968; Müller and Schlein 2005; Müller and Schlein 2006; Müller et al. 2010; Gu et al. 2011), there has only been one study to examine the attraction of mosquitoes to plants that produce nectar that readily lend themselves to genetic transformation (Chen and Kearney 2015).

Study analyzing attraction of mosquitoes to nectar-producing plants that are also readily transformable. This a study compared four nectar-producing plants that are known from the literature to attract mosquitoes with nectar and be readily-transformable with *Agrobacterium,* The four nectar plants were chosen from a field of 37 candidate plants which were screened for mosquito attractiveness, nectar production and nectar

protein production. *Impatiens walleriana* was shown to sustain mosquito populations and be preferentially fed upon by mosquitoes better than the three other top nectar plants (Chen and Kearney 2015). Chen and Kearney demonstrated that *Aedes aegypti* populations feeding on impatiens nectar live significantly longer on impatiens than any other sugar source tested. Other top nectar-producing plants tested were *Passiflora edulis* (passion flower)*, Asclepias curassavica* (milkweed) *,* and *Ricinus communis* (castor bean), with *Nicotiana benthamiana* (tobacco) and *Beta vulgaris* (beet) serving as a negative controls due to inaccessible nectar or no flowers, respectively. In direct competition studies between top nectar plants, mosquitoes preferred feeding on impatiens nectar over nectar from *Campsis radicans,* and *Ricinus communis* (Chen and Kearney 2015). As an experimental design, either 20 *Aedes* or 20 *Culex* mosquitoes were placed in a microcosm environment with all three plants, with one them having its nectar dyed red with food coloring, for a total of three days. At the end of each day the percentage of mosquitoes dyed red was recorded. Impatiens dyed mosquitoes via imbibition at least two times better than *Campsis* and *Ricinus* for each day of the trial for both *Aedes* and *Culex* mosquitoes. The preference for impatiens was measured in two ways. First, mosquitoes live significantly longer on impatiens than the other plant sugar sources, demonstrating long term feeding preference. Second, mosquitoes prefer to feed on impatiens nectar over other sugar sources under direct competition. This implies that uptake of the heterologous toxin peptide into a mosquito population will be greater using impatiens than other sugar sources, perhaps even sugar-bait traps, while not requiring a regular government-sponsored spray program and avoiding pesticides altogether. If a mosquito-specific toxin were to be expressed in impatiens, a further benefit would be the avoidance of the

destruction of non-target species, currently an unavoidable fallout of pesticide spray programs

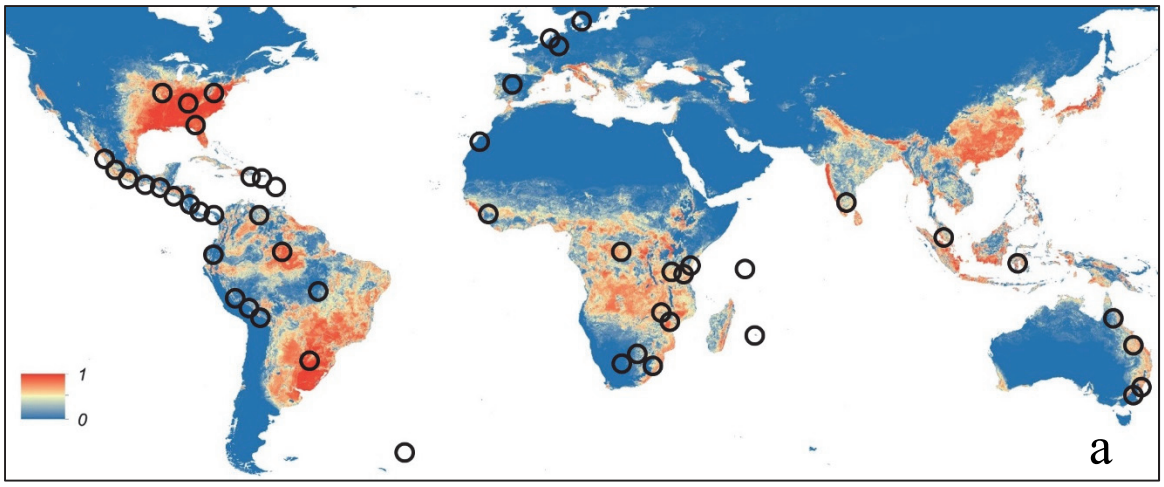
Intro to Impatiens

Impatiens, a prodigious nectar-producing plant belonging to the Balsam family, members of the asterids clade of eudicots, possesses qualities that make it a highly attractive plant for developing a nectar model system to attract and deliver toxins to mosquitoes. First, it is readily transformed using *A. tumefaciens*, with a protocol for transformation already in place (Dan et al. 2010), and was demonstrated as being able to strongly attract mosquitoes (Chen and Kearney 2015). Next, the native growth patterns of *I. walleriana* and *Impatiens spp.* are tropical—originating in East Africa—and spans parts of North, Central and South America, Africa, Asia, and Australia, and strongly correlate with the natural habitats of mosquitoes. As demonstrated by Figure 1-1 the natural sites of *Impatiens* overlaps many of the locations of *Aedes aegypti* and *Aedes albopictus*. The natural environment of impatiens matches well with mosquito populations, implying a wide range of areas that impatiens can be used to bait and drug mosquitoes. Thus, the specific modification of nectar content in *I. walleriana* could prevent the spread of mosquito-vectored diseases in North, Central and South America, Africa and South East Asia.

Another notable feature of impatiens is its remarkable quantity of nectar protein. As mentioned previously, impatiens total nectar protein concentration exceeds every other recorded amount. Additionally, Chen and Kearney discovered that a single-sized protein band accounted for roughly 70% of the total nectar protein amount. In this study, the identity of this nectar protein was discovered.

Phylloplanins

The major nectar protein, migrating at approximately 21 kDa on SDS-PAGE, is a homolog of the previously characterized *Nicotiana tabacum* protein, T-phylloplanin. T-phylloplanin was initially characterized from *N. tabacum* on the surface of leaves (Shepherd 2005; Shepherd 2005). They are secreted by short glandular trichomes and act as an anti-fungal protein, protecting the plant from fungal sporulation (Shepherd 2005). Phylloplanins are characterized by a hydrophobic side and a polar side, similar to other antimicrobial peptides (Lee et al. 2011). Transgenic tobacco plants expressing GUS or GFP fused to the phylloplanin promoter and signal peptide show exclusive expression in a special short glandular trichome population on tobacco leaves (Kroumova et al. 2013). A different trichome population, the tall glandular trichomes, secrete hydrophobic terpenes and diester sugars that solubilize phylloplanin from short glandular trichomes, which then migrates onto the leaf as exudate (Shepherd 2005).



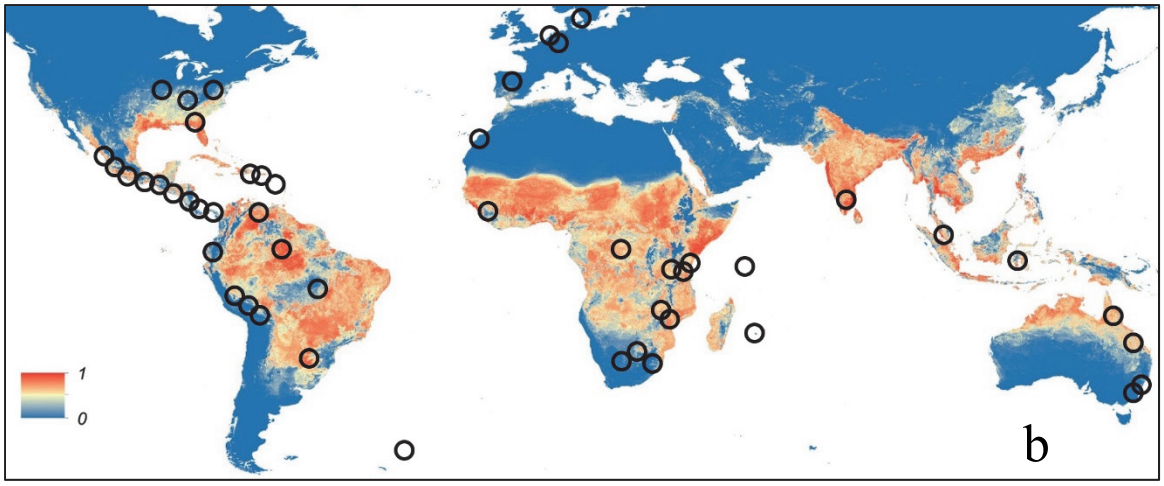


Figure 2. Global distribution comparison of *Aedes aegypti* and *A. albopictus* with *Impatiens walleriana*. These maps illustrate the probability of global occurrence of *Aedes aegypti* (A) and *albopictus* (B) (Kraemer et al. 2015). The colors represent the probability of occurrence of these mosquitoes from 0 blue to 1 red with a spatial resolution of 5 km × 5 km. The black circles indicate the known locations of *Impatiens walleriana* that have been identified in the wild (http://discoverlife.org).

Significance of Research

Prior to this research, very little was known about the molecular biology of *Impatiens walleriana* nectar and nectaries. This study has produced several important data that will henceforth serve for the development of impatiens into a model nectar system. It was shown that transgenesis of impatiens nectar is possible, through expression of the histochemical marker enzyme GUS using the promoter element from the eudicot-conserved nectary transcription factor, Crabs Claw, and the signal peptide from CARN2. A native nectary promoter for impatiens, piwSWEET14 was isolated through the use of RNA sequencing and inverse PCR. The isolation of this native promoter may increase expression of payload genes in impatiens nectar. Finally, the major nectar protein previously characterized by Chen and Kearney (2010), IW23, was discovered to be a homolog of T-phylloplanin and has been given the name iwPHYL21. iwPHYL21 may prove useful as a fusion partner for some payload proteins that are difficult to express.

CHAPTER TWO

Materials and Methods

Impatiens: DNA, RNA, and Protein Experiments

Non-Transgenic Plant Growth Conditions

Impatiens walleriana seeds were purchased from outsidepride.com. They were planted in Professional Growing Mix from Sungro®. Plants were fertilized using Plus® Pellets from Sure Soft®. The plants were grown under continuous light using Sun System Galaxy Grow AMP® T-8 bulb until about 2 months of age. Then they were transferred to a 16 h–8 h light-dark photoperiod under 400 W metal halide lamps with approximately 6ft of separation from the lamp. They were watered daily with approximately 200 ml of deionized water using drip-irrigation. Ambient temperature remained within a range from 25 °C–29 °C. Impatiens seeds were planted in groups, with each 4in pot sprouting 15–30 plants. Between 2–4 weeks old the seedlings were transferred to separate pots and fertilized. They were fertilized again every 4–6 weeks. After plants grew to 6 inches tall, they were repotted into 5 inch pots.

Genetic Engineering of Impatiens and Evaluation of Transgenesis.

Modifications to pORE vectors. The final vectors used to transform impatiens with to express GUS in their nectar were modified from two vectors, pLBW4 and pLBW5, a generous gift from Dr. Clay Carter. Both pLBW4 and pLBW5 were created

by Dr. Carter from the pORE4 vector published by Coutu, C., Brandle, J., Brown, D., et al. (Coutu et al. 2007). To create pLBW4 and pLBW5, nectary promoter elements Crabs Claw (CRC) and Cell Wall Invertase 4 (CWINV4) from *A. thaliana* were inserted into the pORE-04 vector at the SacII and AclI restriction sites in the multiple cloning region (MCR) to form pLBW5 and pLBW4 respectively (See Figure 2.2). The genes CRC and CWINV4 from *Arabidopsis thaliana* will hence be referred to as atCRC and atCWINV4, respectively, and the promoters for these genes will be referred to as patCRC and patCWINV4, respectively.

Antibiotic resistance is used to select for explants that have successfully transformed. pORE-04 contains the Neomycin phosphotransferase II (nptII) gene that confers resistance to the antibiotic kanamycin. The cryptic constitutive promoter PENTCUP2 from *N. tabacum* drove nptII expression.

A gene block was synthesized containing a 5' PacI restriction site, the CARN2 signal peptide (CARN2sp), the first 394bp of the GUS open reading frame (ORF). The GUS ORF contains a SnaBI restriction site that was located at the 3' end of the geneblock. The sequence of this geneblock, including visualization of the new PacI site, CARN2 signal peptide, 5' portion of the GUS ORF, and SnaBI site, can be seen in Figure 2.1. The viral vector SHEC74GUS contains a full length ORF for GUS. SHEC74GUS was digested using PacI and SnaBI, resulting in loss of the 5' end of the GUS ORF. The CARN2sp-GUS gene block was also digested with PacI and SnaBI and ligated into SHEC74GUS to form SHEC74CARNspGUS, as shown in Figure 2.2.

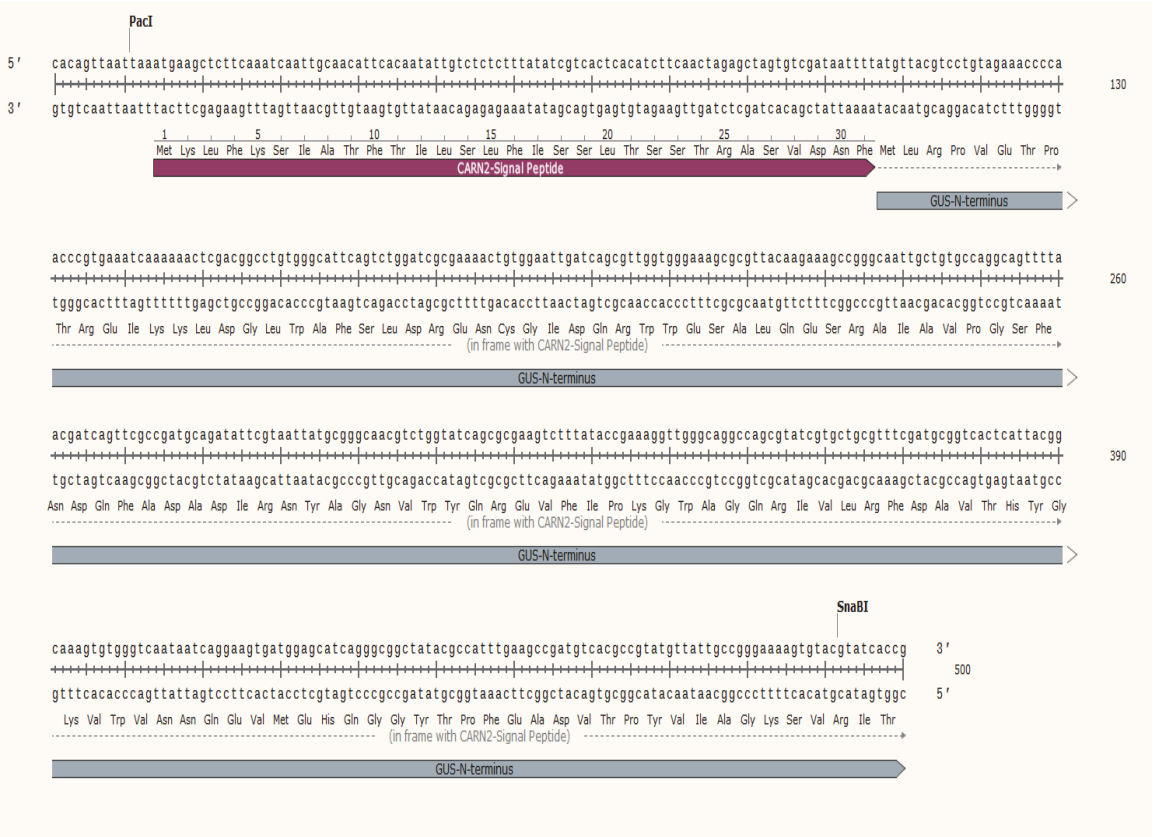


Figure 2.1. Gene Block of CARN2sp and Partial GUS ORF. The geneblock contained a 5' PacI site, the CARN2 signal peptide, the first 304bp of the GUS ORF. The translated amino acid sequence is shown below the DNA sequence.

The entire CARN2sp-GUS was cloned into pLBW4 and pLBW5 by a collaborator, Grace Pruett, by digestion of the CARNsp-GUS from the vector by the using PacI and BssHII restriction enzyme sites. XmaI and EcoRI sites were added to the CARNsp-GUS, at the 5' and 3' end respectively, via PCR amplification with primers containing the restriction sites of XmaI and EcoRI.

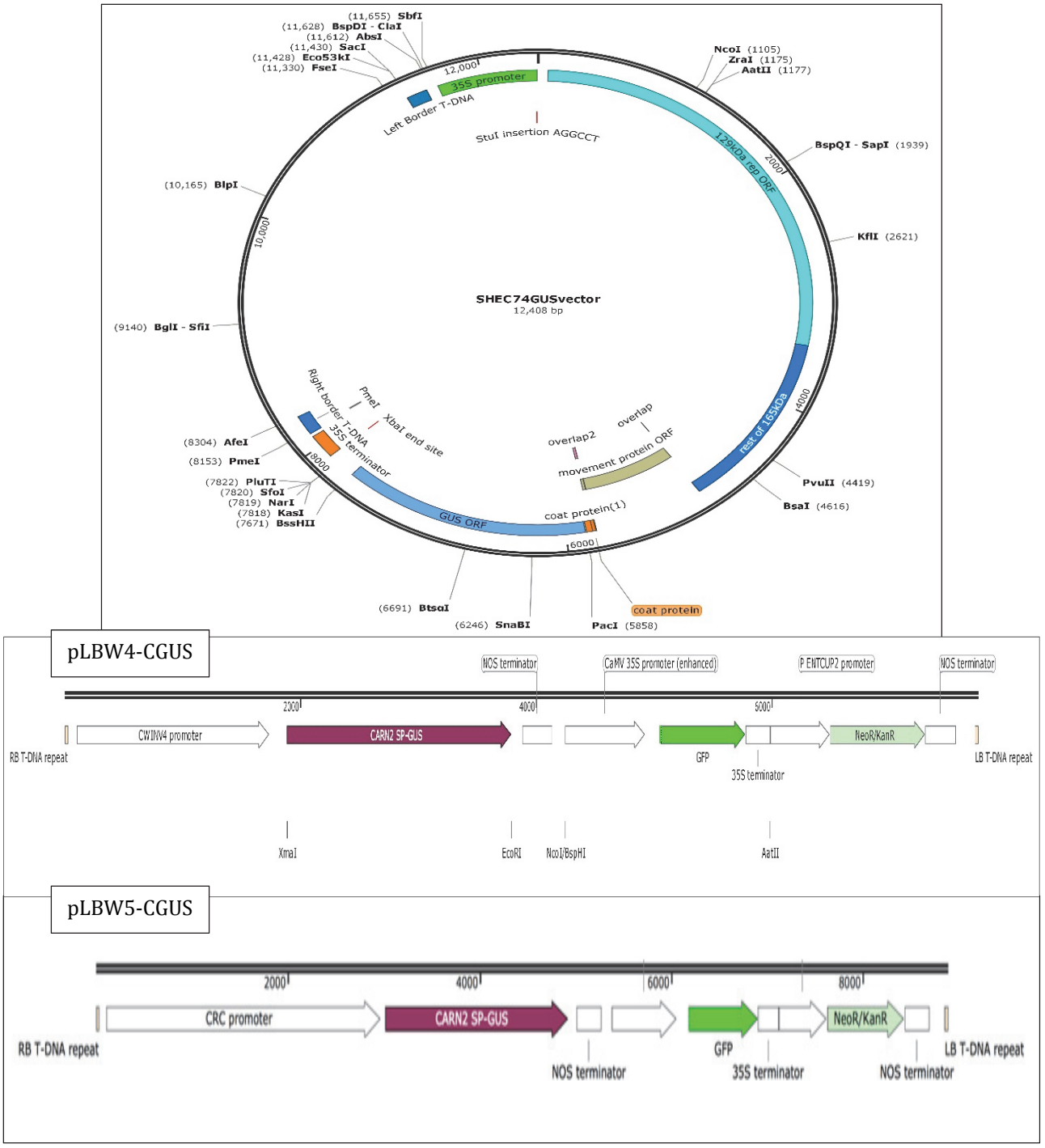


Figure 2.2. Genetic Engineering of CARN2sp-GUS into pLBW4 and pLBW5. The vectors displayed from top to bottom are SHEC74GUS, pLBW4-CGUS and pLBW5-CGUS. Both pLBW4-CGUS and pLBW5-CGUS were created from the pORE-O4 series by addition of a CARN2-GUS fusion into the restriction sites XmaI and EcoRI and the addition of a CaMV 35S enhanced promoter (pCaMV35S) fused to GFP into NcoI and AatII. Ligation of pCaMV35S-GFP at the NcoI site deleted the NcoI sequence and created a BspHI restriction sequence.

Additionally, the enhanced Cauliflower Mosaic Virus (CaMV) promoter driving expression of GFP was inserted into pLBW4 and pLBW5 at the restriction sites NcoI/BspHI and AatII to form pLBW4-CGUS and pLBW5-CGUS respectively.

Transgenesis and growth conditions. Impatiens were grown, transformed, and cultivated until soil-ready by Dr. Yinghui Dan at Virginia Polytechnic Institute. In Dr. Dan's laboratory, *I. walleriana* red accent seeds were gas sterilized for approximately 15 minutes with chlorine gas, grown on germinating media, transferred to media to promote growth into multiple bud clusters and transformed using *Agrobacterium tumefaciens*, producing explants. The explants were transferred to media that promotes shoot formation. Once shoots appeared, the shoot and the multiple bud cluster of cells attached to the shoot were excised and transferred to shoot elongation media. Once shoots were approximately 2 cm, they were transferred to media that promotes root formation and elongation. Once root hairs appeared and grew to appropriate length, the explants were transferred into pea soil pods and stored inside sterile Magenta boxes. When the plants were approximately 1 month post-Magenta box transfer, about 4 cm tall, they were shipped to Baylor University. Upon receiving the plants they were allowed a 24 h acclimation period in room temperature under a 16 h- 8 h photoperiod under indirect light, in a large climate and light-controlled incubator. The incubator had five T-8 fluorescent bulbs in the door and the plants were shielded from the light through the placement of a large plastic tray. The explants proved sensitive to the direct light, and responded better with the shielding during this acclimation period.

After acclimating, the plants were transferred to a growth room at which the ambient temperature ranges from 26 °C–29 °C. The explants were transferred to 8 ” soil pots containing Miracle Grow Garden Soil and sealed from the atmosphere for 24–36 hours by covering them with a clear plastic barrier, in the form a 20 oz. plastic cup. From this point onward, the lighting conditions of the plants corresponded with the lighting conditions detailed for normal impatiens growth, meaning they were under fluorescent bulbs after soil transplantation for approximately 2 to 3 months, and then transferred under metal halide lamps. Once in soil, the plants were watered by hand everyday with approximately 200 ml of water. Plants were slowly acclimated to atmospheric conditions through slight openings in the container covering them. After 24 h of slight acclimation, the containers were completely removed and plants were exposed to open air. Explants were fertilized 6 weeks post soil-transfer.

Histochemical staining of tissue. 3 month-old transgenic *I. walleriana* were used for tissue collection. GUS interacts with the substrate 5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid (X-Gluc) to produce colorless glucuronic acid and a bright blue chloro-bromoindigo precipitate (bromoindigo dye) (Jefferson et al. 1987). The X-Gluc staining protocol used was adapted from the one mentioned in (Jefferson et al. 1987) and published by the Stockinger lab from Ohio State University was followed (Stockinger 2014), but scaled down for the relatively small size of the tissue. Staining solution was composed of 100 mM sodium phosphate pH 7.0, 10 mM EDTA, 0.1% Triton X-100, 1 mM potassium ferricyanide, and 2 mM X-Gluc. For nectaries, 100 µL of the staining solution was used for one nectary, instead

of the mentioned 1 ml. For leaf and stem tissue, 20 mg of tissue were stained with 1mL of staining solution. All tissue were then incubated overnight at 37 °C. After the incubation, the staining solution was removed and tissues were fixed with several changes of ethanol for 10 h at a time. The use of ethanol also removed much of the chlorophyll from tissue. Tissues were imaged using a stereomicroscope in Baylor's microscopy center.

Chemical detection of GUS in transgenic impatiens nectar. Nectar from pLBW5-CGUS (patCRC-GUS) plants were used for the assay. 6 droplets of nectar were collected on the end of a plastic pipet tip and dissolved in 20 µL of staining solution as described above, on Parafilm ®. Additionally, 6 droplets of nectar were collected in the same manner and 20 µL of staining solution was added. The droplets were covered and incubated at 37 °C overnight. The sites where the nectar evaporated were imaged using a stereomicroscope in Baylor's microscopy center. Sites were examined in full for the presence of bromoindigo crystals. Both at low magnification at 10 x and ranging to the 110 x to search for trace amounts of bromoindigo crystals.

