ABSTRACT

Mass spectrometry-based peptide sequencing and an evaluation of the enzymatic properties of the *I. walleriana* extrafloral nectar proteome

Luke Richardson

Director: Dr. Chris Kearney, Ph.D.

For many years, ecologists have observed how plants mediate mutualist interactions with aggressive defenders against herbivores, such as ants, through production of metabolite-rich extrafloral nectar (EFN) as a benefit to the ant defender. However, this nectar can also support unwanted microbial contamination detrimental to the plant. Recent studies have elucidated enzymatic mechanisms in nectar by which plants protect vulnerable tissues from microbial infestation. As a novel subject of interest, relatively few nectar proteomes have been fully evaluated and characterized. The EFN proteome of *Impatiens walleriana*, a common flowering plant native to East Africa, falls into this group. Given its abundant secretion of EFN, attractiveness to insects, including mosquitoes, and robust expression of nectar proteins, *I. walleriana* was determined by the Kearney lab to be an excellent candidate to serve as a transgenic model system for study of nectar-mediated delivery of insecticidal proteins. In pursuance of this objective, the purposes of this study are twofold: firstly, to generate and identify nectar protein, and then genomic, sequence data, and, secondly, evaluate the enzymatic properties of the *I. walleriana* EFN proteome in its defense of nectary tissues against microbial infestation. We collected extensive peptide sequence data via bottom-up proteomic analysis of *I. walleriana* EFN proteins with tandem mass spectrometry in expectation of identifying peptides with a transcriptome-derived protein sequence database. Additionally, I found that the EFN had mild antimicrobial activity on *E. coli* but a much more pronounced effect on yeast growth. Furthermore, activity indicating the presence of peroxidase class enzyme was detected and quantified in the *I. walleriana* EFN and implicated in microbial defense.

APPROVED BY DIRECTOR OF HONORS THESIS:

_____

Dr. Chris Kearney, Department of Biology

APPROVED BY THE HONORS PROGRAM:

_____

Dr. Elizabeth Corey, Director

DATE:_____

MASS SPECTROMETRY-BASED PEPTIDE SEQUENCING AND AN EVALUATION OF THE ENZYMATIC PROPERTIES OF THE *I. WALLERIANA* EXTRAFLORAL NECTAR PROTEOME

A Thesis Submitted to the Faculty of

Baylor University

In Partial Fulfillment of the Requirements of the

Honors Program

By

Luke Richardson

Waco, Texas

May 2016

TABLE OF CONTENTS

LIST OF FIGURES

Figure 1. Impatiens walleriana produces an abundance of nectar from extrafloral nectary tissues located on the petioles and leaf edges.

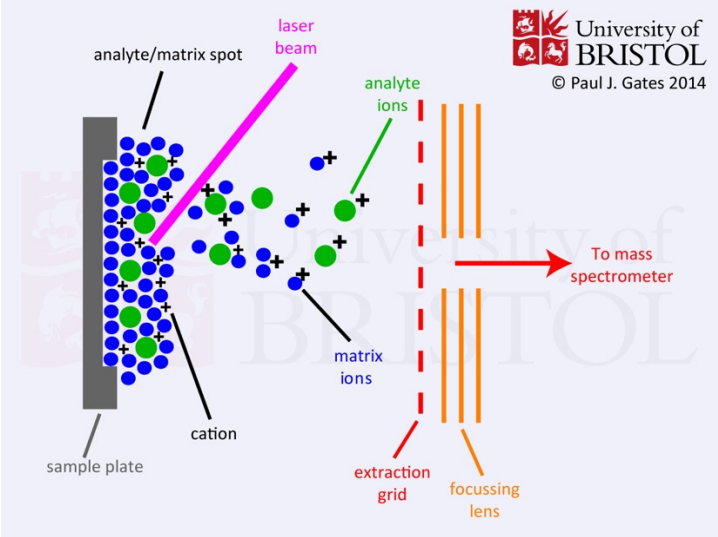

Figure 2. Schematic showing the process of analyte ionization by the laser desorbed MALDI matrix. Charge is transferred from the ionized matrix molecules to the analyte prior to entry to the mass analyzer (From University of Bristol School of Chemistry).

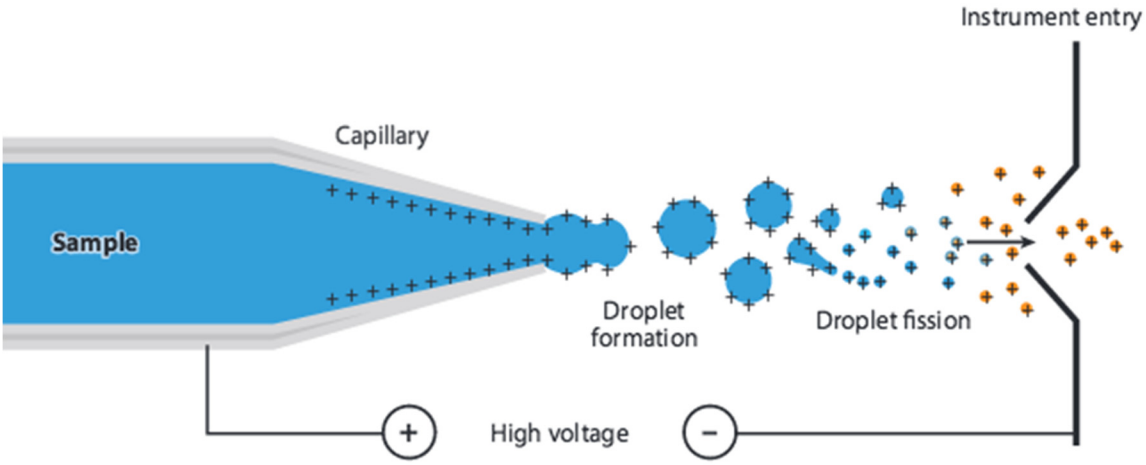

Figure 3. Charged droplets containing gas-phase ion precursors evolve from the capillary and undergo fission prior to instrument entry due to electrostatic repulsions prior to instrument entry (From Mehmood et al., 2015).

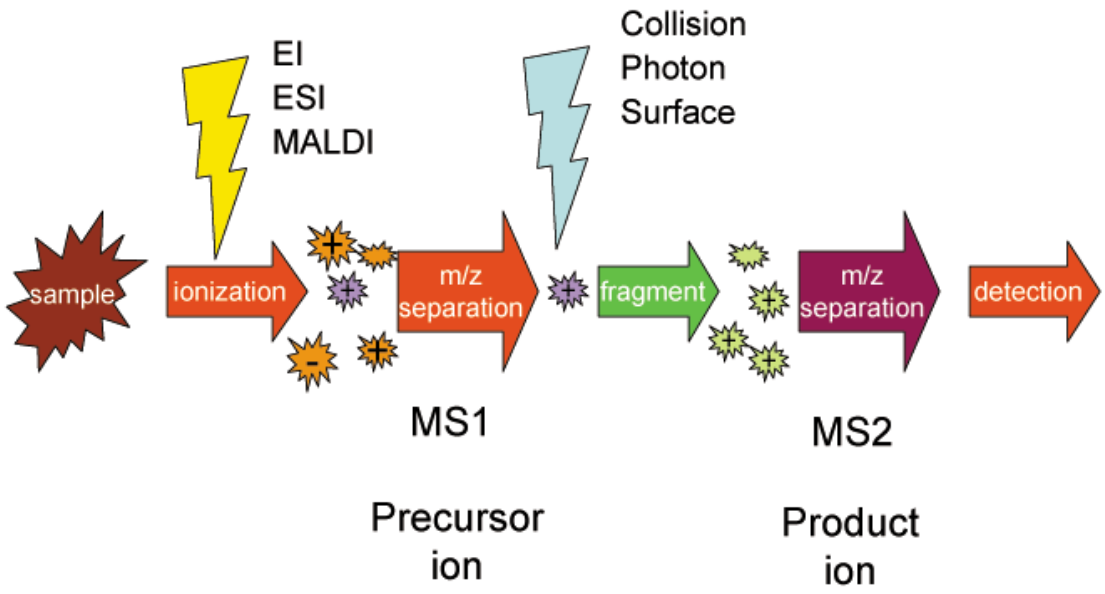

Figure 4. Simple schematic of MS/MS process showing the selection and successive fragmentation of a certain precursor peptide ion for further analysis and sequencing. By K. Murray (Kkmurray) - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1943319.

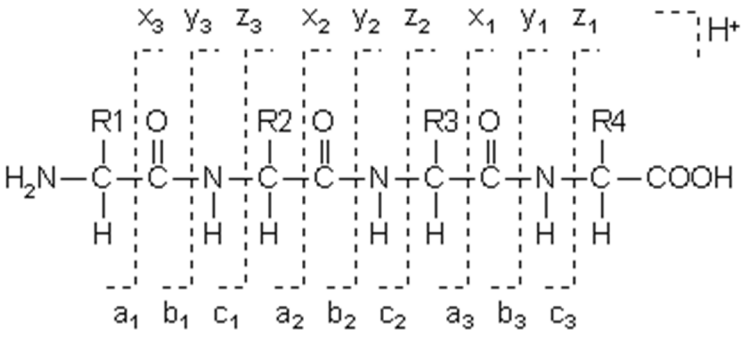

Figure 5. Observed peptide cleave sites for peptide fragmentation methods in mass spectrometry. CID generally produces b and y fragment ions and ETD selectively produces c and z fragment ions (From Johnson et al., 1987).