Impatiens RNA Experiments

RNA isolation protocol. RNA was isolated by following a user-developed protocol for Qiagen ® RNeasy ® Plant Mini Kits, specific for isolating RNA from tissue with large relative percentages of phenolics or polysaccharides adapted from a protocol published in Plant Disease (MacKenzie et al. 1997). The procedure was as follows: 50 mg–100 mg of *I. walleriana* nectaries from 5-month-old adult plants

were collected for RNA sequencing. As nectaries are quite small, ranging from 0.2 mg–1 mg in a fully grown (at least 4 months old) impatiens, we pooled between 100 and 150 nectaries for one sample. All nectaries for one sample came from the same plant, hence why the plants needed to be fully mature in order to provide the sufficient size and number of nectaries. Nectary collection was not a simple process and had to be optimized over several iterations. The method that was deemed most efficient was removing nectaries from the plant with sterile tweezers 5–10 at a time (nectaries stick to the tweezers quite readily) and immediately submerged into a mortar containing liquid nitrogen. After several seconds the nectaries became frozen and were gently scraped with another set of sterile tweezers into the pool of liquid nitrogen. Additional liquid nitrogen was added to the mortar so that the liquid nitrogen never evaporated in entirety for more than 5 seconds. This process was repeated until sufficient nectaries were collected, around 30 minutes. The nectaries were ground in liquid nitrogen with a pestle to a fine powder. The powder was transferred to an Eppendorf tube. After liquid nitrogen evaporated but before the powder could thaw, 600 µL of Lysis buffer was added and vortexed with the powder. Lysis buffer is contained: 4 M guanidine isothiocyanate; 200 mM sodium acetate, pH 5.0; 2 5mM EDTA; 2.5 % (w/v) polyvinylpyrollidone (molecular weight ~40 g/mol); and 1 % (v/v) β-mercaptoethanol (immediately before use). After vortexing, 60 µL of 20 % sarkosyl was added and samples were incubated in a 70 °C water bath for 10 minutes with intermittent mixing. The powder solution, containing the ground nectaries, lysis buffer and 1 % BME, was then pipetted into a QIAshredder ™ Spin Column and centrifuged for 2 minutes at 18,000 x g. This

produced a pellet of cellular debris at the bottom of the flow through tube, along with supernatant containing the lysed-cell contents. The supernatant of the flow-through was carefully pipetted into a new Eppendorf tube, being careful not to disturb the cell-debris pellet at the bottom of the flow-through tube. Then, a half-volume of 100 % ethanol was added and mixed via pipetting. This mixture was transferred to an RNeasy® spin column as instructed in step 5 of the Qiagen® RNeasy® Plant Mini Kit Quick-Start Protocol and processed according to the protocol until pure RNA was isolated (See Appendix E for Qiagen® Quick-Start Protocol). *I. walleriana* stem and leaf RNA was isolated exactly as according to the Qiagen® Quick-Start Protocol (using step 1a instead of 1b). All RNA was analyzed using a two reads from a NanoDrop spectrophotometer. All RNA samples required between a 1.8-2.1 260 nm: 280 nm light absorbance ratio. For RNAseq a minimum of 500 ng of RNA is recommended. All samples contained 30 µl of RNA with a concentration ranging between 50 ng/µl–200 ng/µl.

RNA sequencing. A total of 2 µg of RNA for each sample was submitted for RNA sequencing. Two biological replicates each of total RNA from stem, leaf and nectary tissue from *Impatiens walleriana* were submitted to the Minnesota Genomics Center for RNA sequencing. RNA was first quality-checked. All RNA passed the quality check inspection. Mature messenger RNA (mRNA) from total RNA was separated through binding of the poly-adenosine (poly-A) tail, a characteristic of mature mRNA, to oligo-dT primers. The mRNA samples were then sequenced on an Illumina® HiSeq 2500 with 50 bp reads. The transcripts for each tissue were de novo assembled by Trinity RNA-seq software as published by (Haas

et al. 2013) at the University of Minnesota Genomics Center. The Trinity program assembled contigs using the 50 bp reads from the mRNA sequencing. *De novo* assembly relies on the use of multiple De Brujin graphs to produce contigs. This process has been shown to possess highly accurate de novo predictions of transcriptomes that lack a reference genome (Haas et al. 2013). Potential translation sites were also predicted by Trinity and searched against a SWISSPROT database by a collaborator at Minnesota State University.

DNA Experiments and Promoter Isolation

Isolation of Impatiens walleriana genomic DNA. DNA was isolated from 1 g of I. walleriana leaf tissue using the protocol as published by (Dellaporta et al.). Following the protocol, 1 g of leaf tissue was immediately frozen in liquid nitrogen and ground to a fine powder in a mortar with a pestle. The tissue was then directly transferred into 15 ml of extraction buffer 1 (EB1), which contained: 50 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0, 100 mM tris pH 8.0, 500 mM sodium chloride and 0.7 % fresh 100 % β-mercaptoethanol, BME (the BME was added after autoclaving the prior ingredients). After mixing, 1 ml of 20 % SDS was added to the sample. The samples were vortexed vigorously and placed in 65 °C water bath for 10 minutes. Next, 5 ml of 5 M of potassium acetate was added. The sample was vortexed and immediately placed on ice for 20 minutes. Then, the tubes were centrifuged for 20 minutes at 25,000 x g at room temperature. A pellet formed consisting of cellular debris should form at the bottom of the tube. Careful not to disturb this pellet, the supernatant is poured over cheesecloth into a from tube

containing 10 mM 100 % ice-cold isopropanol (note: a DNA supercoils should be very visible at this step. They should appear clearish-white). The sample was mixed by inversion and left overnight in -20 °C (although the protocol states 30 minutes is sufficient). Next, the supernatant isopropanol solution was spun at 20,000 x g for 15 minutes at 4 °C . A pellet of DNA collected at the bottom of the tube. The supernatant was gently poured off and the pellet was dried for 10 minutes by inverting the tube on a paper towel. The pellet was then dissolved in 700 µl of extraction buffer 2 (EB2), which was made of: 10 mM EDTA pH 8.0 and 50 mM tris pH 8.0. Resuspending the pellet in EB2 required an overnight incubation at 4 °C followed my several minutes of hand mixing. The resuspended DNA was pipetted into a 1.7 ml Eppendorf ™ tube and centrifuged for 10 minutes at 22,000 x g at room temperature. This step helps to clear any remaining contaminants, which form as a pellet at the bottom of the tube. The DNA, suspended in solution in the supernatant, was transferred into a new tube and 75 µl of 3 M sodium acetate and 500 µl 100 % isopropanol were added. The solution was thoroughly mixed by vortexing and then centrifuged for 30 seconds at 22,000 x g at 4 °C. A pellet of pure DNA formed at the bottom of the tube and was washed with 500 ml of 80 % ethanol. The pellet was air-dried, just enough to ensure no residual ethanol leaked from the pellet and was resuspended in 100 µl of TE buffer, which contained: 10 mM tris HCl pH 8.0 and 1 mM EDTA. The sample was verified for purity (a 260/280 ratio of at least 1.9) and concentration was measured using Nanodrop spectrophotometry.

Primers, thermocycler and gel electrophoresis. All primers for PCR and sequencing were ordered and analyzed through IDT. Gene blocks for cloning in the

pORE vectors were purchased from IDT. All PCR reactions were carried out in an Eppendorf Mastercycler. 1 % agarose (Sigma) gels containing 0.005 % Gel Star™ (Lonza) were placed in a FisherBio Tech Mini Horizontal Unit. DNA was separated as described in (Olivera et al. 1964). For inverse PCR, voltage was lowered to provide between 30–50 mA current.

For this study 11 primers were constructed. A list of these primers can be found in Table 2.1. Six primers were made for iPCR of the RuBisCO large subunit gene in impatiens and five primers were made for iPCR of the SWEET14 gene in impatiens. Starting and ending positions are included in Table 2.1, and correspond to the sequences for impatiens rbcL and iwSWEET14 included in Appendix C.

Purification and sequencing of individual PCR products. Gel bands were purified using either Promega Wizard ® SV Gel and PCR Clean-Up Kit (REF#A9281) or Sigma ® GenElute ™ Gel Extraction Kit. All centrifuge spins were at room temperature. Modifications were made to the Promega Wizard ® SV gel and PCR Clean-Up Kit recommended protocol entailing both: 1) a drying spin for 3 minutes at 8,000 x g with the centrifuge lid open rather than the recommended 1-minute spin at 16,000 x g with the lid closed to dry the columns proceeding the two wash spins; as well as 2) eluting DNA with only 30 µL water rather than 50 µL. The protocol for the kit from Sigma ® was followed as written. The author notes a significant decrease in DNA yielded from the kit from Sigma ® compared to DNA yielded using the kit from Promega.

All DNA sequencing was performed by Macrogen USA, using their EZseq service. This service uses 5–7.5 µL of purified DNA, which is recommended to have a

concentration between 50 ng/µl and 200 ng/µl. However, the concentration of the DNA samples submitted never exceeded 27.1 ng/µl. Macrogen's EZseq guidelines for template recommends using 5 µl, but since the template concentration was less than optimal additional template was used. However, the amount of DNA template for sequencing was still below the total amount recommended. Therefore, this generating sequences that were shorter than optimal (<1,000 bp obtained from sequences using sufficient template). Sequences were scored at each position, and a file showing these scores can be found in Appendix in Figure A.1, A.2, and A.3. 2.5 µl of one 10 mM primer was used for each sequencing event. Primers for sequencing were between 20 bp and 25 bp long, as per Macrogen USA EZseq recommendations. Primer melting temperature was between the 55 °C–60 °C recommended for sequencing.

Table 2.1. Primer Names and Sequences

Primer Name	Primer Sequence – primers are oriented from their 5' to 3' – Numbers denote where in the sequence the primer binds. Any nucleotides outside the numbers do not bind to the target. Primers with "RUB" bind to the RbcL sequence and primers with "14" bind to the iwSWEET14 sequence		
RUB-UP1	186 –	TGCTAGTTCCGATCATTCAGGTAGCCAG	- 158
RUB-UP2	155 –	GTGTCTACAGGTACATGGTCATCTTCTA	- 129
RUB-UP3	125 –	GCGTCGATGGCGTCGTGGACGAAGAAGT	- 98
RUB-UP4	82 –	ATCTTGGCAGCATTCCGAGTAAGTCCTC	- 54
RUB-UP5	1173 –	GTCCTAAAGTTCCGCCACCA	- 1154
RUB-DN	328 –	TCAAAGCCCTGCGCGCTCTA	- 347
14A	GAGG – 53 –	TGGAAGAACGAGGAATTACG	- 33
14B	GTAT – 102 –	AACCCATTGCCTTGCGTCGG	- 82
14C	GGAG – 305 –	CGACCAAAGAGTTTCAACC	- 325
14D	CGCC – 305 –	TGGTGTTGTTATAGAGTCCC	-324
14E	CGAG – 583 –	GTGACAATGGCAATGGC	-566

Inverse PCR. Inverse polymerase chain reaction (iPCR) is a technique of amplifying unknown sections of DNA adjacent to known sections of DNA, using specific restriction enzyme sites and ligation protocols. The technique works well for isolating upstream promoter sequences, especially in organisms without reference genomes (Triglia 2000). Figure 3.1 illustrates the approach of iPCR. The technique relies upon utilizing a restriction enzyme sequence in the known section of DNA. Genomic DNA is digested with the enzyme creating many fragments, most of which are non-target sequences. These fragments are then ligated in conditions that promote self-ligation, creating circularized fragments of DNA. This pool of DNA is used as a PCR template for amplification using outward-facing primers corresponding to the known section of DNA.

The specific digestion and ligation conditions for this experiment were as follows. Digestion reactions were set up in 100 µl total volume and used 10 µg of *I. walleriana* genomic DNA in 1 x NEB Cutsmart ™ buffer with 50 U of enzyme, and incubated in corresponding temperatures for optimal enzyme efficiency for 18 h. This long digestion time is used to encourage complete digestion of the large quantity of DNA used. Enzymes were heat inactivated at their respective temperatures for 15 minutes in a water bath. Then, ligation reactions were set up using 500 ng of cut genomic DNA in total volume of 500 µl. NEB ligation buffer was added to a 1 x final concentration, and 1,600 U of T4 DNA ligase from NEB was added to the ligation reaction. The ligation reactions lasted 18 h overnight in 15 °C and for 2 h immediately after at 24 °C. The large volume and cool reaction temperature encourage circularization of the DNA fragments. T4 DNA ligase was

inactivated by incubating the ligation reaction mixtures in a 65 °C water bath for 10 minutes as per NEB's recommendation.

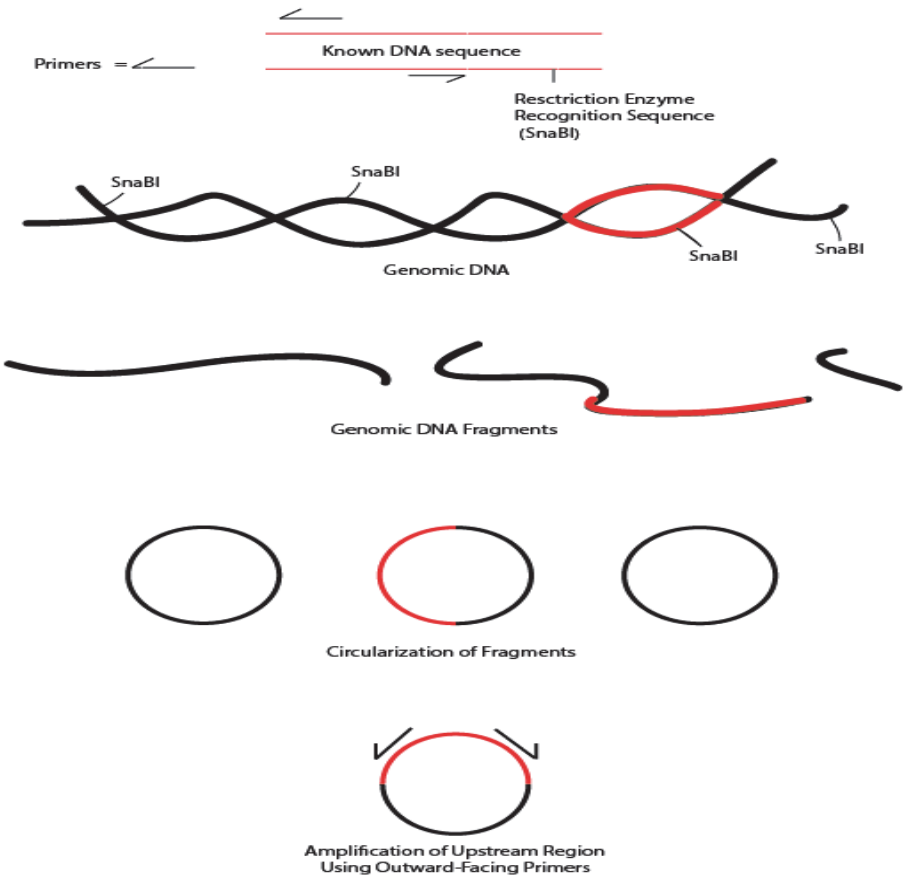


Figure 2.3. Illustration of Inverse PCR to Isolate Unknown, Upstream DNA

Circularized DNA was used as a template for iPCR reactions. Several attempts were made trying different amounts of reactants, different annealing and extension times, and different PCR additives like bovine serum albumin (BSA) and dimethyl sulfoxide which have been known to improve PCR reactions in certain cases (Rochelle et al. 1997). The conditions that proved to produce reliable iPCRs were as follows. A standard program of 5 minutes at 95 ˚C followed by 30 cycles of: 95 ˚C for

10 s; 50 ˚C–57 ˚C (depending on primers) for 30 s; 72 ˚C for 2 minutes. Annealing temperature varied depending on the primers used. All reactions were 50 µL and contained 5 ng–20 ng template; 500 nM of each primer; 250 µM dNTPs; 1 U Q5 polymerase. Initial inverse PCR products were either gel purified or diluted 1:100 in water and used as template for the second round of iPCR.

A positive control iPCR—RuBisCO large subunit promoter for Impatiens. In order to establish a protocol for performing iPCR on impatiens genes, a positive control experiment was performed using a gene where the upstream sequence was known. The RuBisCO large subunit gene (rbcL) is a non-nuclear gene located in plant chloroplasts, and is highly conserved in plants. It is often sequenced, along with the upstream ATPase beta-subunit gene (atpB) and the intergenic space, to determine phylogenetic relationships between different plant taxa (Savolainen et al. 2000). The rbcL sequence (accession: AB043508.1), and the upstream intergenic sequence (accession: DQ147892.1) are available on NCBI's GenBank. Primers were designed using the sequence for rbcL.

Promoter sequence analysis. In silica prediction of promoter sequences was performed using Plant Care promoter prediction software (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The software uses 417 verified promoter element sequences from plants, mostly monocots and eudicots, and uses "clustering, motif recognition and a probabilistic approach based on Gibbs Sampling" to predict upstream *cis*-regulatory elements (Lescot et al. 2002). Other experiments have relied upon this predictive approach to determine promoter elements before

experimental verification (Dojcinovic et al. 2005). A known promoter element for rbcL was isolated using iPCR and analyzed to serve as an *in silico* prediction comparison for the unknown iwSWEET14 promoter element.

Nectar Protein Experiments

Nectar protein purification with ethanol. Nectar was collected from extrafloral nectaries by swabbing nectar onto strips of filter paper and dissolving in 10 mM sodium phosphate buffer pH 6.8. Cold ethanol was added 1:1 and the solution was centrifuged at 20,000 x g for 20 minutes at 4 °C. The supernatant was gently poured off and the protein pellet was resuspended in water.

SDS-PAGE gel electrophoresis. 12 % acrylamide SDS-PAGE gels were cast in BioRad gel pouring stations. Electrophoresis used a BioRad Mini Trans-Blot Cell rig. The protocol used for loading and separating samples can be found in (Peterson 1977).

Two-dimensional gel electrophoresis of proteins. Two-dimensional gel electrophoresis was carried out to determine if the large 21 kDa band found from impatiens nectar protein SDS-PAGE was composed of multiple proteins.

To prepare for 2D separation, impatiens nectar protein was concentrated by ethanol precipitation as mentioned previously. It should be noted that nectar protein was collected from plants only 3 months old producing significantly less nectar than the plants used for regular SDS-PAGE.

For isoelectric focusing, 50 µl of concentrated nectar protein was added with 100 µl De Streak rehydration solution from GE Life Sciences (product code: 17600319), and 2 µl of 1 M dithiothreitol (DDT) to facilitate protein reduction of disulfide bonds. This solution was placed in the bottom of a ceramic coffin (GE Ettan™ IPGphore) and an 11 cm immobilized pH gradient (IPG) strip ranging from pH 3–10 was placed on top of the solution. The strip was incubated on top of the solution for 15 h at room temperature until the IPG had sufficiently absorbed the protein sample. Then, voltage was applied, using a GE IPGphore, to encourage migration of proteins along the pH gradient. Proteins migrate to their isoelectric point, therefore enabling resolution between two proteins of similar size. The amount of voltage applied varies according to the size of IPG strip being used, as well as the pH range. The GE IPGphore has recommended protocols depending on the IPG strip used. In this experiment, voltage was applied as follows: 300 V for 30 minutes; 1000 V for 30 minutes; 5000 V for 1 h.; 5000 V for an additional 500 cumulative volt hours. Total volt hours accumulated for the entire run was 14,726 V-hours. Temperature was kept at 23 °C during the entire isoelectric focusing. The IPG strip was then removed and prepared for 2nd dimension separation using SDS-equilibration buffer from GE Life Sciences. To equilibrate, the IPG strips were placed into 15 ml conical tubes using sterile tweezers and sufficient equilibration buffer was added to cover the IPG strip. Strips were incubated in two changes of SDS-equilibration buffer while rocking for 30 minutes. each. Then IPG strips were placed atop Bio-Rad precast SDS-PAGE gradient gels, 4 %-16 % acrylamide, specifically made for 2nd dimension separation. The IPG strip was sealed in place using molten

1% agarose at approximately 55 °C administered via pipet. Once the agarose had polymerized, SDS-PAGE was carried out as described previously.

SDS-PAGE gel band purification for mass spectrometry and de novo *analysis*. A 20 kDa to 25 kDa protein band from SDS-PAGE gels was excised in a sterile hood on Parafilm ® and placed into Eppendorf tubes to prevent keratin contamination. Keratin is a protein that is abundant on the outer layer of human skin, and composes much of human hair. Therefore, there without using sterile conditions there a high likelihood of keratin contamination of the sample. This will cause mass spectra for keratin to overshadow the mass spectra for the target protein. Additionally four peptide spots from a 2D-SDS-PAGE were excised from the gel in the same manner. These were destained to remove Coomassie Blue, then they were reduced of disulfide bonds with tris(2-carboxyethyl)phosphine (TCEP) and alkylated with iodoacetamide at free cysteine residues to prevent disulfide bonds from reforming. Then, the peptides were digested with trypsin as described in the protocol for the in-gel digestion kit (Thermo Fisher product #81879X). No modifications were made to the standard protocol that is included with the kit. Peptides were cleaned with C-18 clean-up columns (Thermo-Fisher), vacuum dried and reconstituted in 0.1 % formic acid.

UPLC and mass spectrometer. Peptides were processed by ultra-pressure liquid chromatography (UPLC) in tandem with mass spectrometry. UPLC separation used a 170 Å C18 column. Eluted peptide fragments fed into the Synapt G2 and were ionized by electrospray. A MSe analysis was performed to generate mass spectra for

a peptide databank search. Additionally, a data-dependent acquisition (DDA) was performed. The DDA fragmented peptide ions that produced enough signal to surpass an intensity threshold. This generated amino acid profiles for peptide fragments. UPLC and mass-spectrometric analysis were performed in Baylor University's Mass Spectrometry Center under the guidance of Dr. Alejandro Ramirez.

A databank of all SWISSPROT proteins (552,259) was used for a peptide databank search of the MSe mass spectra for the SDS-PAGE nectar protein band. The databank search was done with Waters Protein Lynx Global Server software. For the details of the search parameters for the databank search, please refer to Figure D.3 in Appendix D.

For the detailed parameters of the data-dependent acquisition, please refer to the data file under "DDA Data Acquisition Parameters_Cox, Andrew" at http://andmcox19.wixsite.com/thekearneylab/data.

De novo analysis of MS/MS spectra was performed using Waters PepSeq program, part of the Waters BioLynx software, a package included in Masslynx software version V4.1. MS/MS spectra were centered using the lock mass value of glu-1-fibrinopeptide B (Glu-Fib), exact mass = 746.2642 kDa. Peaks were selected and deconvoluted using the Maxent3 function in the mass spectra viewer of the Mass Lynx software. The centered and deconvoluted spectra were analyzed using the PepSeq program.

As a standard, mass spectra from known proteins alcohol dehydrogenase and enolase were sequenced in the same manner described above, using spectra from runs provided by Dr. Ramirez.

CHAPTER THREE

Results

Impatiens nectaries express GUS via arabidopsis promoters

Impatiens were constructed using nectar promoters derived from *Arabidopsis thaliana* as a first test of the concept of impatiens as a model nectar system. It was discovered the atCRC-CARN2sp-GUS expression cassette produced GUS in impatiens nectaries and nectar and young stem and leaf tissue in areas where nectaries typically form, but had not yet emerged. Bromoindigo dye, the product of GUS interacting with X-Gluc, is pictured in Figure 3.1. The presence of GUS in non-nectary tissues is not surprising, since CRC functions as a transcription factor that is necessary for nectary formation, and GUS was only observed in nearby tissue where the nectaries typically form. Regrettably, this data is not displayed as it was lost in the process of creating cross-sections for imaging.

Plants containing the patCWINV4-CARN2sp-GUS expression vector (patCWINV4-GUS vector) were received later, and as such, nectaries and nectar could be assayed in time for this thesis. However, leaf and tissue were assayed in young plants and unsurprisingly did not display GUS activity in pre-nectary tissue. Since the function of atCWINV4 is to cleave sucrose just prior to nectar secretion, it would have been surprising to observe its activity outside of nectary tissue. Regrettably, this data is not displayed as it was lost in the same process mentioned above of creating cross-sections for imaging.