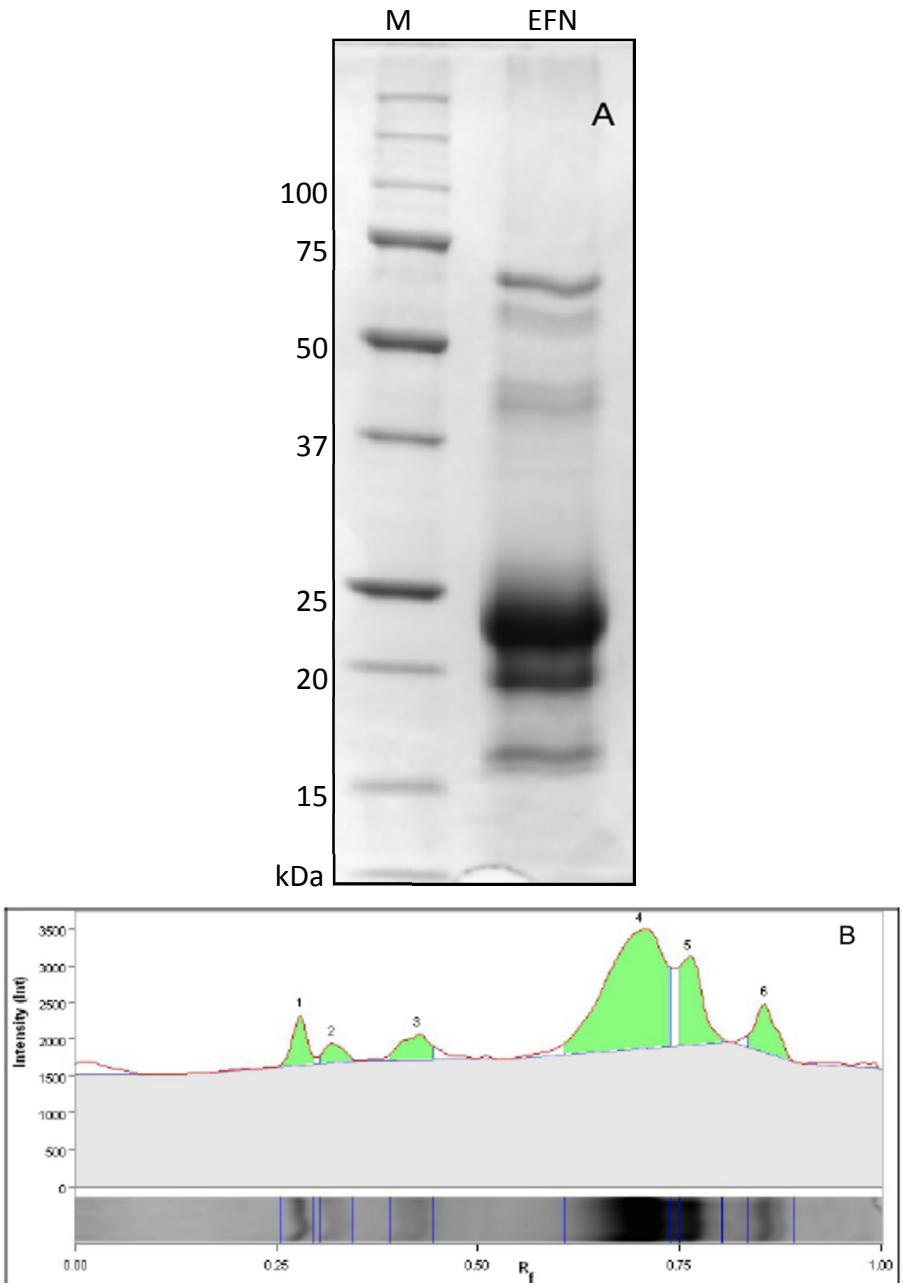

Figure 6. SDS-PAGE of *I. walleriana* extrafloral nectar proteins. Nectar was chloroform/methanol precipitated in order to concentrate proteins. The EFN lane (panel A) represents the equivalent of 20µl pure nectar. Gel Doc™ EZ (Bio-Rad) lane analysis (panel B) shows the relative migration of nectars (data in **Table 1**) and a visual representation of the integration used to determine quantity.

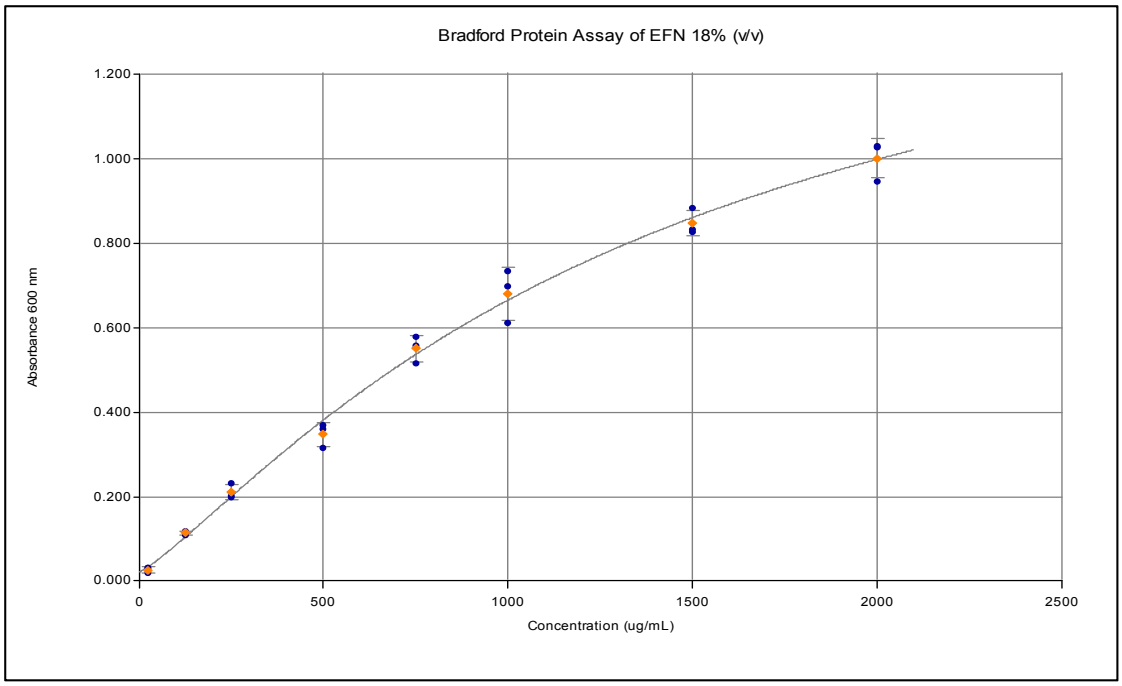

Figure 7. BSA standard absorbance curve used in Bradford assay to determine total protein content of an 18% (v/v) aq. nectar solution. The polynomial regression is described as follows: Y = (0.0215-1.64)/(1+(X/1.41E+03)^1.21) + 1.64; R^2 = 0.998.