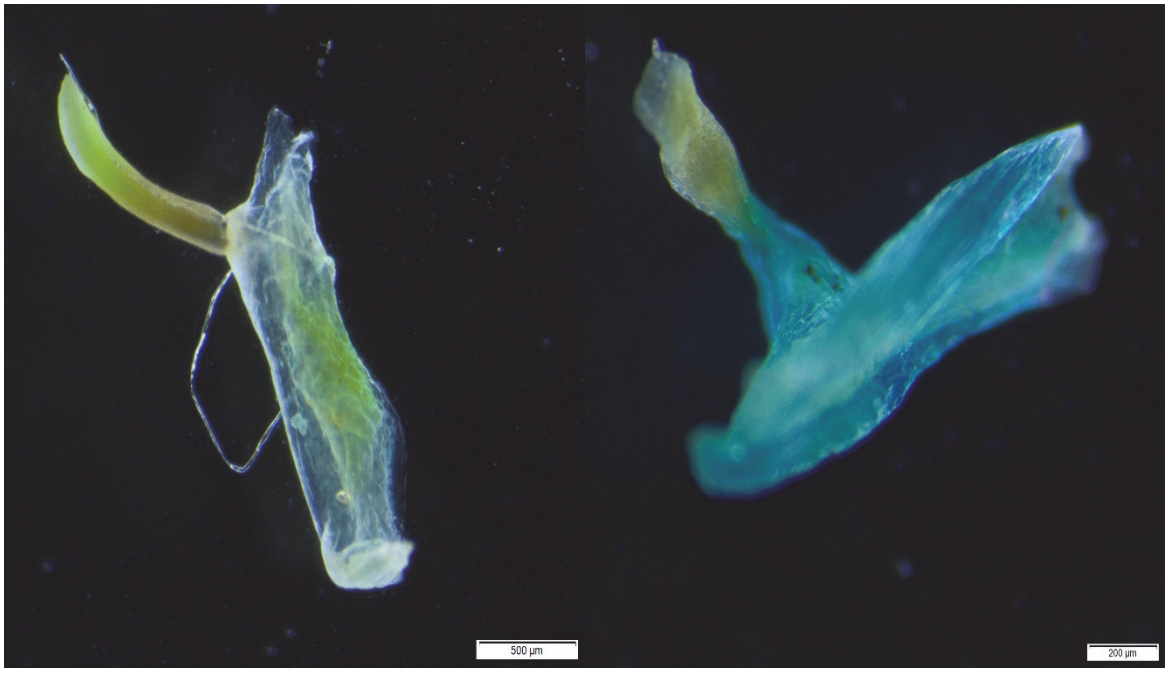


Figure 3.1.GUS Detection in Nectary of patCRC-GUS Impatiens Compared to Control. On the left is a nectary from untransformed impatiens. No bromoindigo dye was observed. On the right image is a nectary from patCRC-GUS transformed impatiens. Bromoindigo dye is apparent in the basal nectary and surrounding stem tissue. The patCRC-GUS nectary was from the specific plant documented by Dr. Dan as 030816-T3-2-2. Both nectaries were imaged after 36 hours of being submerged in ethanol to remove chlorophyll.

The presence of GUS was detected in the nectaries of three different patCRC-GUS impatiens. Pictured above is the highest amount of GUS expression. Additional photos can be found in Appendix D Figures D.1 and D.2. Overall, the patCRC-GUS vector demonstrates the ability to consistently express GUS in nectaries and surrounding tissue of impatiens. This activity is tissue-specific as leaf and stem tissue not located next to nectaries did not contain bromoindigo dye.

Next, an experiment was performed to determine if the CARN2 signal peptide fused to GUS was sufficient for localization of GUS to impatiens nectar. Nectar from the 030816-T3-2 patCRC-GUS impatiens pictured above on the right was evaluated

for the presence of GUS. Bromoindigo crystals were observed on the site of incubation of nectar and staining solution of patCRC-GUS nectar, where non-transformed nectar produced no bromoindigo crystals. Figure 3.2 shows the crystals detected from the patCRC-GUS *Impatiens* and shows a larger field of view for the negative control reaction. An additional photo showing another cluster of bromoindigo crystals from the reaction of patCRC-GUS nectar with staining solution can be found in Appendix D Figures D.

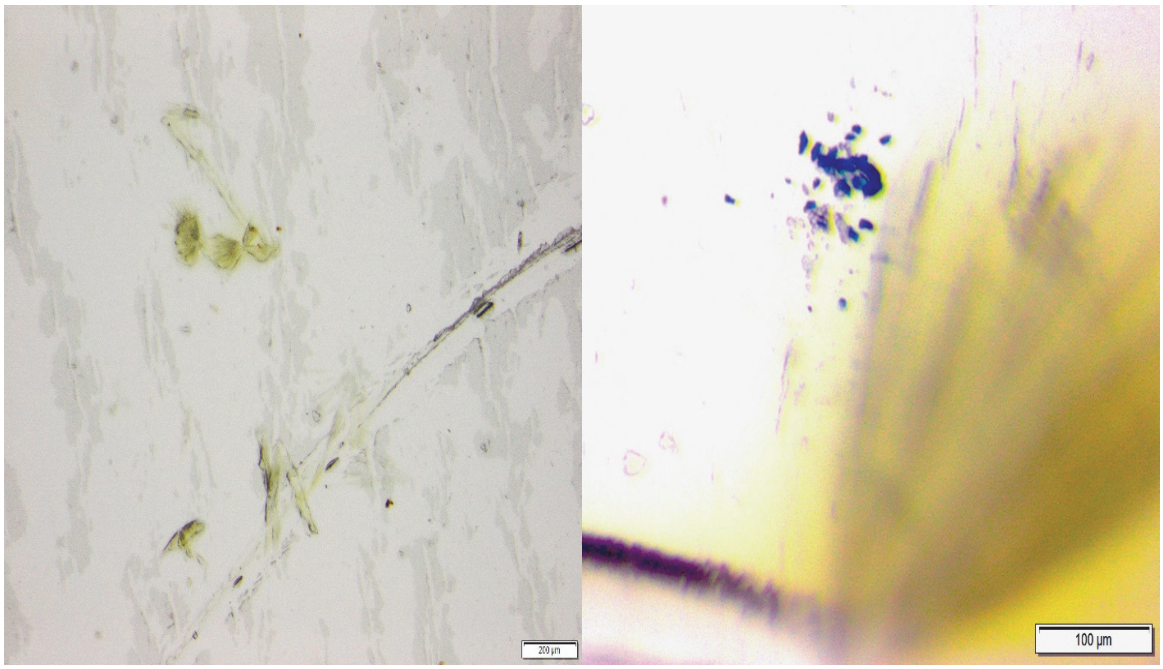


Figure 3.2. GUS Detection in Nectar of patCRC-GUS Impatiens Compared to Control. The left panel shows a negative GUS reaction, with unreacted potassium ferricyanide assay crystals left over after evaporation of non-transformed *I. walleriana* nectar mixed with X-Gluc and potassium-ferricyanide. No bromoindigo crystals were visible at any magnification level indicating there was not a confounding false-positive. The right panel shows a positive GUS reaction, with blue bromoindigo crystals produced using the same assay but with patCRC-GUS nectar.

The presence of such large amount of bromoindigo crystals is highly indicative of the presence of GUS in patCRC-GUS impatiens nectar. Additionally, no

bromoindigo crystals were observed from the negative control. The false positive rate for GUS is 5 % and the false negative rate is less than 1 % (http://www.x-gluc.com/xgluc.htm).

These experiments demonstrate the first successful transgenesis of impatiens nectaries and nectar, the first known transgenesis of extrafloral nectar, and the third recorded transgenesis of nectar.

RNA Sequencing Results

Impatiens RNA Sequencing—Sorting the Data to Identify Nectary-Specific Genes

Ultimately, RNA sequencing of impatiens leaf, stem and nectary revealed transcripts that are abundant in impatiens nectary tissue. One of these transcripts, impatiens walleriana SWEET14 (iwSWEET14), proved to be highly abundant in nectary tissue and was transcribed in very low levels in leaf and stem tissue. Other transcripts of note were the T-phylloplanin homolog which was the most transcribed in nectary tissue, but was also transcribed highly in stem tissue, and a non-specific lipid-transfer protein homolog. Even with these valuable proteins identified there remains many other unexplored componenets of the RNAseq data due to the vast amount of transcripts identified.

In total there were 39,900 unique contigs assembled by the program Trinity, which assembles RNA sequencing reads without a reference genome. From the 39,900 contigs searched, 13,983 possessed viable predicted translation products that matched to a SWISSPROT protein.

To make sense of the vast amount of mRNA for the purpose of identifying a good candidate to isolate a nectar promoter, transcripts were sorted by two thresholds: 1) a minimum number of 1,000 reads (averaged from the two biological replicate reads) and 2) a nectary to any-other-tissue read ratio of at least 10:1 (i.e. 100 reads in nectary, <10 reads in both stem and leaf tissue). The thresholds allowed for visualization of transcripts that would make ideal candidate for isolating nectary-specific promoter elements. A transcript with a read value less than 1,000 might not possess a promoter element worth isolating, since it would not be highly transcribed even if it was nectary-specific. The nectary to tissue transcription ratio ensures that even highly-transcribed genes in nectary are nectary-specific. Fig 3.3 shows the fourteen transcripts with mapped homologs in SWISSPROT that fit both parameters. Among the proteins listed in Figure 3.3 the bidirectional sugar transporter SWEET14 and the non-specific lipid-transfer protein stand out significantly for their high transcription rates and nectary-specific transcription relative to the other 12 genes. However, several highly transcribed mRNAs in nectary were excluded using these sorting criteria.

To ensure that highly abundant mRNAs were not overlooked, the transcripts were also ordered just according to the greatest amount of transcription in nectary tissue. To do so the two biological replicate reads for each technical replicate were averaged and ranked from most reads in nectary tissue to least. Figure 3.4 shows the 19 most abundant transcripts in nectary tissue. Of note is the most highly transcribed mRNA in nectary tissue, a homolog of T-Phylloplanin. This transcript was also highly

transcribed in one of the stem sample reads, thereby causing it to be overlooked using the nectary-specific transcription criteria from the first sorting.

Considerations of the RNA Sequencing Experiment

There were some specific considerations analyzing the RNA sequence data from this experiment. With only two biological replicates performed for each tissue type, the difference in transcription rate between tissues requires further validation, e.g. qPCR. It is for this reason that that the 10:1 ratio between nectary and next-highest tissue was used. This large ratio provide room for error in the reported transcription values. This way, even if moderate inaccuracies exist, selection for nectary-specific mRNA is achievable. For example, say if the mRNA in the tissue with the second highest transcription rate was actually transcribed twice as much as reported. This would still equate a five-fold greater increase in transcription in nectary tissue.

Another concern was the reliability of the *de novo* assembled contigs. The assembly of 50 bp reads into larger contigs without a reference genome was a concern when relying on the sequence data provided to be accurate. Therefore, a PCR was performed for the proposed sequence for iwSWEET14 from the *de novo* assembly. The PCR yielded a product, which was sequenced and matched with the *de novo*-assembled sequence for iwSWEET14. The PCR product sequence matched with the *de*-novo sequence. PCR experiments should still be used to verify sequences of other transcripts.

Using the RNAseq Data to Find a Nectary-Specific Promoter

The RNAseq data showed that iwSWEET14 is both highly transcribed in nectary tissue and highly nectary-specific relative to all the other transcripts from the RNAseq experiment. This can be seen in both Figures 3.3 and 3.4. Therefore, it was selected as a candidate for iPCR in hopes of isolating a nectary-specific promoter element.

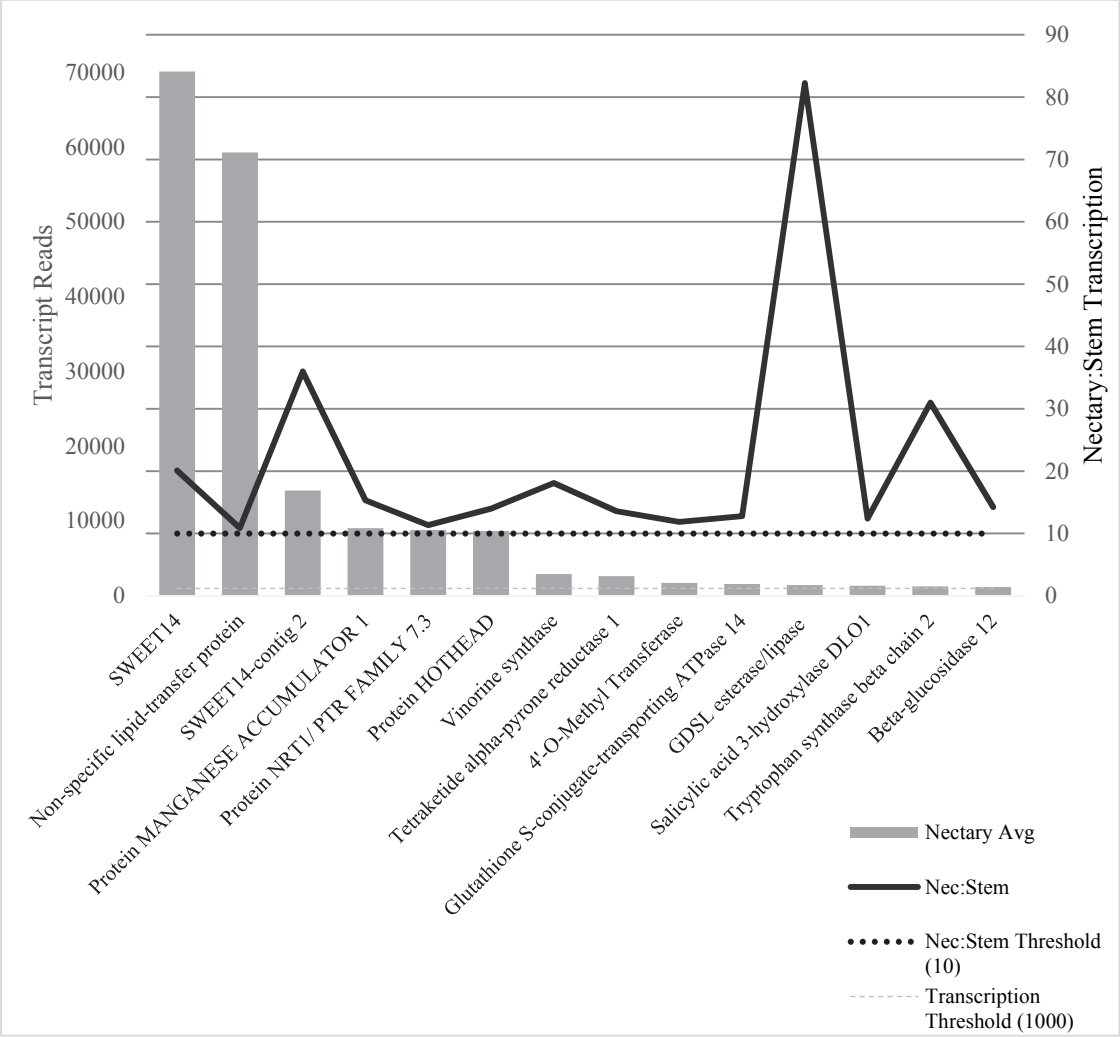


Figure 3.3. Transcripts with High Nectary Transcription and High Nectary Transcription Specificity. This graph displays the fourteen transcripts matched to proteins in SWISSPROT that possessed RNAseq reads above 1,000 and a nectary:other-tissue read ratio greater than 10:1.

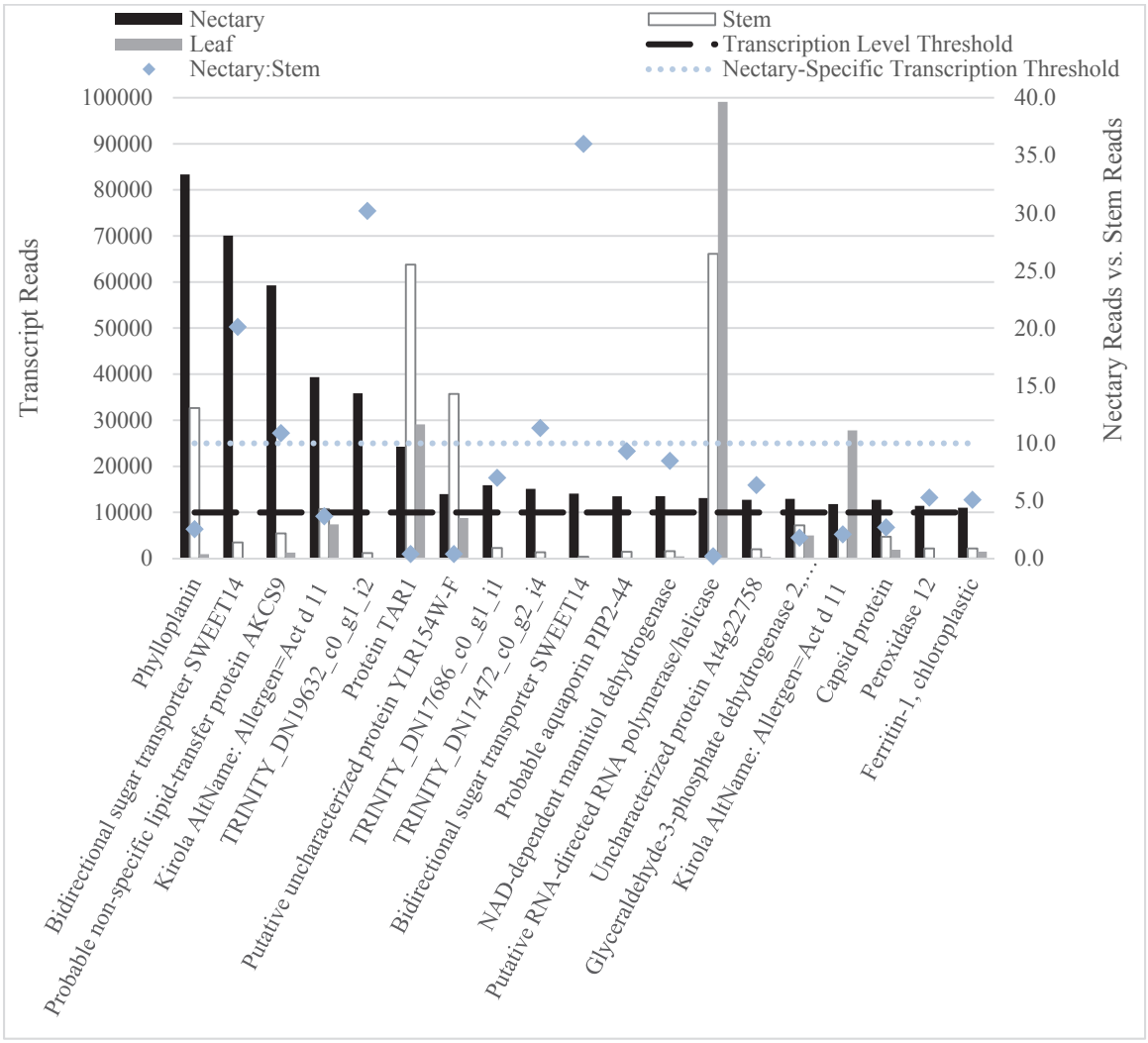


Figure 3.4. Transcripts with Most Transcription Nectaries. The 19 mRNA with the most reads in nectary tissue. Transcription rates in nectary, stem and leaf tissue are displayed for comparison.

Isolating a Native Nectary-Specific Impatiens Promoter Element

Positive Control iPCR with rbcL

An iPCR experiment was performed using the known sequence for the RuBisCO large subunit (rbcL) gene for impatiens and successfully isolated the upstream region of rbcL. To begin a set of 5'-facing nested primers were designed

near the 5' end of the sequence. Additionally, two internal primers were designed. The SnaBI restriction site was identified as a unique restriction site in the rbcL sequence, and thus was selected as the enzyme used for digesting impatiens genomic DNA (gDNA). All of these features are illustrated in Figure 3.5.

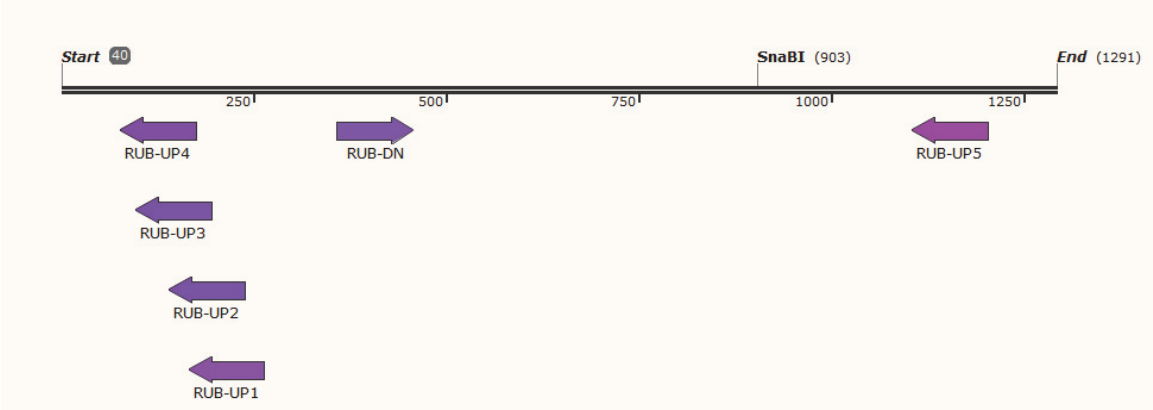


Figure 3.5 Primers and Restriction Enzyme Site for rbcL iPCR. Primers RUB-UP1, UP2, UP3, and UP4 are nested primers located at the 5' end of the sequence and face in the 5' direction. RUB-DN is both an internal primer and used as the only 3' facing primer for iPCR. RUB-UP5 is an internal primer to test for positive digestion of gDNA with SnaBI. Primers are not to scale for illustration purposes, and thus the nested primers at the 5' end do not actually overlap. See Figure B.1 for the actual locations and sequences of all the primers in this Figure.

Two rounds of iPCR, utilizing the nested primers, were performed and produced a PCR product that was submitted for sequencing and matched to the upstream rbcL sequence deposited on GenBank. gDNA that had been digested with SnaBI and circularized (lgDNA-SnaBI, because it is ligated genomic DNA cut with SnaBI) was used as a template for iPCR amplification. The circularized template was calculated to be at least 903 bp in length, as shown in Figure 3.5.

In the first round, primers RUB-UP1 and RUB-DN were used produced three distinct bands in agarose gel that were the minimum required size. These bands can

be seen in Figure 3.6a, and are marked with asterisks. In an attempt to establish a custom protocol for iPCR on impatiens gDNA, different PCR conditions were tested.

Different amounts of template (lgDNA-SnaBI) were tested to determine an optimal amount for iPCR. Since the entire gDNA was present, the large amount of template used would ensure sufficient copies of the target circularized DNA containing rbcL were present. The amount of template tested were 30 ng, 3 ng, and 300 pg. PCR reactions using 30 ng and 3 ng generated visible PCR products between 1.5 kb and 2 kb, while 300 pg did not generate any visible PCR products. Using 3 ng of template produced a large, novel band around 6 kb that was not present when using 30 ng of template.

Additionally, PCR additives that have been shown to enhance target-specific amplification were tested. These included 2 % DMSO, 6 ng/µl BSA, and the NEB high GC-content additive (1X final concentration). Additives were only tested using 30 ng of template. DMSO appeared to reduce overall amplification in first round iPCR. The GC enhancer appeared to function similarly to DMSO by reducing overall amplification. BSA slightly reduced PCR amplification but revealed two distinct PCR products that were not visible in the reaction without BSA.

Four PCR bands on agarose gel were purified as indicated with asterisks in Figure 3.6a, diluted 10x, and used as a template in a 2nd round of iPCR, this time using the nested primer rUP2 with the same downstream primer, rDN. This 2nd round resulted in PCR products for all 4 templates isolated from the 1st round. These 2nd round products appeared slightly smaller on agarose gel electrophoresis, as

would be expected by replacing RUB-UP1 for RUB-UP2. In total there appeared to be 5 PCR bands. These different bands are shown in Figure 3.6b.

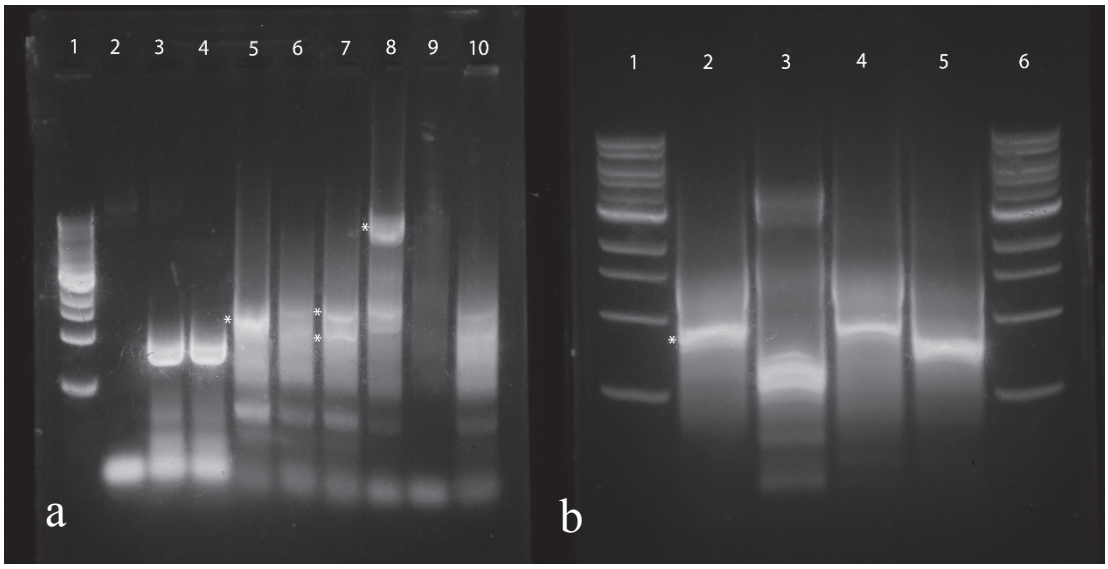


Figure 3.6. Agarose gels for 1st and 2nd round iPCR for rbcL. A) Lane 1: 1 kb DNA ladder (in order of ascending size 500 bp, 1 kb, 1.5 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb). Lanes 2-4 are unrelated to the iPCR of rbcL. Lane 5: 30 ng lgDNA-SnaBI with primers [RUB-UP1, RUB-DN]. Lane 6: 30 ng lgDNA-SnaBI plus 2 % DMSO with primers [RUB-UP1, RUB-DN]. Lane 7: 30ng lgDNA-SnaBI plus 300 ng BSA with primers RUB-UP1, RUB-DN]. Lane 8: 3 lgDNA-SnaBI with primers [RUB-UP1, RUB-DN]. Lane 9: 300 pg lgDNA-SnaBI with primers [RUB-UP1, [RUB-DN]. Lane 10: 30 ng lgDNA-SnaBI plus 1X NEB GC-enhancer® with primers [RUB-UP1, RUB-DN].
B) 2nd round iPCR. Lanes 1 and 6 are DNA ladders of the corresponding size as those mentioned in A. Lane 2 used the marked band from lane 5 in 1st round iPCR as a template. Lane 3 used the marked band from lane 8 in 1st round iPCR as a template. Lane 4 used the larger marked band from lane 7 in 1st round iPCR as a template. Lane 5 used the smaller marked band from 1st round iPCR as a template.

The 2nd round iPCR product that corresponded to the 1st round of 30 ng template, no additives, was gel purified and submitted for sequencing using primer RUB-UP3. The sequence from this product matched for 470 consecutive nucleotides, excluding 4 mismatches near the 3' end of the sequence as shown in Figure 3.7. The final 42 nucleotides before the first N, did not match with the sequence deposited in GenBank. Whether this is a genetic difference in the plants used or a sequencing

error is not yet known. A BLAST of the 696 nucleotide of the sequence produced hits with an E value of 0.0 to 73 atpB-rbcL intergenic spacer sequences from the genus *Impatiens*. These BLAST results can be found at http://andmcox19.wixsite.com/thekearneylab/data. The match of the 2nd round iPCR product to the upstream sequence deposited in GenBank established a working protocol for isolating unknown upstream regions of DNA for other impatiens genes.

Figure 3.7. Sequence and Alignment of 2nd round iPCR Product for rbcL. The sequence above is shown in reverse complement from the manner in which it was sequenced. Blue corresponds to nucleotides matching the intergenic region sequence in GenBank. Red denotes a mismatch flanked by matching nucleotides. Grey is the not included in the GenBank sequence and therefore neither matches or mismatches White is a mismatch to the intergenic region sequence.

Inverse PCR of iwSWEET14 Homolog

An iPCR targeting the upstream region of iwSWEET14 identified 671 bp of upstream DNA, which was determined to contain a core promoter that includes a TATA box and several CAAT box elements. Additional promoter motifs were predicted.

There were 3 restriction enzyme sites that were selected as candidates for iPCR: MfeI, BamHI, and SfcI. Two sets of nested primers, one set facing upstream

and one set facing downstream were designed and an additional primer at the 3' end of the transcript facing upstream (Figure 3.8 top).

Initially, an iPCR was carried out for iwSWEET14 using the restriction site BamHI in the same manner as for rbcL. This however produced no visible PCR products for all possible primer combinations, except 14D and 14E which produced a product approximately 550 bp. It then became clear to the author that the genomic sequence for iwSWEET14 might contain intronic sequences that would not have been present in the mRNA sequence.

To test this hypothesis, an exon prediction for iwSWEEt14 was performed. The exon prediction revealed 5 conserved exon regions using the top 8 hits from a protein BLAST of iwSWEET14 (Figure 3.8 bottom). The top 8 hits with reference genomes were analyzed for their specific splice sites. These hits are included in Table 3.1. Notably, all 8 proteins revealed conserved splice sites at all but one codon for the first three quarters of the mRNA. Even more so, the supposed BamHI site was composed of nucleotides spanning the junction between exon 3 and exon 4. The exon gap was in fact the reason for failure of iPCR initially, as an internal sequence of iwSWEET14 confirmed the exon predictions for exon 3, 4 and 5 to be accurate. Primers 14C and 14E produced a PCR product slightly larger than the expected size of 396 bp. The band was gel purified and sequenced using primer C. The reported sequence matched to iwSWEET14, and showed the junctions between exons 3 and 4 and exons 4 and 5 were accurate. The sequence revealed a 146 bp intron between exons 3 and 4 and a 84 bp intron between exons 4 and 5. The annotated internal sequence of iwSWEET14 is shown in Figure 3.9.

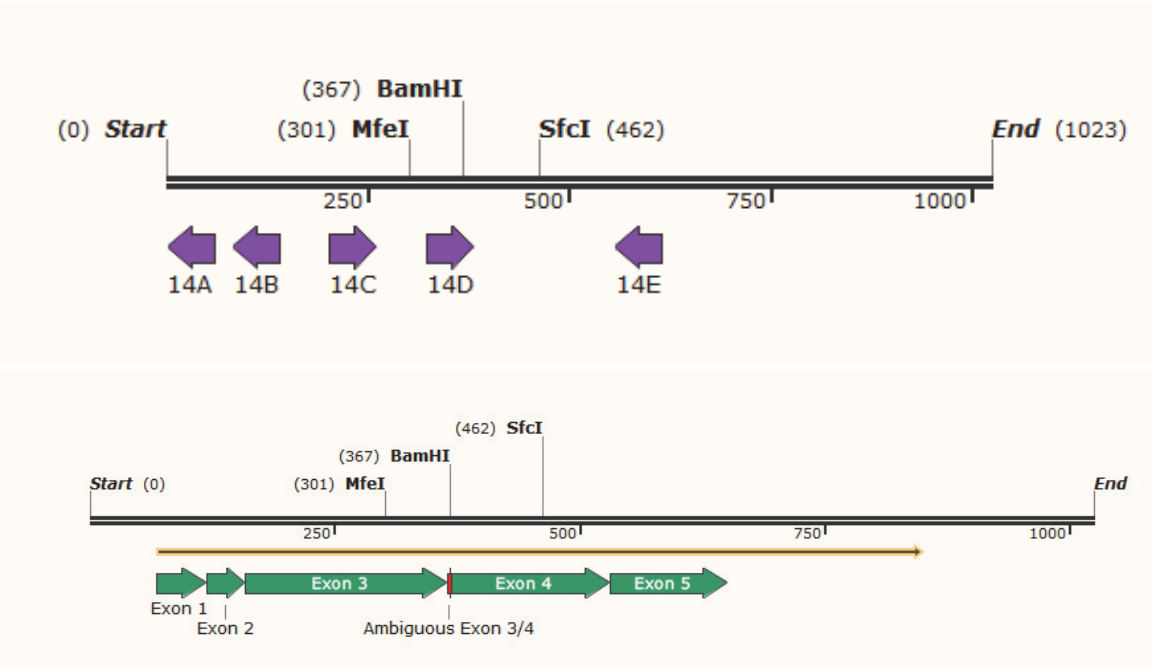


Figure 3.8. Inverse PCR primers and Exon Prediction for iwSWEET14. iPCR primers (top) and predicted exon sequences (bottom) for iwSWEET14. Primers are not to scale, but rather enlarged for visualization. The predicted exon regions are to scale. For instance, primer 14D does not overlap MfeI or BamHI, but exon 4 does overlap BamHI.

Table 3.1. Organisms and Proteins used in Exon Prediction of iwSWEET14

Organism	Protein
Amborella trichopoda	SWEET14 (Predicted)
Populus euphratic	SWEET12-like (Predicted)
Nicotiana tomentosiformis	SWEET12-like (Predicted)
Nicotiana sylvestris	SWEET10-like (Predicted)
Populus trichocarpa	POPTR (Hypothetical)
Vitis vinifera	SWEET14 (Predicted)
Gossypium raimondii	SWEET12-like (Predicted)
Nicotiana sylvestris	N3-like bidirectional sugar transporter (predicted

Assuming the remaining exon sites were predicted correctly, the MfeI and SfcI sites, as shown in Figure 3.8, were selected as good restriction sites for iPCR of iwSWEET14 because they were located in the middle of exons. Two separate

digestions of gDNA using MfeI and SfcI were carried out and the digested fragments were circularized. Two 1st round iPCRs for each lgDNA-MfeI and lgDNA-SfcI were carried out. The primer pairs [14A,14C] and [14B,14C] were used for lgDNA-MfeI. The primer pairs [14A,14C] and [14B,14C] were used with lgDNA-SfcI. As a positive control primer pairs [14A,14C] and [14B,14C] were used in separate standard PCRs that used undigested gDNA for the template.

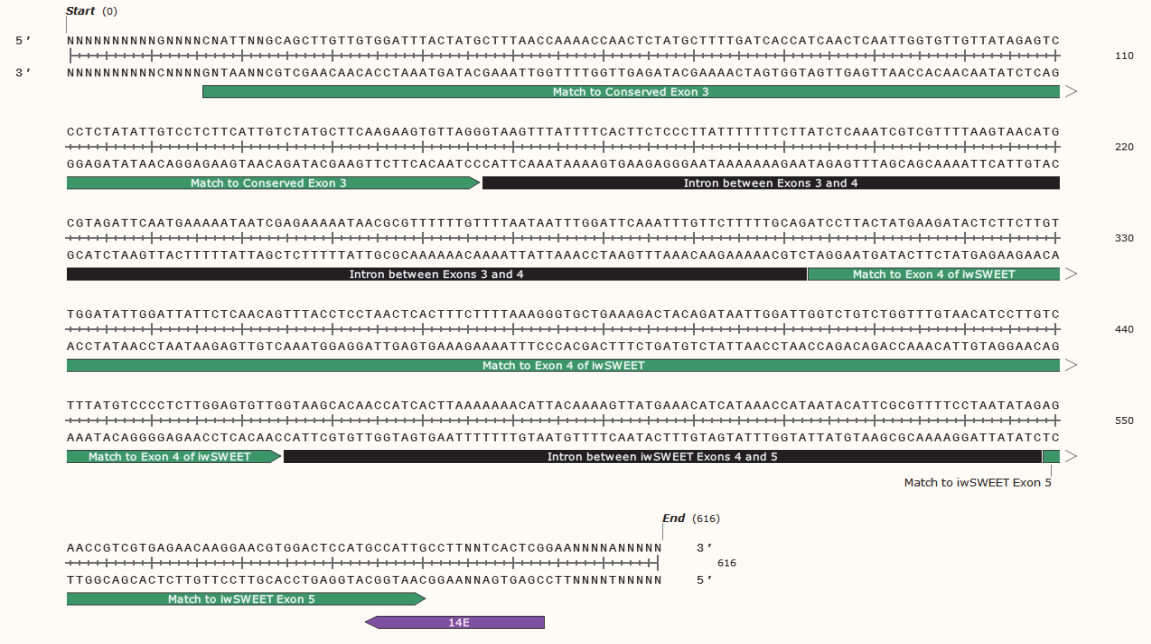


Figure 3.9.Sequence of iwSWEET14 with Introns. Internal sequence of iwSWEET14 showing two identified introns between exons 3:4 and exons 4:5.

DNA electrophoresis of these PCR products did not reveal any prominent bands, but two PCR reactions were chosen for 2nd round iPCR in the hopes that nested primers would facilitate specific amplification of the iwSWEET14 target.

The products from reactions using lgDNA-MfeI with primers [14B, 14C] and lgDNA-SfcI with primers [14A,14C] were used as products for 2nd round iPCR. The entire 1st round iPCR reactions were diluted 100 x. The 2nd round iPCR utilized

nested primers by using primers [14A,14C] for the 1st round lgDNA-MfeI reaction (lane 8 in Figure 3.10a) and primers [14A,14D] for the 1st round lgDNA-SfcI (lane 5 in Figure 3.10a). The 2nd round iPCR produced three distinct products, one for lgDNA-MfeI and two for lgDNA-SfcI. These all excised and submitted for sequencing using both primers 14A and 14D independently for lgDNA-SfcI products and both primers 14A and 14C independently for lgDNA-MfeI.

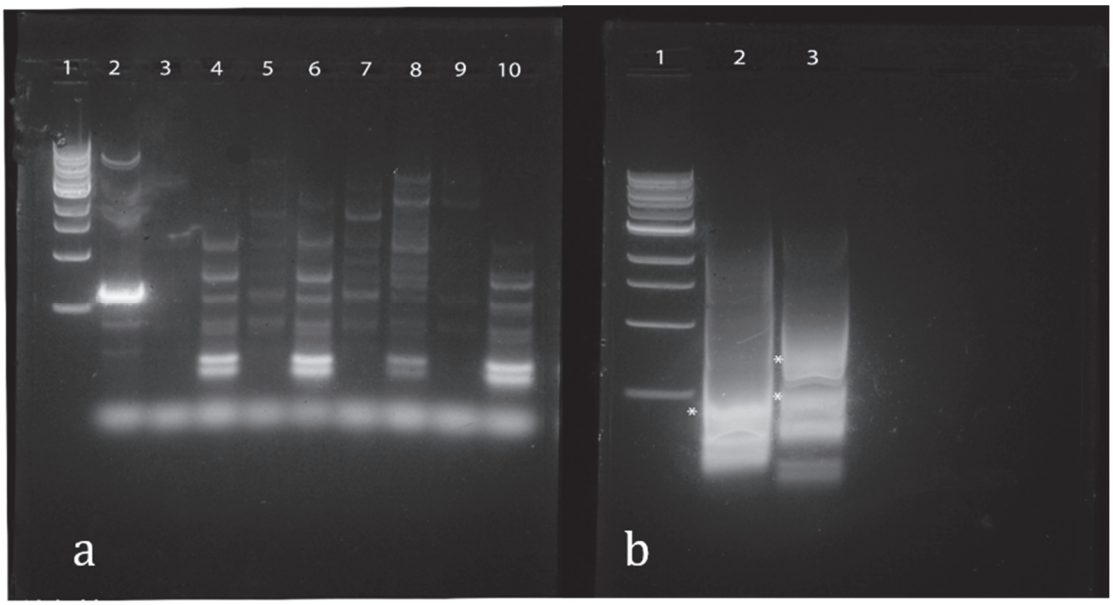


Figure 3.10. Agarose gels for 1st and 2nd round iPCR for iwSWEET14. A) The 1st round of iPCR for SWEET14. Lane 1: 1 kb DNA ladder (in order of ascending size 500 bp, 1 kb, 1.5 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb) Lane 2: lgDNA with primers [14C, 14E[. Lanes 3&4: lgDNA with primers [14A,14C] and [14B,14C], respectively. Lanes 5&6: lgDNA-SfcI with primers [14A,14C] and [14A,14C], respectively. Lanes 7&8: lgDNA-MfeI with primers [14A,14C] and [14B,14C], respectively. Lanes 9&10: lgDNA-HypCH4IV with primers [14A,14C] and [14B,14C] respectively. B) The 2nd round of iPCR for SWEET14. Lane 1: 1 kb DNA ladder with same size bands as mentioned for 1st round iPCR. Lane 2: l g DNA-MfeI with primers [14A,14C]. Lane 3: lgDNA-SfcI with primers [14A, 14D]. Three bands were purified and submitted for DNA sequencing and indicated with asterisks.

The only sequence producing a match to iwSWEET14 came from the product from 2nd round iPCR of lgDNA-MfeI (Figure 3.9b lane 2) using primer 14A. This

sequence revealed a match with the upstream sequence of the IW SWEET14 homolog (Figure 3.11). The sequence returned by Macrogen was 784 bp in length total. Of the 784 bp, 671 bp (excluding multiple non-deciding nucleotides towards the end of the sequence) were upstream of the iwSWEET14 transcript sequence from Trinity. This sequence was inconsistent with the expected size of from the PCR product. However, the use of thick combs, a thick gel, and loading above 100 ng of PCR product into the gel help explain why the expected size differed from the sequence (www.thermofisher.com/us/en/home/life-science/pcr/elevate-pcr-research/agarose-content-with-tips-and-tricks.html). Additionally, the sequenced iPCR product aligned to the 5' region of the SWEET14 transcript, supporting the notion that the seemingly 400 bp iPCR product in Figure 3.10b lane 2, is in fact larger than it appears.

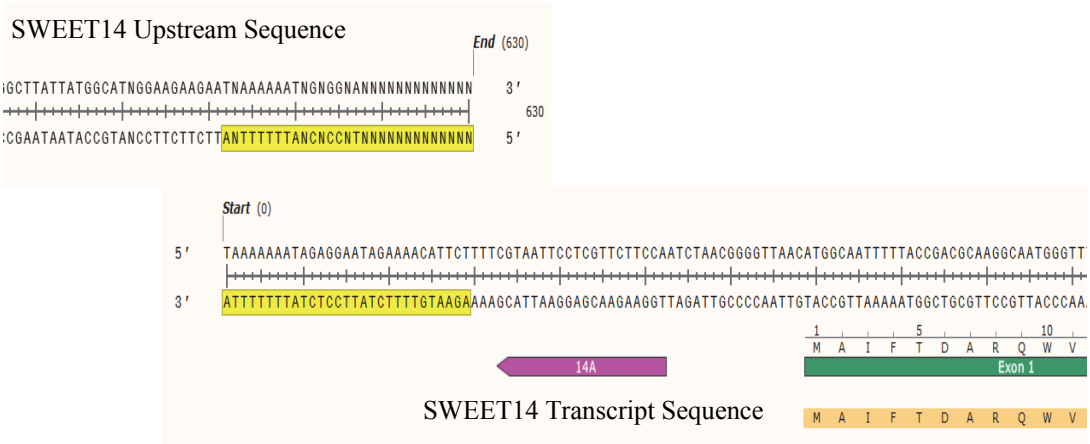


Figure 3.11. Alignment of 2nd round iPCR sequence to iwSWEET14 sequence. The alignment of the upstream region of SWEET14 generated from iPCR to the transcript sequence of SWEET14 from RNAseq.

The 784 bp sequence from iPCR of iwSWEET14 was used in an *in silico* prediction for promoter elements, using the plant promoter prediction software, Plant Care (http://bioinformatics.psb.ugent.be/ webtools/plantcare/html/). The transcription start site (TSS) was determined at the beginning nucleotide from the RNA sequence of iwSWEET14. Figure 3.12 displays the promoter prediction analysis of the upstream region of iwSWEET14.

A core promoter element including several TATA boxes from -5 bp to -58 bp and two CAAT boxes were predicted at -80 bp and -82 bp relative to the TSS of iwSWEET14. The predicted location of both the TATA box and CAAT box are similar with the location of core promoter elements of *Arabidopsis,* wherein a TATA box resides approximately -32 bp of the TSS, and the CAAT box resides approximately -75 bp upstream of the TSS (Molina and Grotewold 2005).

Considerations regarding piwSWEET14. A consideration when working with the SWEET14 promoter element is the fact that RNAseq data for impatiens nectary tissue was from adult, 5-month-old plants. Ergo, the author recommends a follow up experiment using real time PCR to quantify the expression of SWEET14 at age intervals leading up to 5 months. iwSWEET14 was the second-highest transcribed mRNA in nectary tissue from adult impatiens, but that may not be the case for younger plants.

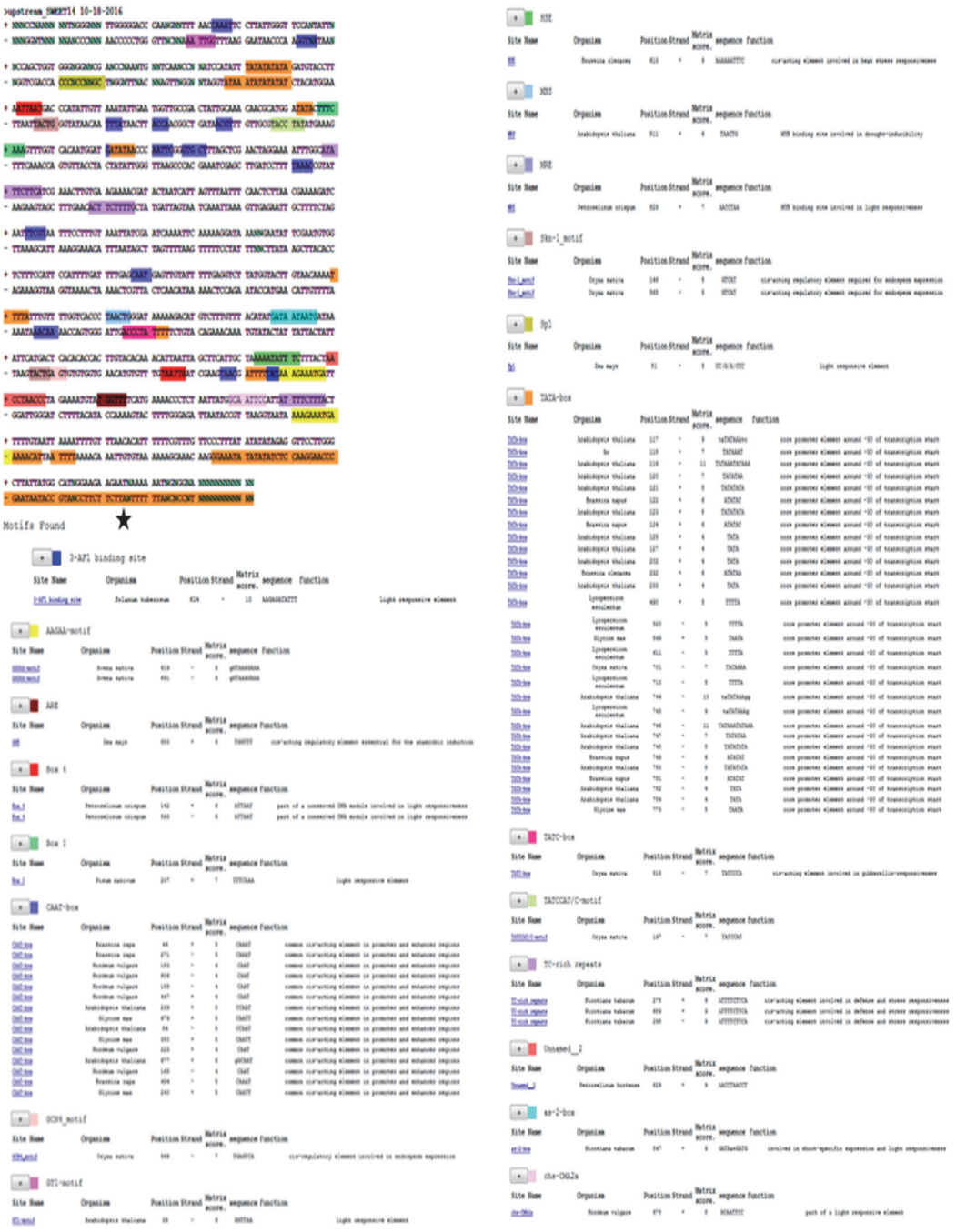


Figure 3.12. A *in silica* analysis of iwSWEET upstream DNA for promoter elements. The 787bp region sequenced with primer A, using the 2nd round iPCR product of genomic DNA digested with MfeI using primers A and C. The black star at the bottom of the sequence denotes the TSS with the start of the SWEET14 transcript residing to the right. The orange highlights denote predicted TATA boxes and the indigo highlights denote predicted CAAT boxes. The closest CAAT box to the TSS is covered by a pink highlight.