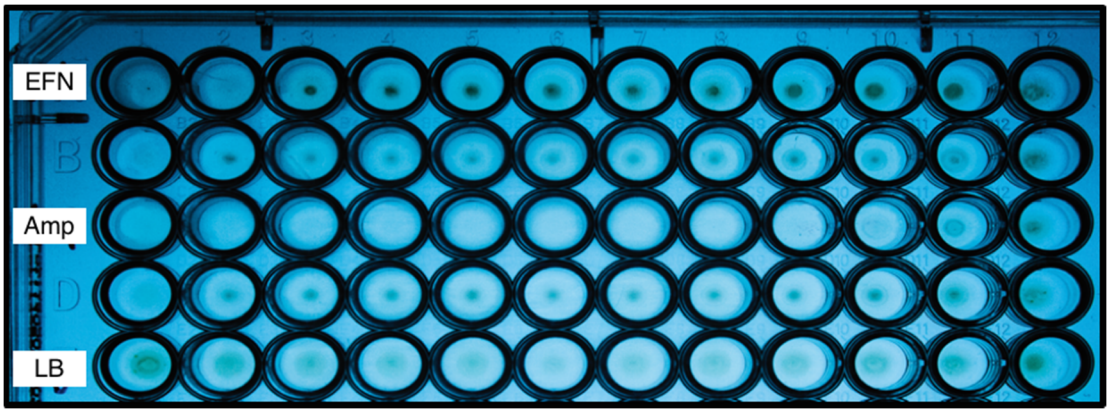

Figure 8. Antimicrobial activity assay conducted with *E. coli*. *Nec* lane contains serial dilutions of 15% EFN aq. solution. Positive control: *Amp* lane contains serial dilutions of ampicillin solution. Negative control: *LB* lane contains lysogeny broth. Unmarked lanes were not used.

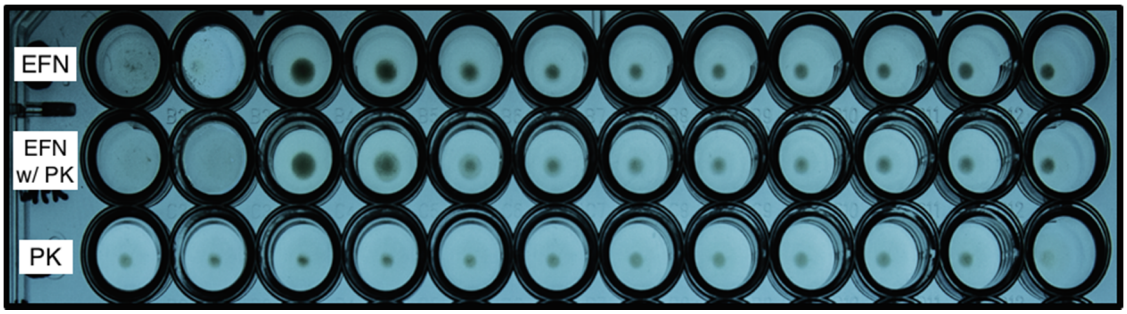

Figure 9. Antimicrobial assay conducted with *E. coli*. *EFN w/ PK* lane contains 15% EFN aqueous solution treated with proteinase K. Positive control: proteinase K-untreated 15% EFN aq. solution. Negative control: lysogeny broth with proteinase K.

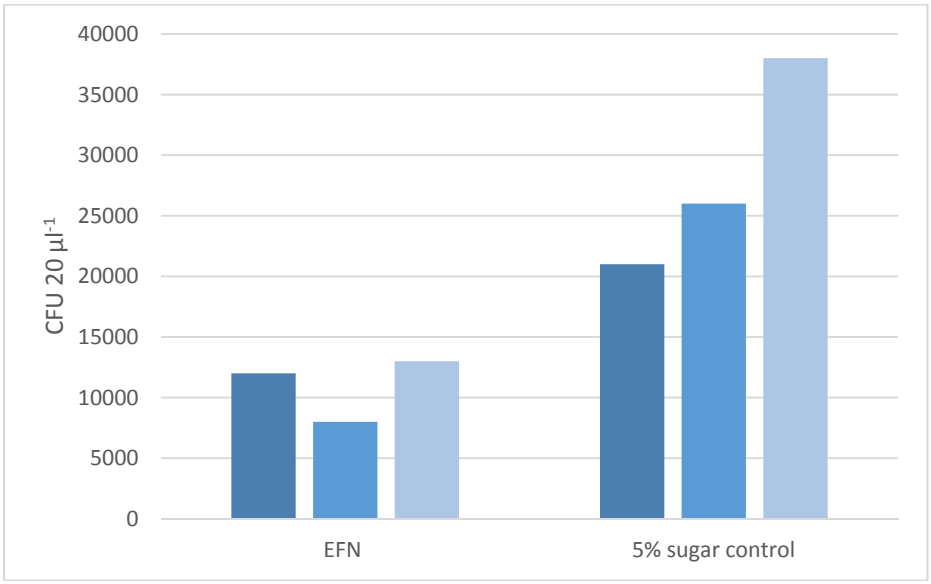

Figure 10. Effect of EFN activities on yeast growth. Yeast growth (number of colony-forming units per 20 µl) in EFN of *I. walleriana* compared to yeast growth in 5% sugar solution (1:1 glucose/fructose). Statistical analysis by univariate ANOVA confirmed the difference between the two groups to be statistically significant ($p = 0.03 < 0.05$).

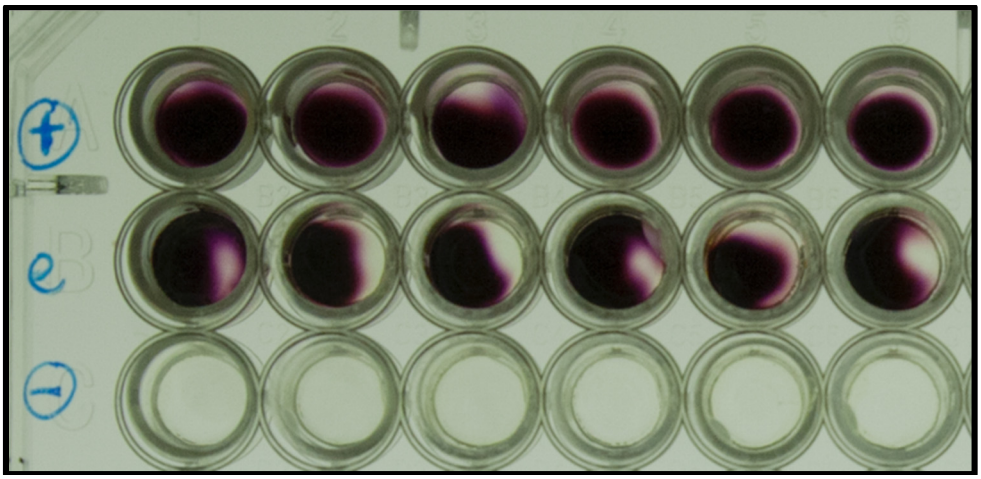

Figure 11. Assay for the colorimetric detection of peroxidase activity. All wells contain a working solution composed of 3% H_2O_2 solution and a developing solution for the colorimetric assay. The solution in *E* lane was introduced to a small volume of EFN protein solution and produced the colorimetric change seen above. Positive control: dilute horseradish peroxidase solution (Amresco) was introduced to the working solution. Negative control: working solution with an additional volume of DI water.

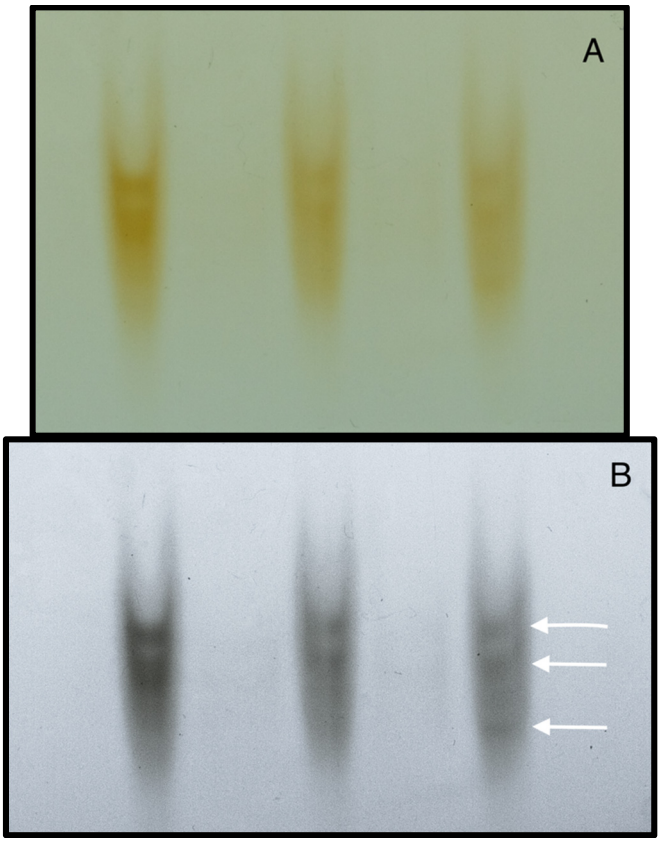

Figure 12. In-gel activity stain (panel A) and subsequent Coomassie Brilliant Blue R-250 stain of native PAGE of *I. walleriana* nectarins. In-gel activity stain using guaiacol (A) as an oxidative indicator revealed several bands of elevated peroxidase activity which were preserved in the Coomassie stain (B).