Table 3.2. Predicted Promoter Elements of DNA Upstream form iwSWEET14

Motif Sequence	Function	Organism	Position Relative to TSS
AAGAGATATTT	Light response element	*Solanum tuberosum*	-145
TGGTTT	Cis-regulatory element essential element for anaerobic function	*Zea mays*	-109
ATTAAT	Part of a conserved DNA module involved in light responsiveness	*Petroselinum crispum*	-617 -166
CAAAT	Common-acting element in promoter and enhancer regions	*Brassica rapa*	-715
CAAAT		*B. rapa*	-488
CAAT		*Hordeum vulgare*	-576
CAAT		*H vulgare*	-153
CAAT		*H vulgare*	-604
CAAT		*H vulgare*	-312
CCAAT		*A. Thaliana*	-620
CAATT		*Glycine max*	-80
CCAAT		*A. Thaliana*	-705
CAATT		*G. max*	-409
CAAT		*H. vulgare*	-536
gGCAAT		*A. thaliana*	-82
CAAT		*H. vulgare*	-594
CAAAT		*B. rapa*	-265
CAATT		*G. max*	-519
TGAGTCA	Cis-regulatory element involved in endosperm expression	*Oryza sativa*	-193
AAAAAATTC	Cis-acting element in heat stress responsiveness	*Brassica oleracea*	-146
TAACTG	MYB binding site involved in drought-inducibility	*A. thaliana*	-248
AACCTAA	MYB binding site involved in light responsiveness	*Petroselinum cripsum*	-130
GTCAT	Cis-acting regulatory element required for endosperm expression	*Oryza sativa*	-613 -194
CC (G/A) CCC	Light responsive element	*Zea mays*	-678

Table 3.2 Continued

Motif Sequence	Function	Organism	Position Relative to TSS
taTATAAAtc	CORE promoter element around -30 of TSS	*A. thaliana*	-642
TATAAAT		*A. thaliana*	-641
TATAAAATATAAA		*A. thaliana*	-640
TATATAA		*A. thaliana*	-639
TATATATA		*A. thaliana*	-638
ATATAT		*Brassica napus*	-637
TATATATA		*A thaliana*	-636
ATATAT		*B. napus*	-635
TATA		*A. thaliana*	-634
TATA		*A. thaliana*	-632
TATA		*A. thaliana*	-557
ATATAA		*B. oleracea*	-527
TATA		*A. thaliana*	-526
TTTTA		*Lycopersicon*	-269
TTTTA		*esculentum*	-239
TAATA		*L. esculentum*	-210
TTTTA		*Glycine max*	-148
TATATATA		*B. napus*	-9
ATATAT		*A. thaliana*	-8
TATA		*B. napus*	-7
TATA		*A. thaliana*	-5
TAATAT		*A. thaliana*	+14
		Glycine max	
TATCCCA	Cis-acting element involved in gibberellin-responsiveness	*Oryza sativa*	-244
ATTTTCTTCA	Cis-acting element involved in defense and stress responsiveness	*Nicotiana tabacum*	-481 -70 -461
GATAatGATG	Involved in shoot-specific expression and light responsiveness	*Nicotiana tabacum*	-212
GCAATTCC	Part of a light responsive element	*Hordeum vulgare*	-81

Evaluation of the iwSWEET14 upstream region reported here for promoter elements will require *in vivo* experiments to validate if the core promoter, TATA box

and CAAT boxes, are sufficient for high-rate, nectary-specific transcription in transgenic impatiens. If transcription of the transgene is less than expected, or occurs in unexpected tissue, then an additional round of iPCR can be used to sequence further upstream of iwSWEET14.

Impatiens Nectar Protein Characterization and Identification

Impatiens Nectar Protein Separation with SDS-PAGE and 2D SDS-PAGE Gels.

To better characterize the major nectar protein from impatiens that migrates around 21 kDa, two independent experiments were performed. First, impatiens nectar proteins were separated by SDS-PAGE. The protein band from 20 kDa-25 kDa was excised, as shown in Figure 3.13 digested with trypsin, and analyzed using UPLC-MS/MS. Mass spectra for the band were searched using a SWISSPROT databank but no matches were identified.

In the other experiment, a two-dimensional gel electrophoresis of IW nectar proteins revealed that the 21 kDa band is in fact made up of at least 4 different proteins species. Figure 3.13 shows the two-dimensional separation of IW nectar proteins. The protein spots from the 2D gel were excised, digested with trypsin and analyzed using UPLC-MS/MS but the instrument did not detect any noticeable peptide mass peaks. It was supposed that sample was lost in the process of preparing the sample for mass spectrometry, due to an unusually absorbent gel when performing the in-gel trypsin digestion. The gel is usually covered with trypsin digestion solution, but in the case of the 2D gel, it absorbed all the solution, leaving the gel pieces looking like there was no liquid in the tube. It is hypothesized

that digestion with trypsin did not occur from insufficient solution, and thus the undigested proteins were too large to be eluted from the gel pieces following the standard procedure for peptide elution.

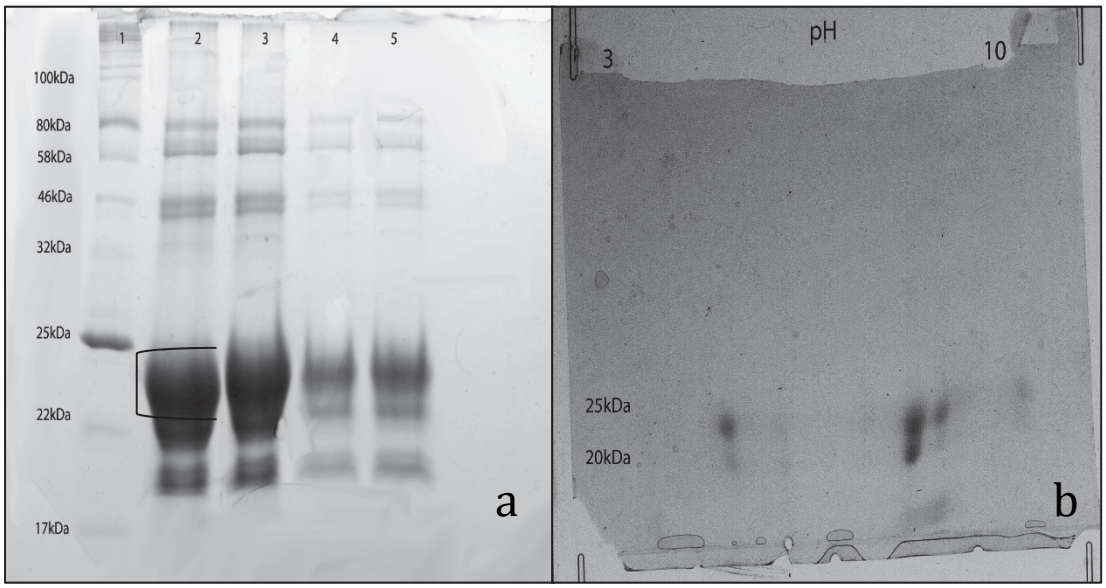


Figure 3.13. SDS-PAGE and 2D Gel Separation of Impatiens Nectar Proteins. a) SDS-PAGE gel of *I. walleriana* nectar proteins prepared by ethanol precipitation. The bracket indicates the portion of the band that was processed for UPLC-MS/MS identification. b) 2D gel of *I. walleriana* nectar proteins.

De novo *prediction of IW23 nectar peptides.*

Since a database search for the impatiens major nectar protein band yielded no hits to any of the 552,259 proteins in SWISSPROT, a data-dependent acquisition (DDA) was performed. The MS/MS spectra from the DDA were then analyzed using Pepseq to predict peptide sequences *de novo*. As a control, MS/MS spectra for two known peptides, ADH and enolase, were analyzed *de novo.*

The *de novo* analysis for the impatiens nectar protein band generated five peptide predictions. Two of these predictions roughly matched to the predicted

protein sequence for T-phylloplanin from the RNA sequence data. The alignments are shown in Figure 3.14. The chromatograms and corresponding MS/MS spectra for the impatiens nectar protein band are included in Appendix D. Additionally, the data for the two *de novo* peptide predictions of the nectar protein band, and the two *de novo* peptide predictions for ADH and enolase, can also be found in Appendix D.

Considerations from de novo *Alignment of Peptides to iwPHYL21.*

Several pieces of evidence support the hypothesis that the 21 kDa nectar protein from Impatiens walleriana is in fact a homolog of trichome exudate protein, T-phylloplanin from Nicotiana. First, RNAseq data indicates that the most highly transcribed mRNA in nectary tissue is translated into a protein that is homologous to T-phylloplanin. Second, SDS-PAGE separation of I. walleriana nectar proteins between 10 kDa-25 kDa mimics the banding pattern found in SDS-PAGE of leaf washes of trichome exudate in tobacco—the bands of T-phylloplanin span 5 different molecular weights and are most intense between 20 kDA and 25 kDa just like nectar proteins from I. walleriana. Third, the de novo sequence of the 20 kDa-24 kDA nectar protein band from I. walleriana roughly matched the protein sequence translated from the RNAseq data for the T-phylloplanin homolog. The difference between the de novo predicted amino acid sequence and the sequence from the RNAseq data could be due to the possible presence of covalent adducts of terpenes and diester sugars with iwPHYL21. The high transcription rate of the non-specific lipid transfer protein in adult impatiens nectary tissue suggests that similar molecules may be secreted in impatiens nectar, and even interacting with iwPHYL21. These covalent adducts are theorized to exist with T-phylloplanin.

Lastly, prior data from [Richardson 2015] (Richardson 2016) demonstrated I. walleriana possessing anti-fungal properties. T-phylloplanin was demonstrated as a powerful anti-fungal protein preventing spore formation on Nicotiana leaves (King 2011). These four independent observations of both I. walleriana nectar proteins and iwPHYL21 support the homology to the T-phylloplanin observed from Nicotiana trichome exudate.

Despite the evidence that the major nectar protein in impatiens is in fact the T-phylloplanin homolog from the RNA sequence data, an unanswered question remains as to why there were inaccurate amino acid sequences from the *de novo* peptide predictions. One possible explanation for the inconsistent mass spectra might be in part to terpenoids and diester sugars binding to iwPHYL21. If terpenoids or diester sugars are covalently bound to iwPHYL21, then they would have disassociated during the fragmentation step when generating the tandem mass spectra. Even then, if the terpenoids or diester sugars were mistaken for amino acids, they would have required identical masses to the b-ions or y-ions of the amino acids they were mistaken for. The answer to this problem is not clear; however isolating the DNA sequence for iwPHYL21 will reveal the true sequence of the protein. If the sequence of iwPHYL21 from impatiens DNA matches the sequence from proposed translation of the RNAseq transcript for iwPHYL21, then further mass spectrometry experiments will be necessary to clarify why the *de novo* prediction was inaccurate.

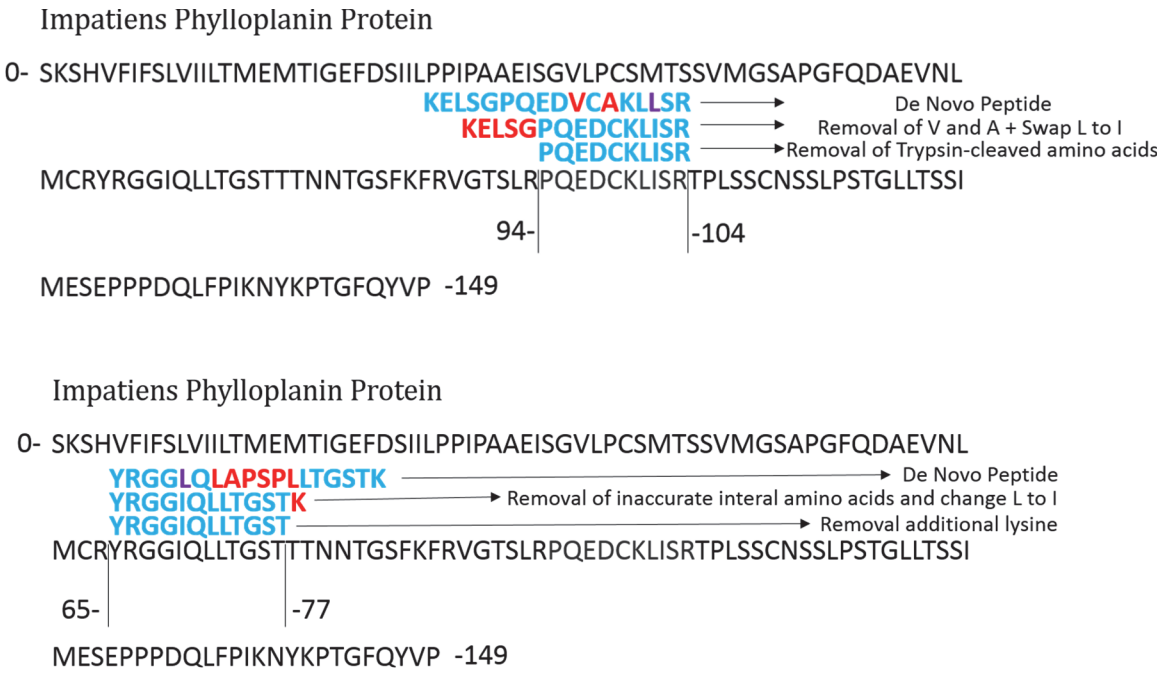


Figure 3.14. Alignment of to *de novo* Predicted Nectar Protein Peptides to iwPHYL.

CHAPTER FOUR

Discussion and Conclusion

Discussion of Results

Confirming the capability of impatiens to produce GUS in both nectary tissue and nectar using the Crabs Claw (crc) promoter element from *Arabidopsis* is an important step towards the establishment of impatiens as a nectar model system. Impatiens transformed with the patCRC-GUS vector, which contained the atCRC promoter driving expression of GUS, produced GUS in non-vascularized nectary tissue and the surrounding stem tissue. patCRC-GUS impatiens also displayed GUS expression in pre-nectary tissue of stem and towards secretory nectaries at the base of leaf tissue (data not included). There was not enough nectar to test for the presence of GUS in nectar colorimetrically during the first few days of nectar secretion, therefore a low-volume approach was utilized relying on the detection of bromoindigo crystals via stereomicroscopy. Bromoindigo is a product when GUS facilitates the cleavage of the bromo-galactoside, X-Gluc. The crystals were easier to detect via microscopy than a solution of bromoindigo in nectar. The presence of these crystals clearly demonstrates the presence of GUS in nectar. According to X-gluc.com, the false positive rate is < 5 % and the false negative rate is < 1 % (http://www.x-gluc.com/xgluc.htm). A colorimetric assay using a known concentration of GUS as a standard will serve as an additional method of verification and quantification of GUS in nectar. A colorimetric assay will be completed, as there is presently a sufficient amount of nectar. The reason this assay was not performed

sooner for there was a fruit fly infestation which was sustained, in large part, by daily nectar secretions from the impatiens. The fruit flies originated from an incubator in a former fruit fly lab that had been dormant for nearly 2 years. This incubator was used to acclimate transgenic impatiens shipped from a collaborator. Apparently, there were undetectable fruit fly eggs that were transported with the transgenic plants to the growth room. These eggs then hatched in the warm and moist temperatures of the growth room. At first there were trace numbers, but once the impatiens began secreting nectar the fruit fly population grew exponentially. Traps were then set, but took nearly two weeks to reduce the number of fruit flies for nectar to once again accumulate. The attractiveness of impatiens nectar to fruit flies serves as evidence for the attractiveness of dipterans in the field, like mosquitoes, and aligns with the previous study by Chen and Kearney (2015) which proved impatiens to be highly attractive to *Aedes* and *Culex* mosquitoes.

The expression of GUS using the atCRC promoter and CARN2 signal peptide proves that impatiens can in fact secrete a transgene into its nectar. This is significant because transgenes have only been expressed in nectar twice—GFP in *Nicotiana langsdorffii* x *N. sanderae* in floral nectar by Dr. Thornburg (unpublished data) and human epidermal growth factor (hEGF) in the floral nectar of the same *Nicotiana* hybrid just mentioned (Helsper et al. 2011)—thus marking the first instance GUS has been expressed in nectar. GUS is significantly larger than either GFP or hEGF, 74 kDa compared to 26.9 kDa and 6.1 kDa, respectively, demonstrating that large foreign proteins can be secreted into nectar. (sizes determined from SWISSPROT sequences calculated by ExPASy ProtParam Tool). When trying to

express protein that are highly-mosquitocidal, it may be necessary to produce concatemer repeats of a transgene to increase toxicity, and a fusion protein may then be necessary for secretion. Knowing that secretion of the 74 kDa GUS transgene is possible, the mosquitocidal transgene can be engineered to a large size and still be secreted.

Future studies will utilize a variety of nectar promoter elements, signal peptides and even fusion partners from impatiens, especially iwPHYL21. Currently, impatiens transformed with the pCWINV-GUS vector, containing the *A. thaliana* CWINV4 promoter driving expression of GUS, have recently become mature enough to perform GUS analysis. A different nectary promoter from *A. thaliana* patSWEET9, driving expression of GUS has just been transformed into impatiens tissue culture by Grace Pruett, a colleague working on the project. The remaining two *Arabidopsis* nectary promoters will soon be tested for expression of GUS in nectar. Of note, the patCRC-GUS vector that successfully demonstrated GUS in nectar from this study utilizes a promoter element from rosids (*Arabidopsis thaliana*), a signal peptide from the core Eudicots (*Dianthus caryophyllus*), and the transcription, translation and cellular chaperones of asterids (*Impatiens walleriana*). This evidence suggests that all clades of Eudicots evolved very similar nectar production mechanisms. Crabs Claw is known to have evolved very early in reference to the emergence of nectar, and is a nectary-initiating transcription factor seemingly conserved in all nectar-producing Eudicots (Lee et al. 2005). In a different study, the CARN2 signal peptide and CARN2 promoter were used to generate expression of a transgene in the nectar of an asterid, a *Nicotiana* hybrid (Helsper et al. 2011). With these two

considerations in mind, it is not surprising that the crc promoter with CARN2 signal peptide successfully generated transcription and localization of GUS. Going further, it may then be possible to create a standard nectar expression cassette that optimizes expression in all Eudicots. This may obviate the need to identify native nectar promoter elements from other species of Eudicots. Without the need to isolate native nectar promoters, transforming the nectar from a wide variety of Eudicots would a rapid process. Notably, this would allow for the efficient, quick production of mosquitocidal plants tailored to specific ecosystems.

On the contrary, there is another explanation for how GUS was successfully expressed in impatiens nectar. Figure 4.1 shows a phylogenetic tree of angiosperms help highlight the evolutionary history of eudicots. The crc promoter is thought to be one of the first evolved genes in nectar development in eudicots (Lee et al. 2005), and thus would theoretically work as a nectar promoter element for any member of the eudicot clade, including impatiens which is an asterid. Additionally, carnation is characterized as a core eudicot as part of the caryophyllid clade. Since phylogenetically, all asterids are also part of core eudicots, asterids may be able to utilize core eudicot promoters and signal peptides, like CARN2, but the reverse may not work. Meaning that, a nectar promoter from asterids may not necessarily work in rosids or even more primitive eudicots.

Whether nectar secretion works universally for all eudicots, or rather works in only forward along the timeline of eudicot evolution remains to be clarified. However, from an evolutionary standpoint it makes sense that nectar would evolve with general characteristics that are shared by all members of the eudicot clade, but

would then specialize at each subsequent taxonomic division. The nectar promoter elements patCWINV4 and patSWEET9 from *Arabidopsis*, provided by Dr. Carter, will soon be able to help elucidate evidence for one of these hypotheses. Again, the CRC promoter is highly conserved in eudicots so its expression in impatiens can be expected.

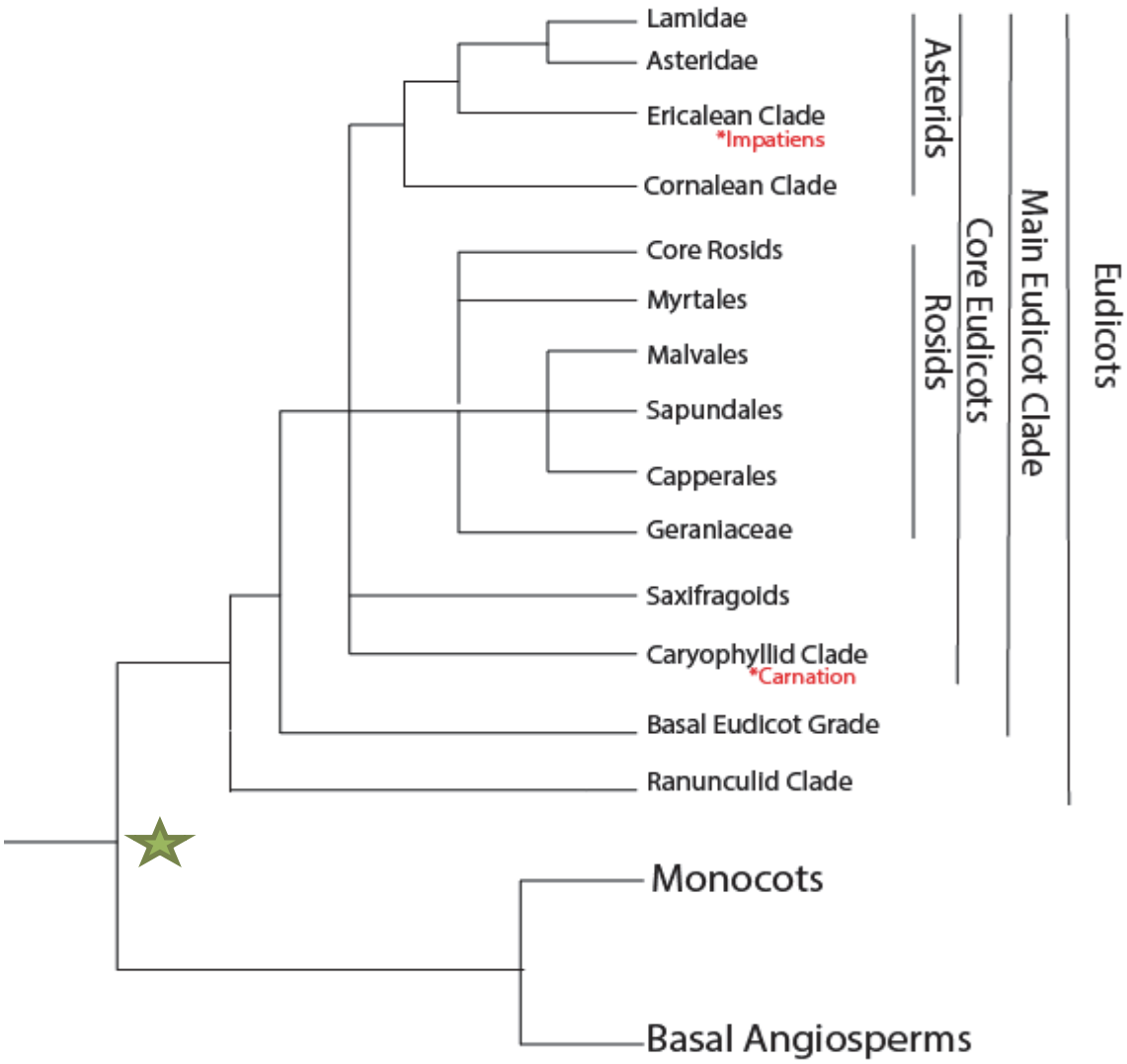


Figure 4.1. This Figure illustrates all angiosperms and the general classification of eudicots. The green star is the estimate evolution of the Crabs Claw transcription factor (Lee et al. 2005). The eudicot phylogeny is based on work from (Magallon et al. 1999).

The evolution of the sucrose transporter SWEET9 has been determined roughly around the development of core eudicots, and it has been shown to produce nectary-specific transcription floral nectar of *N. attenuata,* an asterid (Lin et al. 2014). Therefore, the atSWEET9 promoter may very well work in driving expression of GUS in impatiens. However, CWINV4 has only been characterized in the two rosid genera, *Arabidopsis* and *Brassica*. The lack of a homolog from the RNAseq data for *I. walleriana* suggests that CWINV4 might be evolved in rosids exclusively. Thus, if the atCWINV4 promoter produces GUS in impatiens nectar then the general hypotheses of nectar production would be supported.

The second half of my thesis work sought to obtain an impatiens nectar promoter and identify the major nectar protein in impatiens. In the instance that nectar expression can be optimized using native promoters and fusion partners, my work will provide two ideal candidates—the upstream region containing a core promoter for iwSWEET14 and the sequence for the major nectar protein in impatiens, iwPHYL. The presence of a SWEET14 homolog is significant when considering the possibly efficacy using the *at*SWEET9 promoter in impatiens. Moreover, through the process of characterizing the major nectar protein some insight into the nature between floral nectar proteins and extrafloral nectar proteins may have been gained.

It was originally thought that atSWEET9 would be the most efficient promoter from the three originally received from Carter, *at*CRC, *at*CWINV4, and *at*SWEET9. All three promoters had been used by Dr. Carter for transgene expression in *Arabidopsis* and SWEET9 was reportedly the best. However, the

results from the RNA sequencing of impatiens nectary RNA did not contain a SWEET9 homolog. Additionally, work published in (Lin et al. 2014) contains two phylogenetic analyses of 199 SWEET family proteins, using a neighbor-joining method and a maximum-likelihood method. Both methods map the iwSWEET14 homolog osSWEET14 to a different clade than atSWEET9. Other SWEET homologs were identified in impatiens nectary tissue from RNAseq, but none from the immediate clade as SWEET9 Therefore, there is a possibility that promoter elements of atSWEET9 do not strongly recruit transcription factors in impatiens extrafloral nectaries, due to the seeming difference in utilization of sucrose transporter proteins. Thus, the use of the iwSWEET14 promoter found in this work could recruit transcription machinery more efficiently than atSWEET9 in impatiens nectary cells. As mentioned previously, impatiens transformed with atSWEET9 promoter driving expression of GUS are in tissue culture status, and will be tested within several months.

The characterization and sequence of iwPHYL is important for future work expressing mosquitocidal proteins, due to its potential as a native fusion partner. Fusion partners are often used to increase expression of transgenes, especially when trying to express reactive proteins. Small ubiquitin modifying protein (SUMO) is a fusion protein used in *E. coli* that enables expression of difficult transgenes (Butt et al. 2005) and elastin-like proteins (ELP) facilitates expression of cysteine-stabilized anti-microbial peptides in plant apoplasts (data unpublished by Ghidey and Kearney). In the process of attempting to express mosquitocidal proteins in nectar, a fusion partner might be necessary for expression. Impatiens tissue culture

have already been transformed with a spider toxin, Hv1a. Currently, Hv1a is fused with the signal peptide for a *Juniperus ashei* pollen glycoprotein, named Jun a 3. Jun a 3 signal peptide is known to direct secretion of fused proteins to the apoplast space, the space just outside the plasma membrane (Moehnke et al. 2008). Whether this will result in secretion of Hv1a to nectar remains to be known. However, the discovery of iwPHYL will provide a fusion partner that is known to secrete in impatiens nectar by itself, and therefore may work at directing transgene localization to nectar better than the Jun a 3 signal peptide.

With Hv1a being expressed in upcoming impatiens explants, and the potential pitfalls addressed by the isolation of the work discussed in this thesis, the development of a mosquitocidal impatiens is soon to come. The existence of this plant will generate just as many possibilities as it will questions. Will it still be greatly attractive to mosquitoes? If the attraction level of mosquitoes to impatiens remains similar as published as it did with untransformed plants in (Chen and Kearney 2015) then the possibility to attract and kill mosquitoes in the field could signal the beginning of a new standard in mosquito control. Müller's work on attracting female *Anopheles* in the field with toxic sugar baits demonstrate the immense potential this type of attract and kill technology possesses to control local mosquito populations. Even more so, if impatiens attracts *Anopheles* mosquitoes as well as *Culex* and *Aedes*, then its ability to aid in the control of all mosquito-borne illnesses would be highly useful. One advantage of a mosquitocidal impatiens has over other attract and kill technologies is its ability to naturally increase in size without human support. Impatiens have been observed to grow to prodigious sizes

in our growth room. Impatiens about 1 year old in our growth room conditions grew to be approximately 14ft in length from an 6" soil pot, each plant containing hundreds of nectar-producing nectaries. However, before a mosquitocidal impatiens would be tested in the field, micro and mesocosm attraction and knockdown experiments should be performed with *Aedes*, *Culex*, and *Anopheles*.

Impatiens is not the only plant that possesses the potential to attract and kill mosquitoes, nor would it be beneficially so. While impatiens grows in many areas where mosquito-borne diseases are prevalent, additional plant species could be engineered according to the specific environment. Any plants that also produce nectar, attract mosquitoes, and lend themselves readily to genetic transformation would make good candidates.

Assuming that nectar expression needs to be optimized then this work will provide a template for identifying nectar promoters and nectar proteins, as many of these plants are unlikely to have a reference genome. Transcriptomics, data generated from RNA sequencing, is widely recognized as a tool for working with organism without reference genomes (McGettigan 2013; Shi et al. 2014). The use of mass spectrometry enhances the capability of transcriptomics to identify nectar proteins directly from a sample, at which point the correlating RNA sequencing data reveals locational and temporal information about the transcription of the nectar protein. Therefore, this initial work engineering mosquitocidal impatiens may have a broad application to the creation of hundreds of different mosquitocidal plants.

APPENDICES

APPENDIX A

DNA Sequence Files from Macrogen USA

DNA sequences included data for each nucleotide position. This data can be helpful in assessing the reliability of a sequence. Pictured below are the individual nucleotide data for three sequences mentioned in this study: the rbcL upstream promoter region, the internal sequence of iwSWEET14, and the upstream region of iwSWEET14.

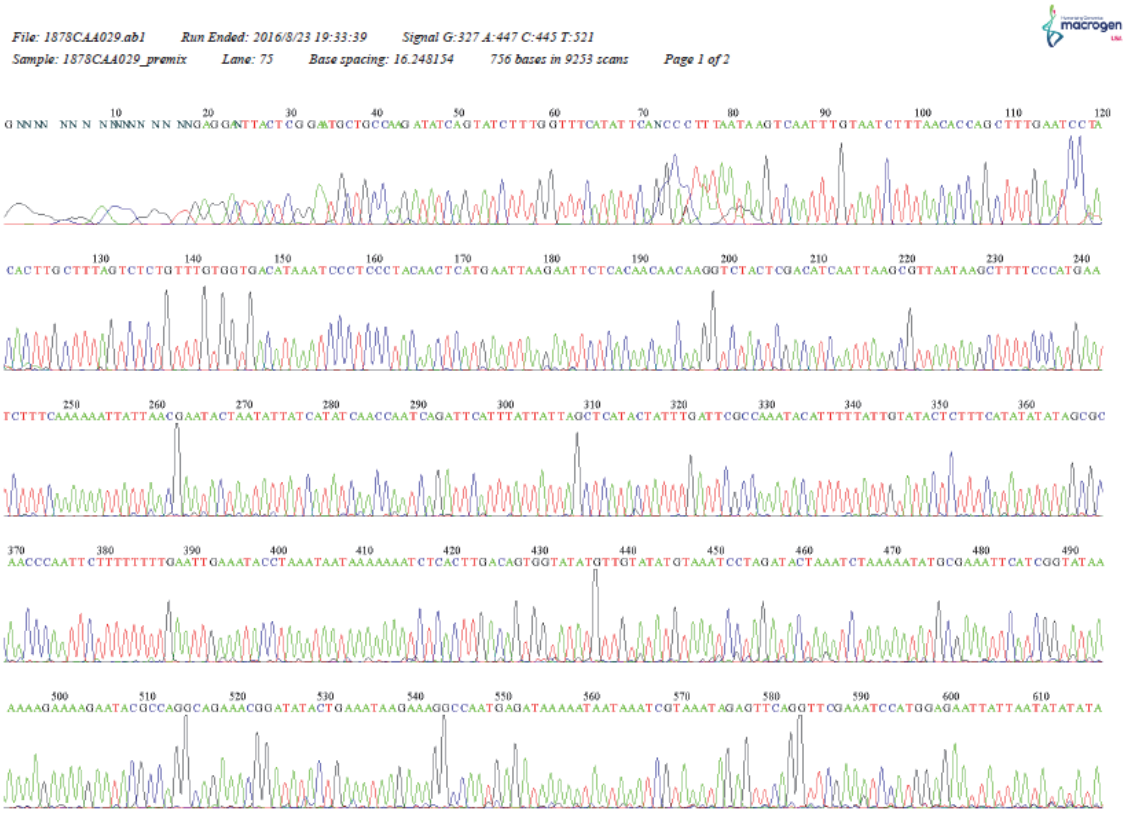


Figure A.2. iwrbcL Upstream Sequence Individual Nucleotide Signal

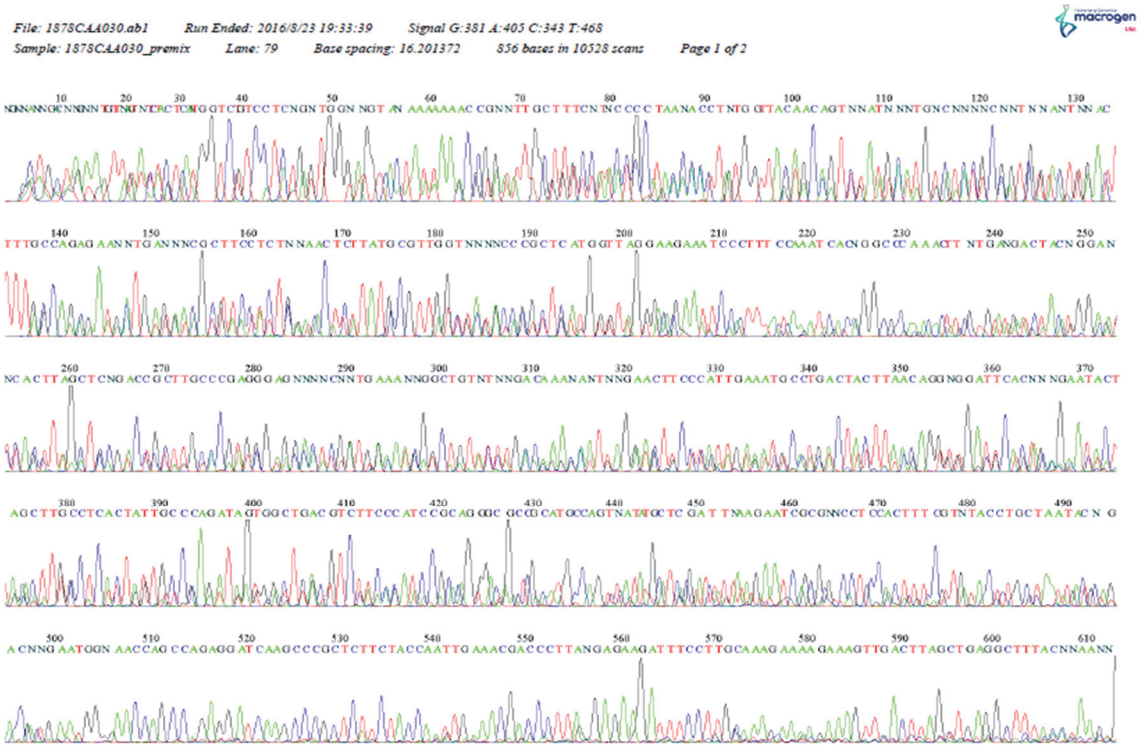


Figure A.2. iwSWEET14 Internal Sequence Individual Nucleotide Signal

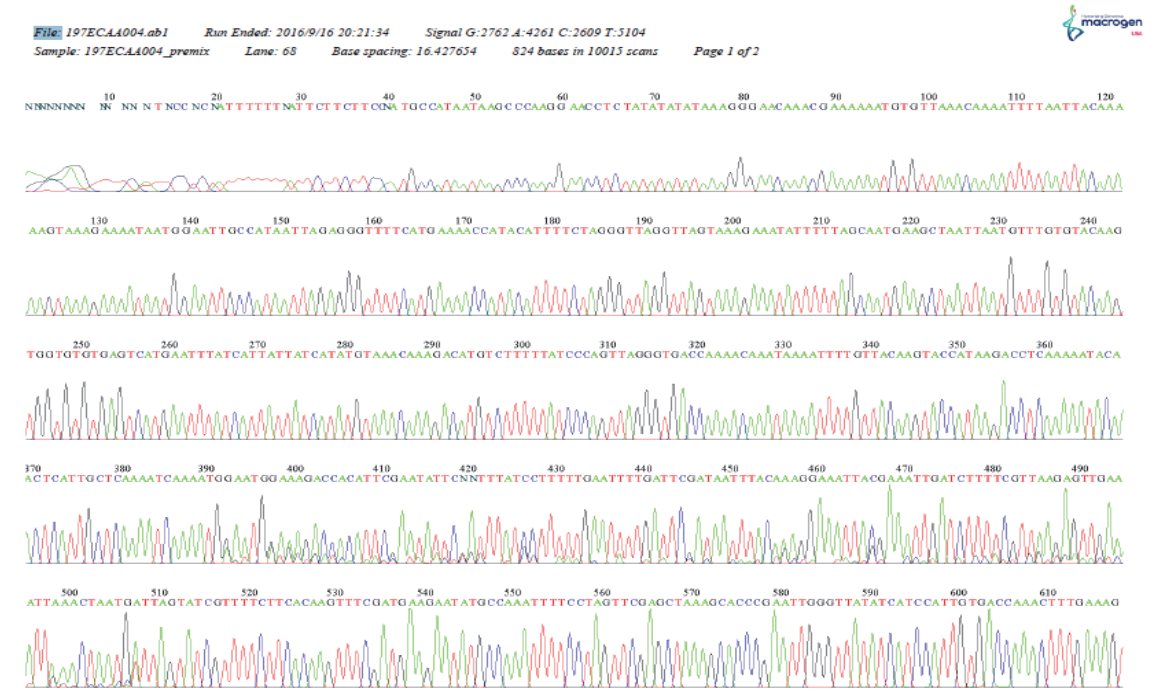


Figure A.3. iwSWEET14 Upstream Sequence Individual Nucleotide Signal

APPENDIX B

Additional Photos of GUS in Impatiens patCRC-GUS Nectaries and Nectar

Nectaries from three patCRC-GUS impatiens were assayed for expression of GUS. The line of patCRC-GUS impatiens with the highest expression of GUS is pictured in the Results section of this thesis. The other two are pcitured here.Additionally, a nectar assay for the detection of GUS was performed on the patCRC-GUS impatiens 030816-T3-2-3, as well as a positive control. Other photos of this assay are pictured in this Appendix.

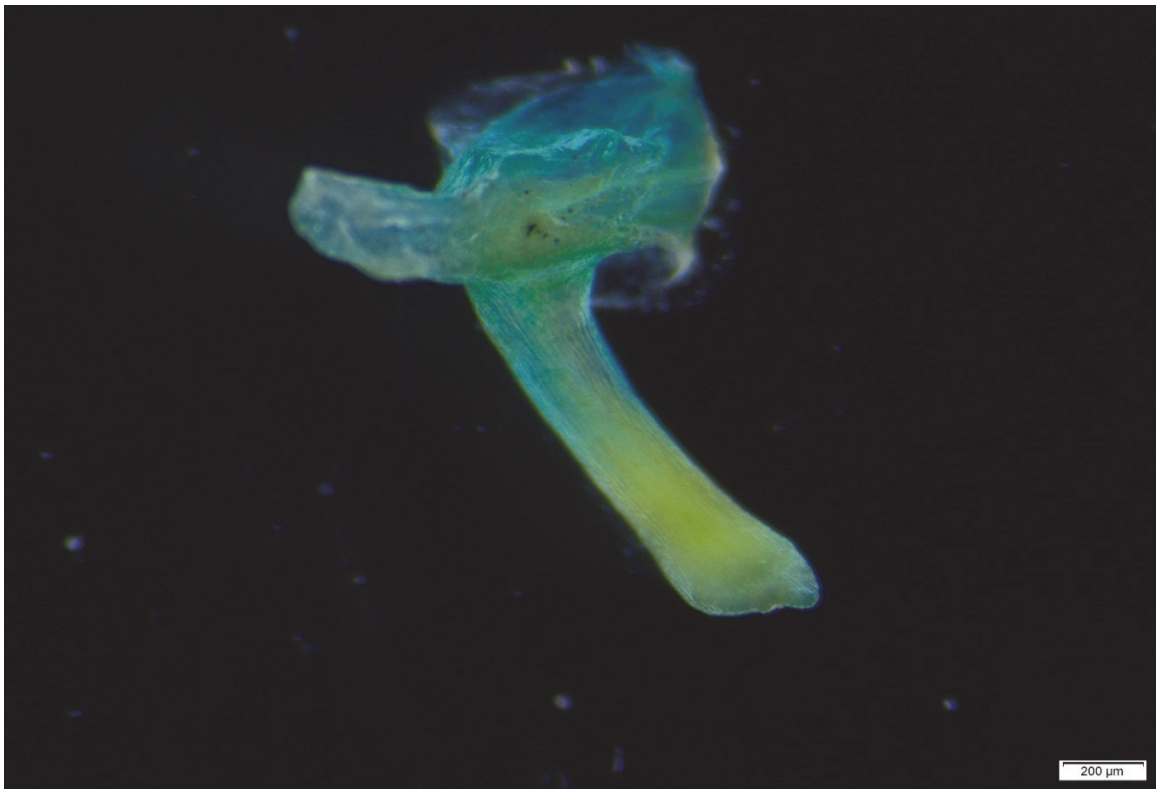


Figure B.1. GUS in the nectary of a patCRC-GUS impatiens nectary from line 030816-T3-2-2

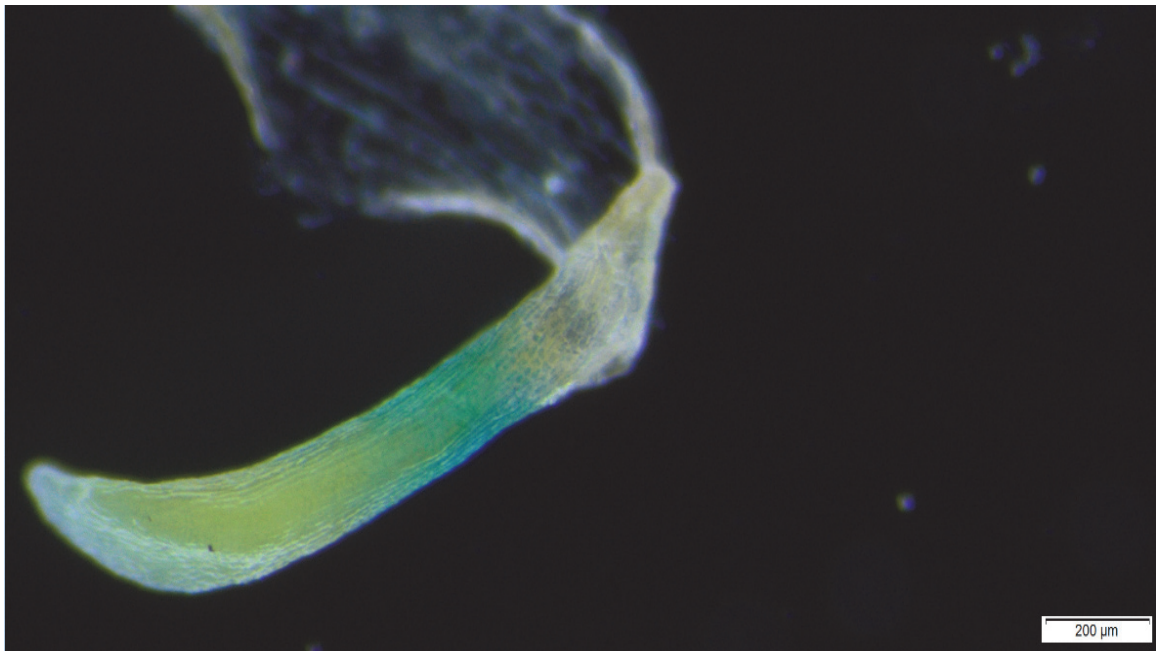


Figure B.2. GUS in the nectary of a patCRC-GUS impatiens nectary from line 030816-T3-2-3

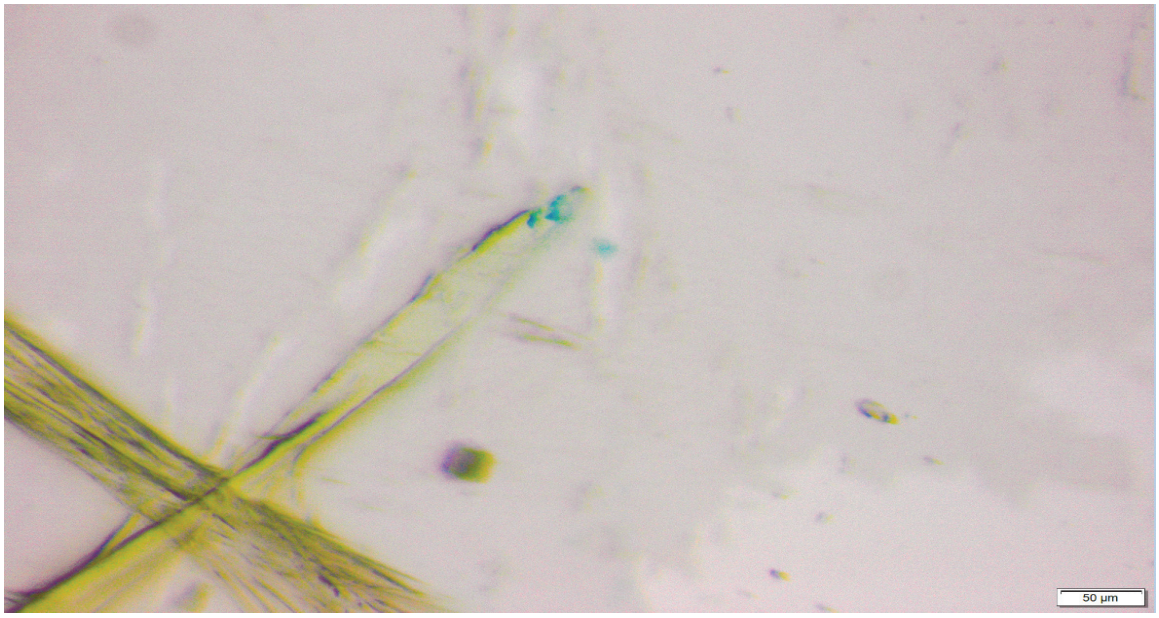


Figure B.3. Photo of Bromoindigo Crystals from patCRC-GUS Impatiens Nectar.

APPENDIX C

Sequences of *Impatiens walleriana* RbcL and SWEET14 Genes

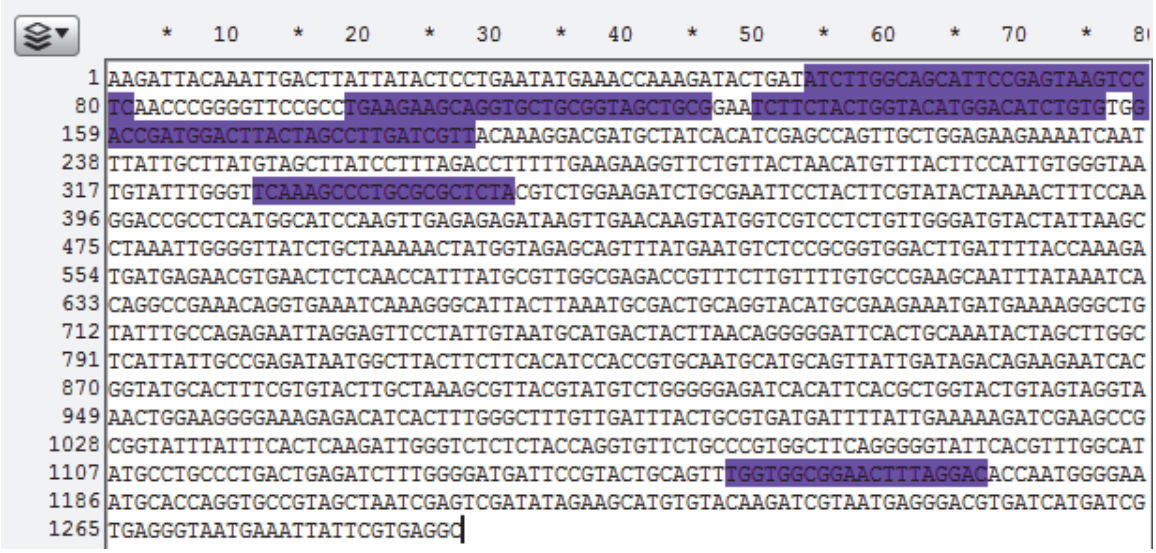


Figure C.1. The RuBiSCo large subunit sequence for *I. walleriana*. GenBank. Accession (AB043508.1). Primers are shown in purple. The 5' end of the sequence is at the top left, with the 3' end being at the bottom-most right.