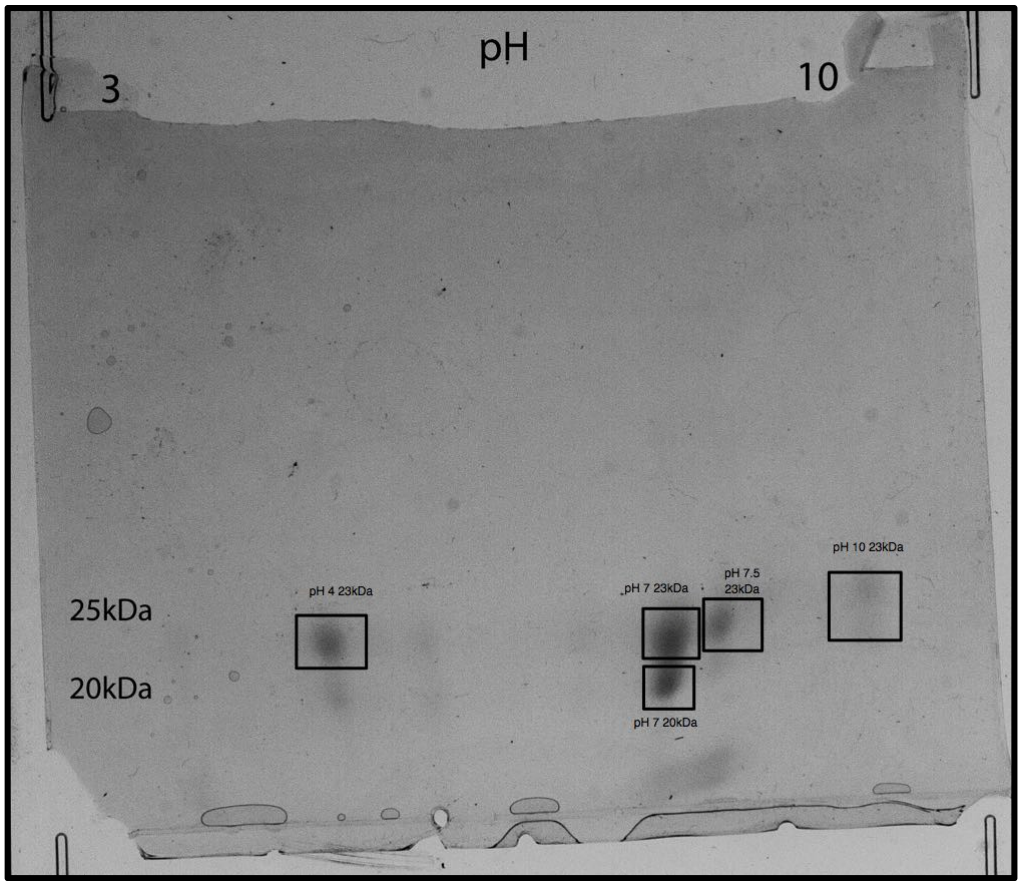

Figure 13. 2D gel electrophoresis separation of *I. walleriana* nectarins. Four bands of 23 kDa in size were isolated with pI's of 4.0, 7.0, 7.5 and 10.0. One band of 20 kDa in size was isolated with a pI of 7.0.

Observed MW: 1392.7811 Precursor ion charge state: 1
M/z tolerance: 0.30 Intensity threshold: 32 (0.750%)
Modifications: Carboxyamidomethylcysteine (+/-)

a	129.11	292.18	406.22	519.30	647.36	760.45	873.53	974.58	1061.61	1162.66	1219.68	1347.77
	0.01	0.01	-0.00	-0.00	0.00	-0.00	0.02	---	0.07	---	---	---
b	157.11	320.17	434.22	547.30	675.36	788.44	901.53	1002.57	1089.61	1190.65	1247.67	1375.77
	0.02	-0.01	-0.00	-0.00	0.00	-0.01	0.03	---	---	---	---	---
	Arg	Tyr	Asn	Leu	Gln	Leu	Leu	Thr	Ser	Thr	Gly	Lys
	66	66	100	100	100	100	100	100	99	99	99	100
y	1393.78	1237.68	1074.62	960.57	847.49	719.43	606.35	493.26	392.21	305.18	204.13	147.11
	---	---	---	---	-0.01	-0.01	-0.00	-0.00	-0.00	-0.01	-0.00	-0.00
z	1376.75	1220.65	1057.59	943.54	830.46	702.40	589.32	476.23	375.18	288.15	187.10	130.08
	---	---	---	-0.01	-0.01	---	---	---	---	---	---	-0.00