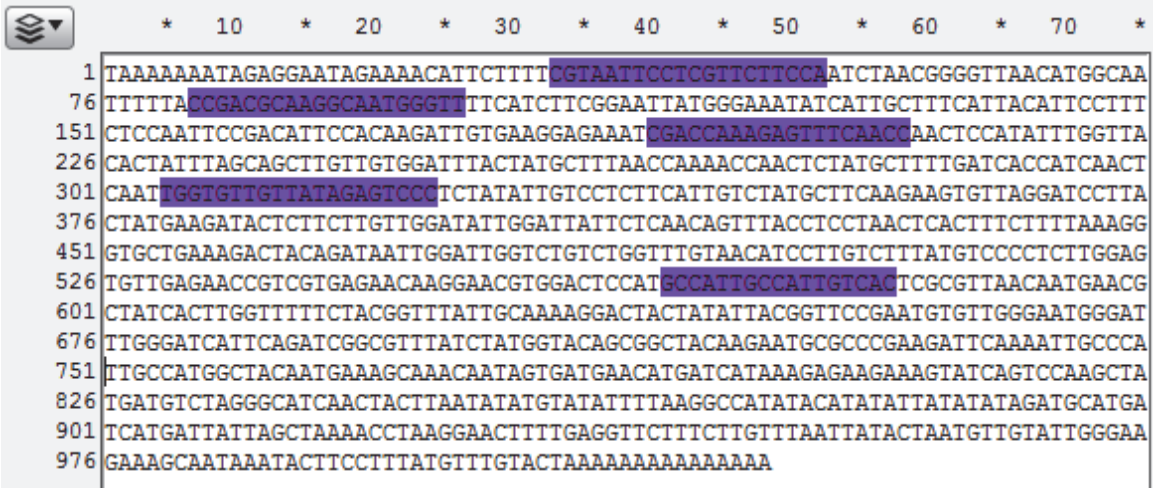


Figure C.2. The sequence from RNA sequencing for *I. walleriana* SWEET14. Primers are shown in purple. The 5' end of the sequence is at the top left, with the 3' end being at the bottom-most right.

APPENDIX D

Additional Chromatograms,Mass Spectra, Databank Search Parameter, *De Novo* Peptide Predictions, and DDA Parameters

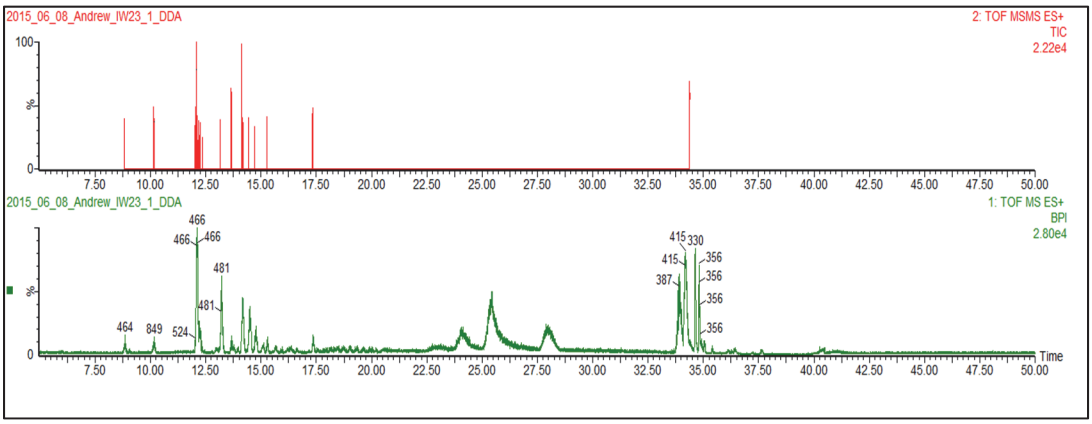


Figure D.1. The Chromatogram and Corresponding Mass Spectra for Impatiens Nectar Protein Band.

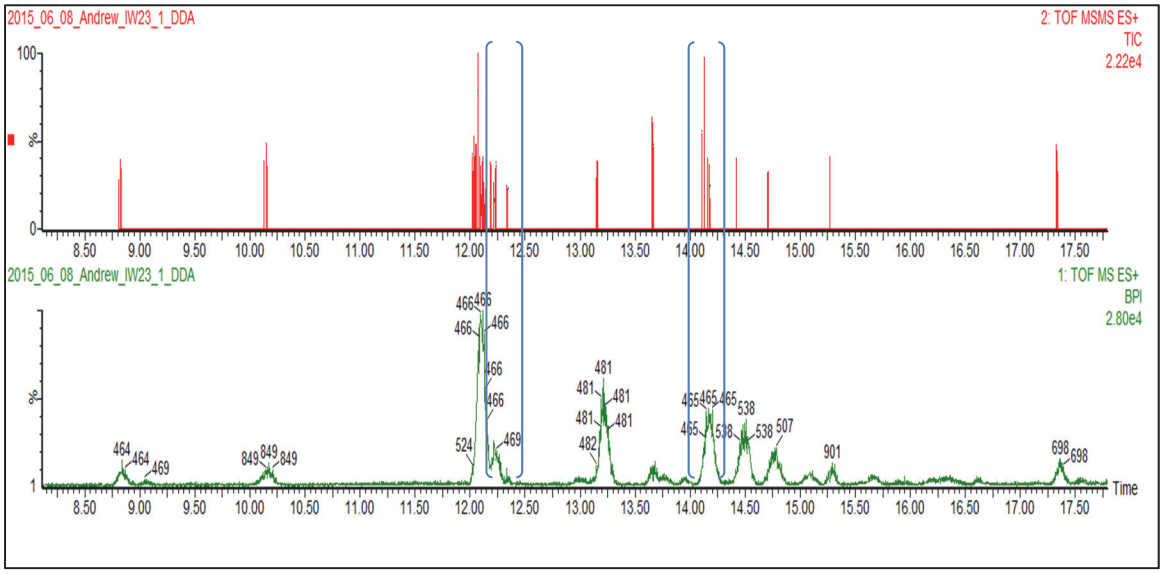


Figure D.2. The MS/MS Peaks Used for *De Novo* Sequencing and Their Corresponding Chromatogram. The parent ion with m/z 469 corresponds with iwPHYL21 peptide 1 and the parent ion with m/z 465 corresponds with iwPHYL21 peptide 2.

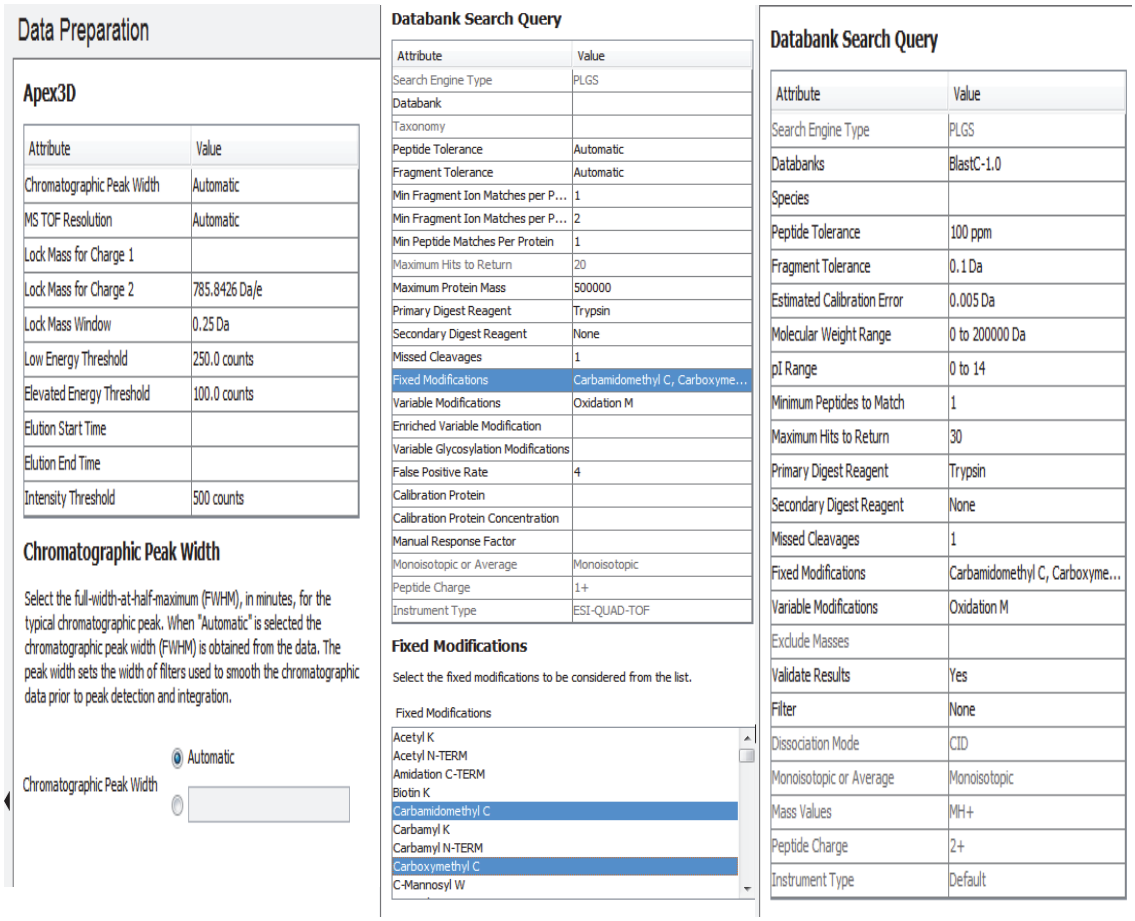


Figure D.3. Data Preparation and Databank Search Query Parameters.

Observed MW: 1415.7103 Precursor ion charge state: 1
M/z tolerance: 0.30 Intensity threshold: 283 (0.750%)

a	87.06	144.08	241.13	342.18	441.25	570.29	669.36	798.40	911.48	1012.53	1113.58	1242.62	1370.72
	---	-0.00	---	---	---	---	---	---	---	---	---	---	---
b	115.05	172.07	269.12	370.17	469.24	598.28	697.35	826.39	939.48	1040.53	1141.57	1270.62	1398.71
	0.00	0.00	0.01	-0.00	-0.00	-0.00	-0.01	---	---	---	---	---	---
	Asn	Gly	Pro	Thr	Val	Glu	Val	Glu	Leu	Thr	Thr	Glu	Lys
	98	98	100	100	100	100	100	100	100	100	100	100	100
y	1416.72	1302.68	1245.66	1148.60	1047.56	948.49	819.45	720.38	591.34	478.25	377.20	276.16	147.11
	---	---	-0.02	-0.01	-0.00	-0.01	-0.00	-0.01	-0.01	-0.01	-0.00	-0.00	0.00
z	1399.69	1285.65	1228.63	1131.57	1030.53	931.46	802.42	703.35	574.31	461.22	360.17	259.13	130.08
	---	---	---	---	---	---	---	0.01	---	-0.03	---	---	-0.01

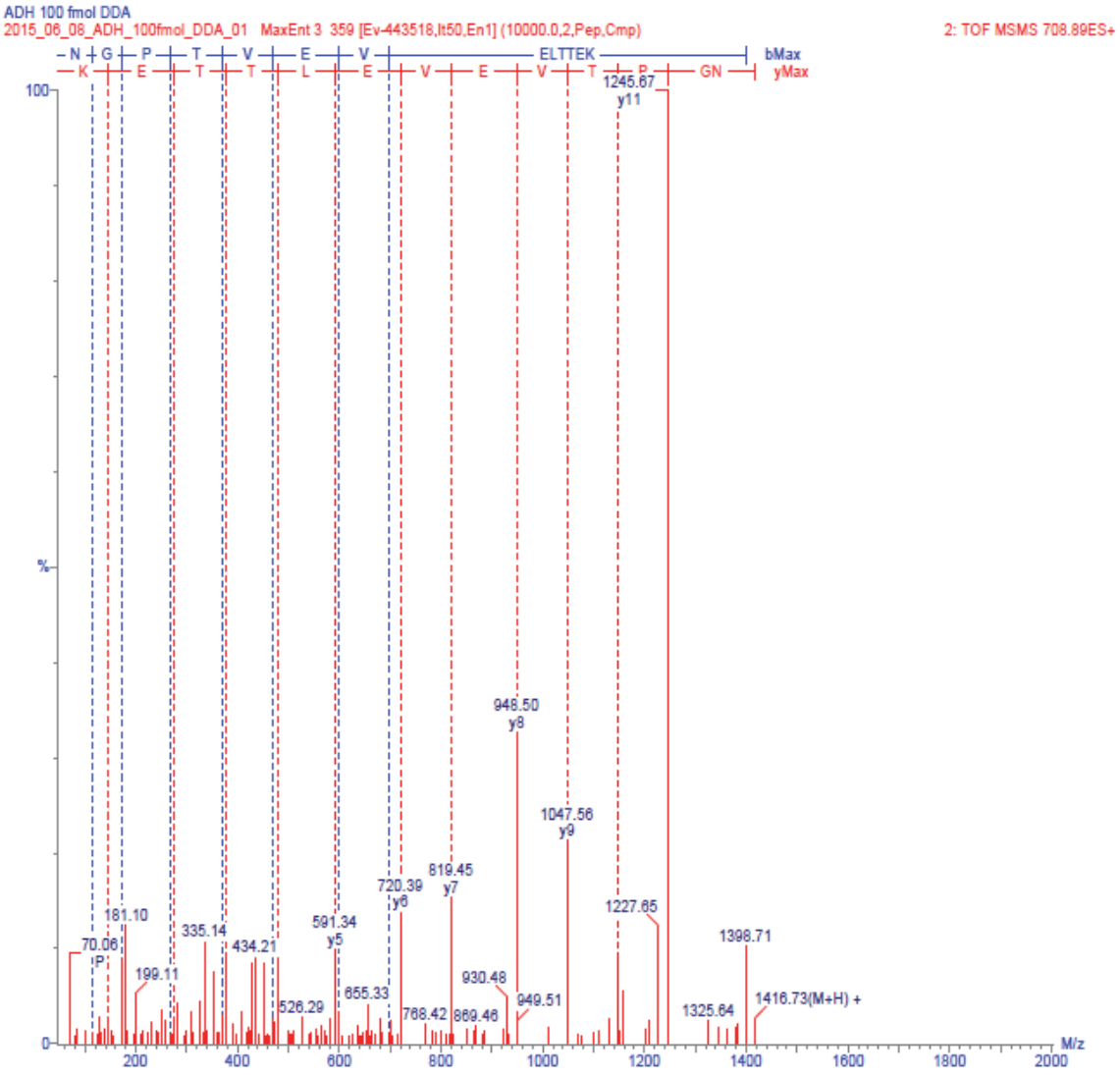


Figure D.4. *De Novo* Prediction of Enolase Peptide

BioLynx - Peptide Sequencing Report Page 2

PepSeq file:
Printed: Mon Dec 07 13:06:28 2015

Observed MW: 1354.6367 Precursor ion charge state: 2
M/z tolerance: 0.30 Intensity threshold: 12 (0.750%)

a	30.03	133.04	236.05	293.07	380.11	495.13	594.20	741.27	855.31	983.37	1082.44	1181.51	1309.60
	---	0.00	-0.01	-0.01	---	---	-0.00	-0.01	0.03	0.17	---	---	---
b	58.03	161.04	264.05	321.07	408.10	523.13	622.20	769.26	883.31	1011.37	1110.43	1209.50	1337.60
	---	-0.18	-0.00	-0.00	0.00	-0.00	-0.00	-0.03	---	0.00	-0.01	-0.28	---
	Gly 48	Cys 48	Cys 94	Gly 94	Ser 100	Asp 100	Val 100	Phe 100	Asn 100	Gln 100	Val 100	Val 100	Lys 100
y	1355.61	1298.59	1195.58	1092.57	1035.55	948.52	833.49	734.42	587.35	473.31	345.25	246.18	147.11
	---	---	0.00	0.14	-0.01	0.00	0.01	-0.00	-0.00	-0.00	0.17	-0.00	-0.00
z	1338.58	1281.56	1178.55	1075.54	1018.52	931.49	816.46	717.39	570.32	456.28	328.22	229.15	130.08
	---	---	0.00	-0.05	-0.21	-0.01	-0.01	-0.01	-0.01	-0.00	-0.01	0.03	-0.01

ADH 100 fmol DDA
2015_08_13_ADH_100fmol_DDA_01 MaxEnt 3 400 [Ev-298205,It50,En1] (10000.0,2,Pep,Cmp) 2: TOF MSMS 678.33ES+

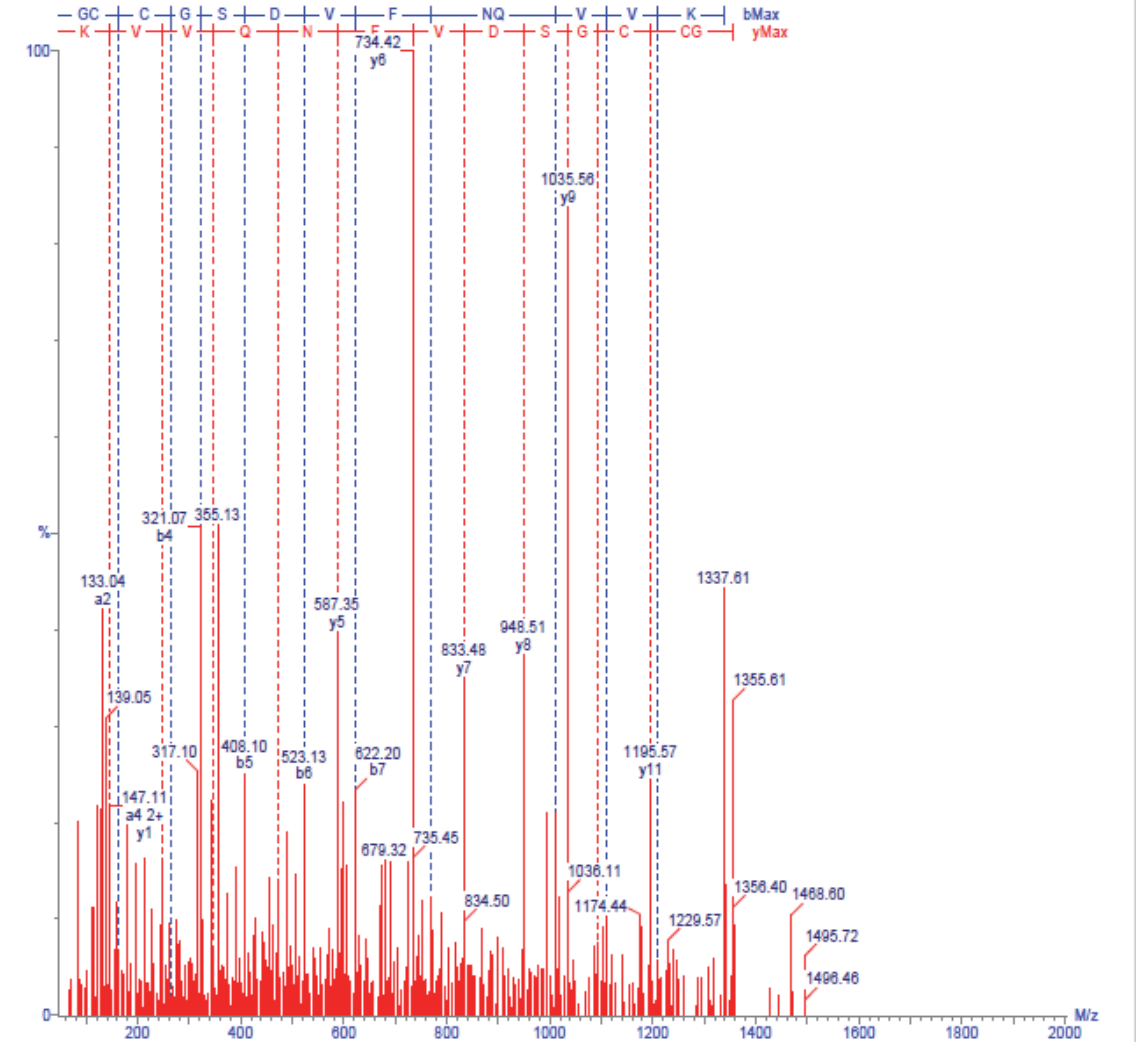


Figure D.5. *De Novo* Prediction of ADH Peptide

BioLynx - Peptide Sequencing Report Page 2

PepSeq file: D:\Andrew\De Novo\IW23_1871_12-183_Nov-18-15.psq
Printed: Mon Dec 07 11:34:50 2015

Observed MW: 1871.9884 Precursor ion charge state: 4
M/z tolerance: 0.30 Intensity threshold: 8 (0.750%)
Modifications: Carboxyamidomethylcysteine (+/-)

a	101.11	230.15	343.23	430.27	487.29	584.34	712.40	841.44	956.47	1055.54	1158.55	1229.58	1357.68
	0.03	0.04	-0.02	0.03	---	---	-0.03	-0.02	---	---	---	---	0.22
b	129.10	258.15	371.23	458.26	515.28	612.34	740.39	869.44	984.46	1083.53	1186.54	1257.58	1385.67
	-0.00	0.01	0.00	0.00	-0.02	0.02	0.01	0.01	-0.01	---	-0.26	---	---
	Lys	Glu	Leu	Ser	Gly	Pro	Gln	Glu	Asp	Val	Cys	Ala	Lys
	68	68	96	97	99	99	93	100	100	100	93	93	100
y	1872.99	1744.89	1615.85	1502.76	1415.73	1358.71	1261.66	1133.60	1004.56	889.53	790.46	687.45	616.41
	---	---	-0.01	---	-0.02	---	---	0.10	---	---	0.02	---	-0.01
z	1855.96	1727.86	1598.82	1485.73	1398.70	1341.68	1244.63	1116.57	987.53	872.50	773.43	670.42	599.38
	---	---	---	---	---	---	---	0.01	---	---	-0.02	-0.02	0.03

a	1470.76	1583.85	1670.88	1826.98
	---	---	---	---
b	1498.76	1611.84	1698.87	1854.97
	---	---	---	---
	Leu	Leu	Ser	Arg
	100	100	100	100
y	488.32	375.24	262.15	175.12
	-0.00	0.00	0.07	---
z	471.29	358.21	245.12	158.09
	-0.00	---	-0.01	-0.01

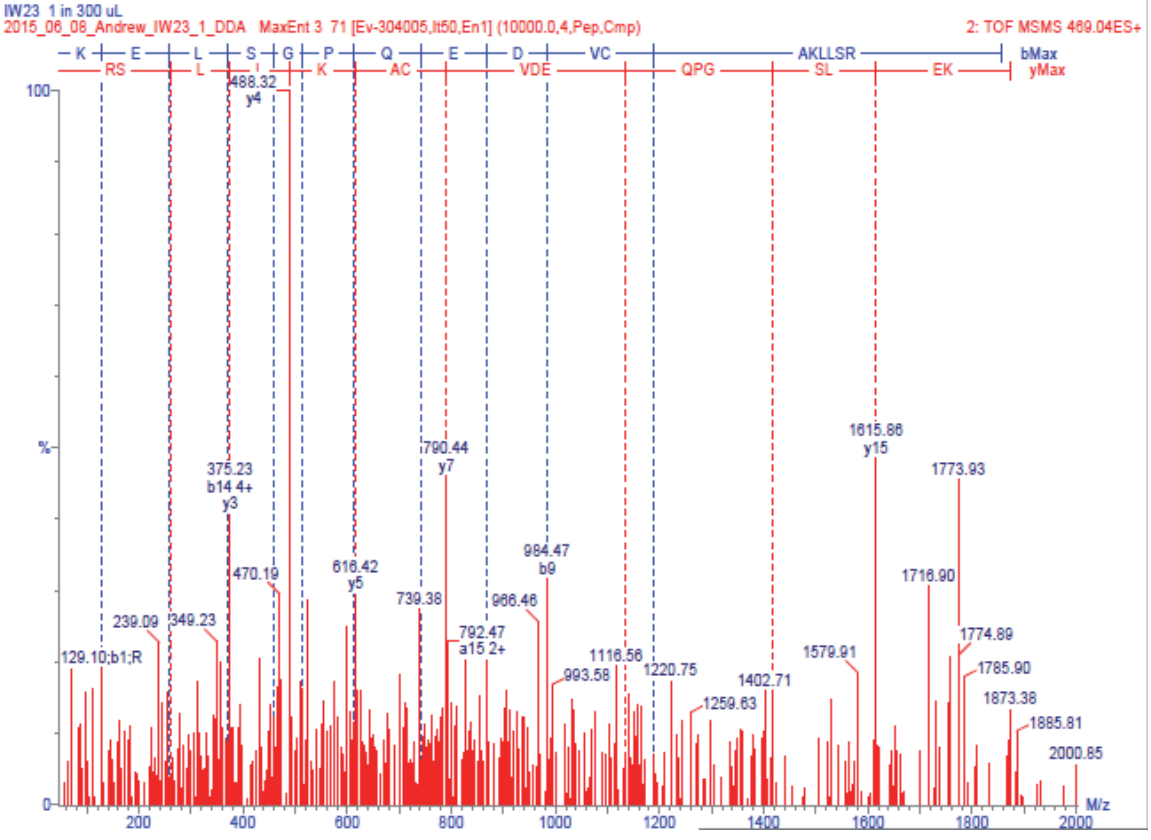


Figure D.6. *De Novo* Prediction of iwPHYL21 Peptide 1.