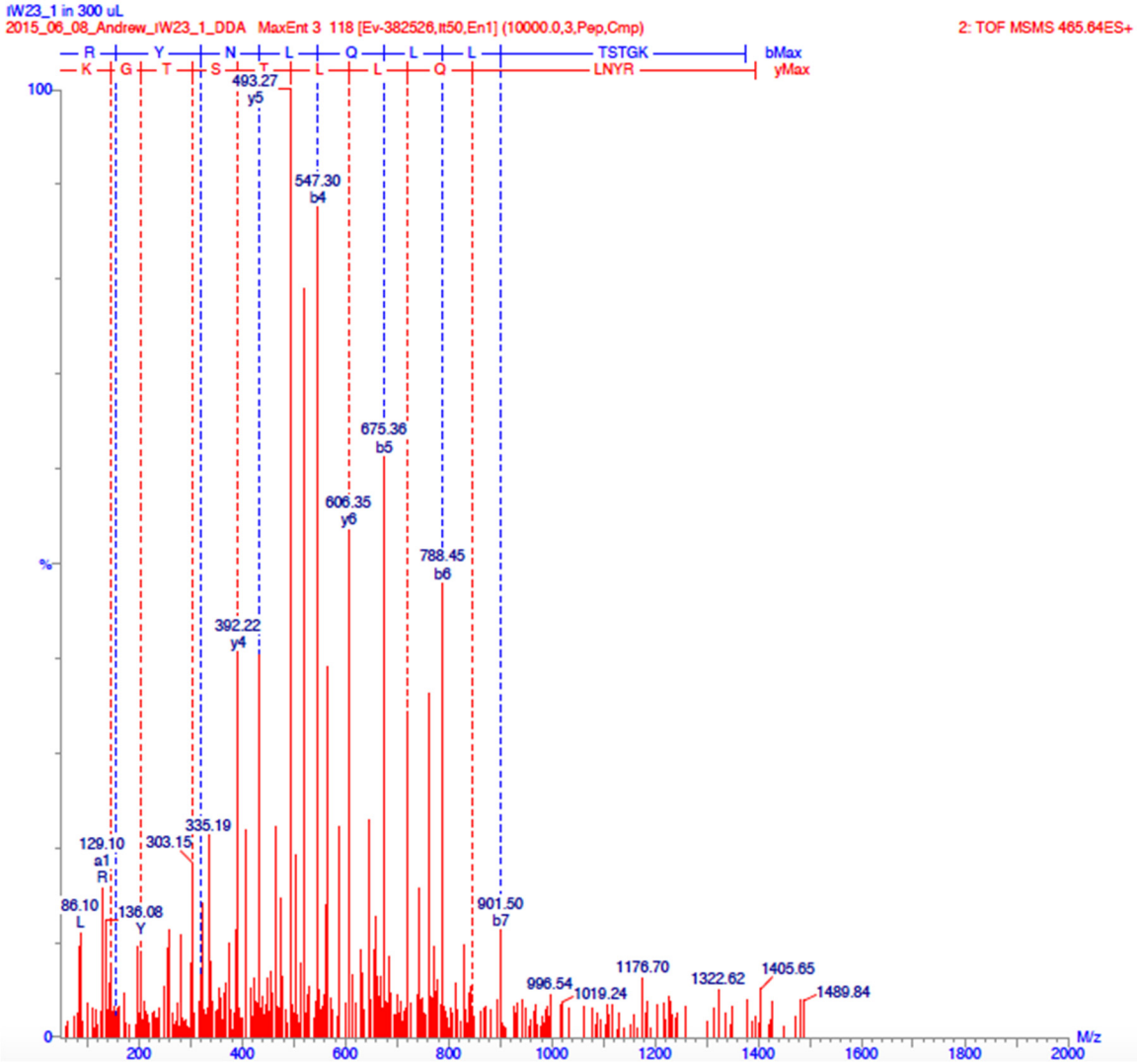

Figure 14. An example of the MS/MS spectra we collected using Waters BioLynx-PepSeq software. The software detected specific mass differences between b and y fragment ions, respectively, and assembles the peptide *de novo* using the information. Noted in the text at the top of the figure, BioLynx takes the precursor ion molecular weight and charge state into account. The red text on the right notes the precursor ion *m/z* peak from which the fragment spectrum was generated.

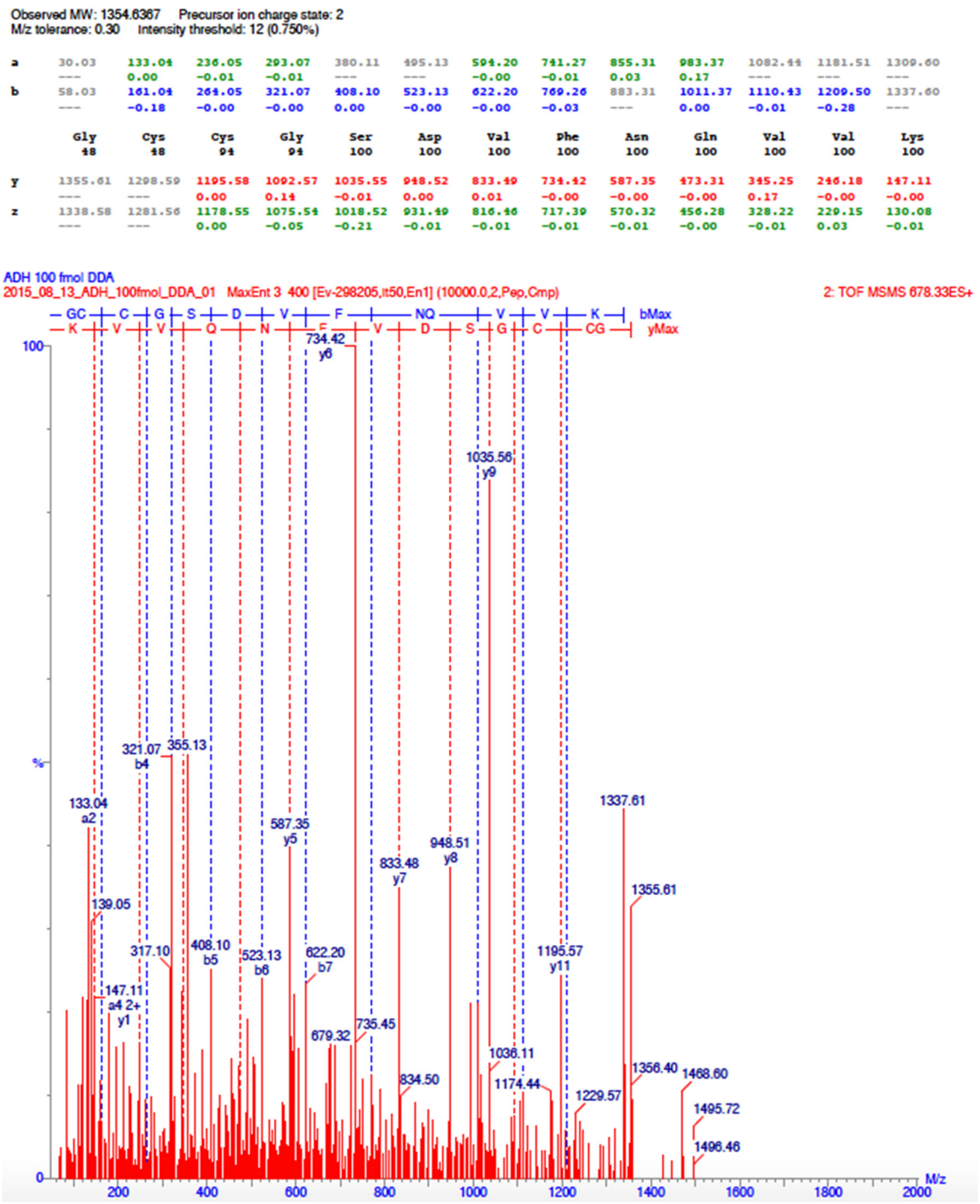

Figure 15. An MS/MS spectrum of a peptide from ADH. Its purposes here are to show a sequenced peptide with full fragment coverage and strong scores.