Observed MW: 1857.9835 Precursor ion charge state: 4
M/z tolerance: 0.30 Intensity threshold: 46 (0.750%)
Modifications: Carboxymethylcysteine (+/-)

a	129.11	292.18	349.20	406.22	519.30	647.36	760.45	859.52	956.57	1055.64	1112.66	1225.74	1338.83
	0.01	---	-0.01	-0.00	-0.00	0.00	-0.01	---	---	---	---	---	---
b	157.11	320.17	377.19	434.22	547.30	675.36	788.44	887.51	984.56	1083.63	1140.65	1253.74	1366.82
	---	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01	---	---	---	---	---	---
	Arg	Tyr	Gly	Gly	Leu	Gln	Leu	Val	Pro	Val	Gly	Leu	Leu
	77	77	100	100	100	100	100	38	38	35	35	97	100
y	1859.08	1702.97	1539.91	1482.89	1425.87	1312.78	1184.73	1071.64	972.57	875.52	776.45	719.43	606.35
	---	---	---	---	---	---	---	---	---	---	0.02	-0.01	-0.00
z	1842.05	1685.94	1522.88	1465.86	1408.84	1295.75	1167.70	1054.61	955.54	858.49	759.42	702.40	589.32
	---	---	---	---	---	---	---	---	---	---	---	---	---

a	1439.87	1496.90	1583.93	1684.98	1813.07
	---	---	-0.26	---	---
b	1467.87	1524.89	1611.92	1712.97	1841.06
	---	---	---	---	---
	Thr	Gly	Ser	Thr	Lys
	100	100	100	100	100
y	493.26	392.21	335.19	248.16	147.11
	-0.01	-0.00	-0.00	-0.00	-0.00
z	476.23	375.18	318.16	231.13	130.08
	---	-0.04	---	---	-0.00

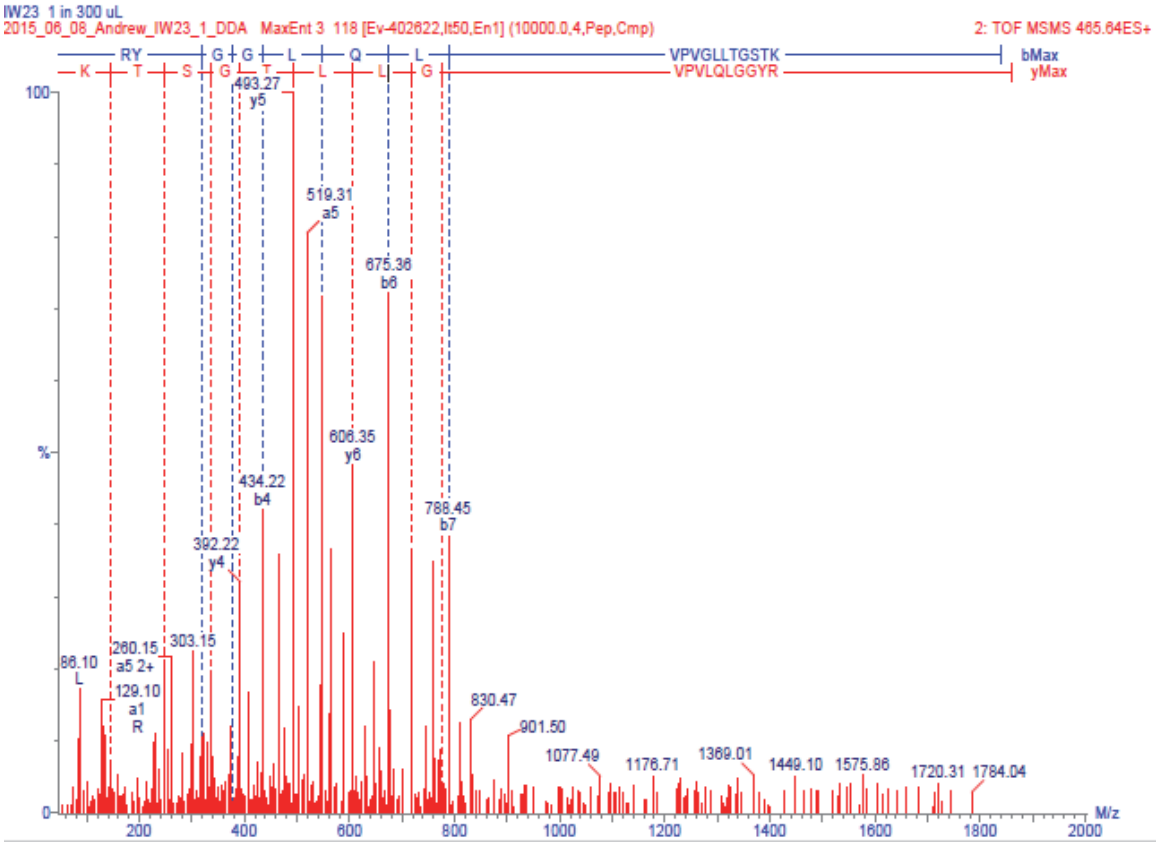


Figure D.7. *De Novo* Prediction of iwPHYL21 Peptide 2.

Mass spectrometry acquisition parameters can be found under “DDA Data Acquisition Parameters_Cox, Andrew” at http://andmcox19.wixsite.com-/thekearneylab/data.

BIBLIOGRAPHY

Agarose Gel Electrophoresis Tips & Tricks. [accessed 2016a Oct 20]. https://www.thermofisher.com/us/en/home/life-science/pcr/elevate-pcr-research/agarose-content-with-tips-and-tricks.html

Baker H, Baker I. 1990. The Predictive Value of Nectar Chemistry to the Recognition of Pollinator Types. Isr. J. Bot. 39:157–166.

Beier JC, Müller GC, Gu W, Arheart KL, Schlein Y. 2012. Attractive toxic sugar bait (ATSB) methods decimate populations of Anopheles malaria vectors in arid environments regardless of the local availability of favoured sugar-source blossoms. Malar. J. 11:31.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. Am. J. Bot. 97:1296–1303.

Bentley B, Elias TS. 1983. The Biology of Nectaries. Columbia University Press.

Brewster JL. 1994. Onions and other vegetable alliums. Wallingford, Oxon, UK: CAB International (Crop production science in horticulture.

Bronstein JL. 1998. The Contribution of Ant-Plant Protection Studies to Our Understanding of Mutualism1. Biotropica 30:150–161.

Butt TR, Edavettal SC, Hall JP, Mattern MR. 2005. SUMO fusion technology for difficult-to-express proteins. Protein Expr. Purif. 43:1–9.

Carter C, Graham RA, Thornburg RW. 1999. Nectarin I is a novel, soluble germin-like protein expressed in the nectar of Nicotiana sp. Plant Mol. Biol. 41:207–216. [accessed 2015 Aug 28]

Carter C, Thornburg RW. 2000. Tobacco Nectarin I purification and characterization as a germin-like, manganese superoxide dismutase implicated in the defense of floral reproductive tissues. J. Biol. Chem. 275:36726–36733.

Carter C, Thornburg RW. 2003. The nectary-specific pattern of expression of the tobacco Nectarin I promoter is regulated by multiple promoter elements†. Plant Mol. Biol. 51:451–457.

Carter C, Thornburg RW. 2004. Is the nectar redox cycle a floral defense against microbial attack? Trends Plant Sci. 9:320–324.

Carter CJ, Thornburg RW. 2004. Tobacco Nectarin V Is a Flavin-Containing Berberine Bridge Enzyme-Like Protein with Glucose Oxidase Activity. Plant Physiol. 134:460–469.

Carter CJ, Thornburg RW. Tobacco Nectarin III is a Bifunctional Enzyme with Monodehydroascorbate Reductase and Carbonic Anhydrase Activities. Plant Mol. Biol. 54:415–425.

Chen Z, Kearney CM. 2015a. Nectar protein content and attractiveness to Aedes aegypti and Culex pipiens in plants with nectar/insect associations. Acta Trop. 146:81–88.

Coutu C, Brandle J, Brown D, Brown K, Miki B, Simmonds J, Hegedus DD. 2007. pORE: a modular binary vector series suited for both monocot and dicot plant transformation. Transgenic Res. 16:771–781.

Dan Y, Baxter A, Zhang S, Pantazis CJ, Veilleux RE. 2010. Development of Efficient Plant Regeneration and Transformation System for Impatiens Using Agrobacterium tumefaciensand Multiple Bud Cultures as Explants. BMC Plant Biol. 10:165.

Deane LM. 1988. Malaria studies and control in Brazil. Am. J. Trop. Med. Hyg. 38:223–230.

Dellaporta SL, Wood J, Hicks JB. A plant DNA minipreparation: Version II. Plant Mol. Biol. Report. 1:19–21.

Dojcinovic D, Krosting J, Harris AJ, Wagner DJ, Rhoads DM. 2005. Identification of a region of the Arabidopsis AtAOX1a promoter necessary for mitochondrial retrograde regulation of expression. Plant Mol. Biol. 58:159–175.

Escalante-Pérez M, Jaborsky M, Lautner S, Fromm J, Müller T, Dittrich M, Kunert M, Boland W, Hedrich R, Ache P. 2012. Poplar Extrafloral Nectaries: Two Types, Two Strategies of Indirect Defenses against Herbivores. Plant Physiol. 159:1176–1191.

Foster WA. 1995. Mosquito Sugar Feeding and Reproductive Energetics. Annu. Rev. Entomol. 40:443–474.

Gallup JL, Sachs JD. 2001. The economic burden of malaria. Am. J. Trop. Med. Hyg. 64:85–96.

Gardener MC, Gillman MP. 2001. Analyzing variability in nectar amino acids: composition is less variable than concentration. J. Chem. Ecol. 27:2545–2558.

Gates B. The Deadliest Animal in the World. [accessed 2016 Oct 21]. https://www.gatesnotes.com/Health/Most-Lethal-Animal-Mosquito-Week

Gaucher D, Therrien R, Kettaf N, Angermann BR, Boucher G, Filali-Mouhim A, Moser JM, Mehta RS, Drake DR, Castro E, et al. 2008. Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. J. Exp. Med. 205:3119–3131.

González-Teuber M, Eilmus S, Muck A, Svatos A, Heil M. 2009a. Pathogenesis-related proteins protect extrafloral nectar from microbial infestation. Plant J. Cell Mol. Biol. 58:464–473.

González-Teuber M, Eilmus S, Muck A, Svatos A, Heil M. 2009b. Pathogenesis-related proteins protect extrafloral nectar from microbial infestation. Plant J. 58:464–473.

Goulson D. 1999. Foraging strategies of insects for gathering nectar and pollen, and implications for plant ecology and evolution. Perspect. Plant Ecol. Evol. Syst. 2:185–209.

Gu W, Müller G, Schlein Y, Novak RJ, Beier JC. 2011. Natural Plant Sugar Sources of Anopheles Mosquitoes Strongly Impact Malaria Transmission Potential. PLOS ONE 6:e15996.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8:1494–1512.

Hadlington JL, Denecke J. 2000. Sorting of soluble proteins in the secretory pathway of plants. Curr. Opin. Plant Biol. 3:461–468.

von Heijne G. 2001. Signal Peptides. In: eLS. John Wiley & Sons, Ltd.

Heil M, Rattke J, Boland W. 2005. Postsecretory Hydrolysis of Nectar Sucrose and Specialization in Ant/Plant Mutualism. Science 308:560–563. [accessed 2016 Sep 19]

Helsper JPFG, Ruyter-Spira CP, Kwakman PHS, Bleeker WK, Keizer LCP, Bade JB, Te Velde AA, Zaat SAJ, Verbeek M, Creemers-Molenaar J. 2011. Accumulation of human EGF in nectar of transformed plants of Nicotiana langsdorffii × N. sanderae and transfer to honey by bees: hEGF in nectar and honey. Plant Biol. 13:740–746.

Hennessy S, Strom BL, Bilker WB, Zhengle L, Chao-Min W, Hui-Lian L, Tai-Xiang W, Hong-Ji Y, Qi-Mau L, Tsai TF, et al. 1996. Effectiveness of live-attenuated Japanese encephalitis vaccine (SA14-14-2): a case-control study. The Lancet 347:1583–1586.

Hillwig MS, Kanobe C, Thornburg RW, MacIntosh GC. 2011. Identification of S-RNase and peroxidase in petunia nectar. J. Plant Physiol. 168:734–738.

Hillwig MS, Liu X, Liu G, Thornburg RW, MacIntosh GC. 2010. Petunia nectar proteins have ribonuclease activity. J. Exp. Bot. 61:2951–2965.

Hocking B. 1968. Insect-Flower Associations in the High Arctic with Special Reference to Nectar. Oikos 19:359.

Ito J, Ghosh A, Moreira LA, Wimmer EA, Jacobs-Lorena M. 2002. Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. Nature 417:452–455.

Jefferson RA, Kavanagh TA, Bevan MW. 1987. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6:3901–3907.

Kang DS, Cotten MA, Denlinger DL, Sim C. 2016. Comparative Transcriptomics Reveals Key Gene Expression Differences between Diapausing and Non-Diapausing Adults of Culex pipiens. PLOS ONE 11:e0154892.

Keller MD, Leahy DJ, Norton BJ, Johanson T, Mullen ER, Marvit M, Makagon A. 2016. Laser induced mortality of Anopheles stephensi mosquitoes. Sci. Rep. 6. [accessed 2016 Oct 21]. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4758184/

Kevan PG, Baker HG. 1983. Insects as flower visitors and pollinators. Annu. Rev. Entomol. 28:407–453.

King BC. 2011. T-phylloplanin and cis-abienol, two natural products from tobacco have broad spectrum, anti-fungal activities.

Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, Moore CG, Carvalho RG, Coelho GE, Bortel WV, et al. 2015. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. eLife 4:e08347.

Kram BW, Bainbridge EA, Perera MADN, Carter C. 2008. Identification, cloning and characterization of a GDSL lipase secreted into the nectar of Jacaranda mimosifolia. Plant Mol. Biol. 68:173.

Kram BW, Xu WW, Carter CJ. 2009. Uncovering the Arabidopsis thaliana nectary transcriptome: investigation of differential gene expression in floral nectariferous tissues. BMC Plant Biol. 9:92.

Kroumova ABM, Sahoo DK, Raha S, Goodin M, Maiti IB, Wagner GJ. 2013. Expression of an apoplast-directed, T-phylloplanin-GFP fusion gene confers resistance against Peronospora tabacina disease in a susceptible tobacco. Plant Cell Rep. 32:1771–1782.

Lee J-Y, Baum SF, Oh S-H, Jiang C-Z, Chen J-C, Bowman JL. 2005. Recruitment of CRABS CLAW to promote nectary development within the eudicot clade. Development 132:5021–5032.

Lee S-B, Li B, Jin S, Daniell H. 2011. Expression and characterization of antimicrobial peptides Retrocyclin-101 and Protegrin-1 in chloroplasts to control viral and bacterial infections. Plant Biotechnol. J. 9:100–115.

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Peer YV de, Rouzé P, Rombauts S. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 30:325–327.

Lin IW, Sosso D, Chen L-Q, Gase K, Kim S-G, Kessler D, Klinkenberg PM, Gorder MK, Hou B-H, Qu X-Q, et al. 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. Nature 508:546–549.

MacKenzie DJ, McLean MA, Mukerji S, Green M. 1997. Improved RNA Extraction from Woody Plants for the Detection of Viral Pathogens by Reverse Transcription-Polymerase Chain Reaction. Plant Dis. 81:222–226.

Magallon S, Crane PR, Herendeen PS. 1999. Phylogenetic Pattern, Diversity, and Diversification of Eudicots. Ann. Mo. Bot. Gard. 86:297.

Map of Impatiens walleriana -- Discover Life. [accessed 2016b Oct 22]. http://discoverlife.org/mp/20m?kind=Impatiens+walleriana

McGettigan PA. 2013. Transcriptomics in the RNA-seq era. Curr. Opin. Chem. Biol. 17:4–11.

Moehnke MH, Midoro-Horiuti T, Goldblum RM, Kearney CM. 2008. The expression of a mountain cedar allergen comparing plant-viral apoplastic and yeast expression systems. Biotechnol. Lett. 30:1259–1264.

Molina C, Grotewold E. 2005. Genome wide analysis of Arabidopsis core promoters. BMC Genomics 6:25.

Monath TP, Liu J, Kanesa-Thasan N, Myers GA, Nichols R, Deary A, McCarthy K, Johnson C, Ermak T, Shin S, et al. 2006. A Live, Attenuated Recombinant West Nile Virus Vaccine. Proc. Natl. Acad. Sci. U. S. A. 103:6694–6699.

Mosquito-Borne Diseases. [accessed 2016c Oct 21]. http://www.mosquito.org/mosquito-borne-diseases

Müller G, Schlein Y. 2005. Plant tissues: the frugal diet of mosquitoes in adverse conditions. Med. Vet. Entomol. 19:413–422.

Müller G, Schlein Y. 2006. Sugar questing mosquitoes in arid areas gather on scarce blossoms that can be used for control. Int. J. Parasitol. 36:1077–1080.

Müller GC, Beier JC, Traore SF, Toure MB, Traore MM, Bah S, Doumbia S, Schlein Y. 2010. Field experiments of Anopheles gambiae attraction to local fruits/seedpods and flowering plants in Mali to optimize strategies for malaria vector control in Africa using attractive toxic sugar bait methods. Malar. J. 9:262.

Müller GC, Kravchenko VD, Schlein Y. 2008. Decline of Anopheles sergentii and Aedes caspius populations following presentation of attractive toxic (spinosad) sugar bait stations in an oasis. J. Am. Mosq. Control Assoc. 24:147–149.

Müller GC, Schlein Y. 2008. Efficacy of toxic sugar baits against adult cistern-dwelling Anopheles claviger. Trans. R. Soc. Trop. Med. Hyg. 102:480–484.

Naqvi SMS. 2005. Nectarin IV, a Potent Endoglucanase Inhibitor Secreted into the Nectar of Ornamental Tobacco Plants. Isolation, Cloning, and Characterization. PLANT Physiol. 139:1389–1400.

Nepi M, Bini L, Bianchi L, Puglia M, Abate M, Cai G. 2011. Xylan-degrading enzymes in male and female flower nectar of Cucurbita pepo. Ann. Bot. 108:521–527.

Nepi M, Soligo C, Nocentini D, Abate M, Guarnieri M, Cai G, Bini L, Puglia M, Bianchi L, Pacini E. 2012. Amino acids and protein profile in floral nectar: Much more than a simple reward. Flora - Morphol. Distrib. Funct. Ecol. Plants 207:475–481.

Olivera BM, Baine P, Davidson N. 1964. Electrophoresis of the nucleic acids. Biopolymers 2:245–257.

Organization WH, others. 2016. World malaria report 2015: summary. [accessed 2016 Oct 21]. http://apps.who.int/iris/handle/10665/205224

Oya A. 1988. Japanese Encephalitis Vaccine. Pediatr. Int. 30:175–184.

Park S, Thornburg RW. 2009. Biochemistry of Nectar Proteins. J. Plant Biol. 52:27–34.

Perret M. 2001. Nectar Sugar Composition in Relation to Pollination Syndromes in Sinningieae (Gesneriaceae). Ann. Bot. 87:267–273. [accessed 2016 Sep 19]

Peterson GL. 1977. A simplification of the protein assay method of Lowry et al. which is more generally applicable. Anal. Biochem. 83:346–356.

Peumans WJ, Smeets K, Van Nerum K, Van Leuven F, Van Damme EJ. 1997a. Lectin and alliinase are the predominant proteins in nectar from leek (Allium porrum L.) flowers. Planta 201:298–302.

Peumans WJ, Smeets K, Van Nerum K, Van Leuven F, Van Damme EJ. 1997b. Lectin and alliinase are the predominant proteins in nectar from leek (Allium porrum L.) flowers. Planta 201:298–302. [accessed 2016 Sep 20]

Phillpotts RJ, Venugopal K, Brooks T. Immunisation with DNA polynucleotides protects mice against lethal challenge with St. Louis encephalitis virus. Arch. Virol. 141:743–749.

Picard-Nizou AL, Grison R, Olsen L, Pioche C, Arnold G, Pham-Delegue MH. 1997. Impact of proteins used in plant genetic engineering: Toxicity and behavioral study in the honeybee. J. Econ. Entomol. 90:1710–1716.

Pimentel D. 2005. "Environmental and Economic Costs of the Application of Pesticides Primarily in the United States." Environ. Dev. Sustain. 7:229–252.

Prevention C-C for DC and. CDC - Malaria - About Malaria - History - Elimination of Malaria in the United States (1947-1951). [accessed 2016 Oct 21]. https://www.cdc.gov/malaria/about/history/elimination_us.html

Richardson L. 2016. Mass spectrometry-based peptide sequencing and an evaluation of the enzymatic properties of the I. walleriana extrafloral nectar proteome [Thesis]. [accessed 2016 Oct 19]. https://baylor-ir.tdl.org/baylor-ir/handle/2104/9763

Rochelle PA, De Leon R, Stewart MH, Wolfe RL. 1997. Comparison of primers and optimization of PCR conditions for detection of Cryptosporidium parvum and Giardia lamblia in water. Appl. Environ. Microbiol. 63:106–114.

Roshchina VV, Roshchina VD. 2012. The Excretory Function of Higher Plants. Springer Science & Business Media.

Rudgers JA, Gardener MC. 2004. EXTRAFLORAL NECTAR AS A RESOURCE MEDIATING MULTISPECIES INTERACTIONS. Ecology 85:1495–1502. [accessed 2016 Sep 19]

Ruhlmann JM, Kram BW, Carter CJ. 2010. CELL WALL INVERTASE 4 is required for nectar production in Arabidopsis. J. Exp. Bot. 61:395–404.

Savolainen V, Chase MW, Hoot SB, Morton CM, Soltis DE, Bayer C, Fay MF, De Bruijn AY, Sullivan S, Qiu Y-L. 2000. Phylogenetics of flowering plants based on combined analysis of plastid atpB and rbcL gene sequences. Syst. Biol. 49:306–362.

Schlein Y, Müller GC. 2008. An Approach to Mosquito Control: Using the Dominant Attraction of Flowering Tamarix jordanis Trees Against Culex pipiens. J. Med. Entomol. 45:384–390.

Seo PJ, Wielsch N, Kessler D, Svatos A, Park C-M, Baldwin IT, Kim S-G. 2013. Natural variation in floral nectar proteins of two Nicotiana attenuata accessions. BMC Plant Biol. 13:101.

Shepherd RW. 2005. Phylloplanins of Tobacco Are Defensive Proteins Deployed on Aerial Surfaces by Short Glandular Trichomes. PLANT CELL ONLINE 17:1851–1861.

Shi S-G, Yang M, Zhang M, Wang P, Kang Y-X, Liu J-J. 2014. Genome-wide transcriptome analysis of genes involved in flavonoid biosynthesis between red and white strains of Magnolia sprengeri pamp. BMC Genomics 15:706.

Silberglied RE. 1979. Communication in the ultraviolet. Annu. Rev. Ecol. Syst. 10:373–398. [accessed 2016 Sep 20]

Stockinger. GUS: Histochemical Staining with X-Gluc - GUS-HistochemicalStaining.pdf. [accessed 2016 Oct 23]. http://www.oardc.ohio-state.edu/stockingerlab/Protocols/GUS-HistochemicalStaining.pdf

Triglia T. 2000. Inverse PCR (IPCR) for Obtaining Promoter Sequence. In: Tymms M, editor. Transcription Factor Protocols. Humana Press. (Methods in Molecular Biology™). p. 79–84. [accessed 2016 Oct 22]. http://dx.doi.org/10.1385/1-59259-686-X%3A79

Wa F, Rg H. 1994. Nectar-related olfactory and visual attractants for mosquitoes. J. Am. Mosq. Control Assoc. 10:288–296.

WHO | Executive summary. WHO. [accessed 2016d Oct 21]. http://www.who.int/whr/1996/media_centre/executive_summary1/en/index9.html

Wiwanitkit V. 2007. Vaccination against mosquito borne viral infections: current status. Iran J Immunol 4:186–196.

Wolff D. 2006. Nectar Sugar Composition and Volumes of 47 Species of Gentianales from a Southern Ecuadorian Montane Forest. Ann. Bot. 97:767–777. [accessed 2015 Nov 10]

Wunnachit W, Jenner CF, Sedgley M. 1992. Floral and extrafloral nectar production in Anacardium occidentale L.(Anacardiaceae): an andromonoecious species. Int. J. Plant Sci.:413–420.

X-GLUC Information ★ Buy XGlucuronide from X-Gluc DIRECT. [accessed 2016e Oct 20]. http://www.x-gluc.com/xgluc.htm

Zha H-G, Flowers VL, Yang M, Chen L-Y, Sun H. 2012. Acidic α-galactosidase is the most abundant nectarin in floral nectar of common tobacco (Nicotiana tabacum). Ann. Bot. 109:735–745.

Zha H-G, Liu T, Zhou J-J, Sun H. 2013. MS-desi, a desiccation-related protein in the floral nectar of the evergreen velvet bean (Mucuna sempervirens Hemsl): molecular identification and characterization. Planta 238:77–89.