ACKNOWLEDGEMENTS

I thank Dr. Chris Kearney for providing the opportunity for me to conduct this research and for his continuous support through out the process. I am also grateful to Grace Pruett, Mishu Islam, Ankan Choudhury, Tommy Zhou, Meron Ghidey and Andrew Cox for their help and guidance during my time in the Kearney group. Lastly, special thanks to Tom Caracciolo for his assistance in assessing the current body of literature in the field of mosquito ecology.

CHAPTER ONE

Introduction

Nectar is the primary means of mediation between plant pollinators and their defenders whose primary purpose is to attract mutualists that properly take part in the plant-pollinator or plant-defender relationship. Plant nectar is an aqueous solution that contains high concentrations of simple sugars ranging from 8% to 80%, (w/w) (Baker and Baker, 1976) as well as other metabolites such as: amino acids (Baker and Baker, 1973), organic acids (Baker and Baker, 1975), terpenes (Ecroyd et al., 1995), alkaloids (Deinzer et al., 1977), flavonoids (Ferreres et al., 1996), glycosides (Roshchina and Roshchina, 1993), vitamins (Griebel and Heß, 1940), phenolics (Ferreres et al., 1996), metal ions (Heinrich, 1989), oils (Vogel) and proteins (Carter and Thornburg, 2004b). Once considered to be relatively simple solution with dissolved sugars and amino acids, studies in the last several decades have revealed complex biochemistry responsible for maintaining the nectar environment and encouraging favorable interactions. For example, production of phenolics (Ferreres et al., 1996) and alkaloids (Adler et al., 2006) in floral nectar seem to repel would be nectar thieves while volatile compounds, such as squalene in *Dactylanthus taylorii* (Ecroyd et al., 1995), play an important role in attracting some pollinators. Without a doubt, the roles of these compounds extend far beyond gratuitously providing pollinators with nutrition.

Plant nectar produced by both floral and extrafloral nectaries hosts a metabolite-rich environment ripe for infestation by respiring microbes. As such, bacteria and fungi

pose a serious threat to the nectar's chemical balance and thus its proper function. Floral nectar (FN) is secreted at the base of the flower surrounding the ovaries and offers a site for microbial infection which could drastically effect fertility (Park and Thornburg, 2009). In one instance, Peter Kevan demonstrated negative effects on insect pollinators due to yeast-infected nectar (Peter Kevan, 1988). For the sake of maintaining the plant-pollinator relationship, it stands to reason that plants would evolve antimicrobial mechanisms in order to maintain the sterility of the nectar. Previous studies by Carter et al. confirm this assumption in demonstrating the antimicrobial activity in ornamental tobacco nectar (Carter et al., 2007). Terpenoids produced in the nectaries of *Arabidopsis thaliana* exhibit significant antimicrobial activity as well (Deans and Waterman, 1993). Thus, it seems that particular compounds present in the nectar play a significant role in defense and actively maintaining a medium unsuitable for microbial growth.

However, it has not been long since researchers have begun to uncover the role of soluble nectar proteins in this defense. To date, only a relative few have been fully characterized through experimental demonstration of biological activity. In one of the few cases, Carter and Thornburg demonstrated the activity of the soluble proteins expressed in the FN of ornamental tobacco (*Nicotiana langsdorffi x Nicotiana sanderae*). The proteins, termed nectarins, have been shown to protect the FN from microbial infection via the Nectar Redox Cycle, which maintains high levels of hydrogen peroxide enough to resist infection and maintain a sterile nectar environment (Carter and Thornburg, 2004a, 2000, 2004b). Additionally, Kram et al. classified *Jacaranda* nectar protein, JNP1, a secreted nectar protein found in *Jacaranda mimosifolia*, as a GDSL

lipase with antimicrobial activity due to its ability to compromise microbial lipid membranes (Kram et al., 2008).

In the interest of retaining fecundity, plants have evolved strong antimicrobial defenses to maintain sterility surrounding otherwise vulnerable reproductive tissues. However, extrafloral nectaries, which are located on the petioles and leafs edges of plants (Chen and Kearney, 2015), have also been shown to serve as openings for phytopathogens (Ivanoff and Keitt, 1941; Keitt and Ivanoff, 1941) and have been observed as such in the case of the fire blight causing bacterium *Erwinia amylovora* (Farkas et al., 2011; T. Buban and Zs. Orosz-Kovacs, 2003). Therefore, from an evolutionary standpoint, plants stand to gain from antimicrobial defensive factors inherent to the nectar produced from these tissues. Prior to investigations concerning nectar proteomes, extrafloral nectar (EFN) was understood to serve the indirect defense of the plant via tritrophic interactions with the attraction of ants and other aggressive defenders (Escalante-Pérez and Heil, 2013). These interactions are facilitated by the presence of sugars and other metabolites in the nectar, the likes of which discussed above. However, in line with our assumption, the EFN of some species possess more complex biochemistry contributing to the defense of nectary tissues. The EFN nectaries of *Acacia* myrmecophytes, which participate in strong mutualist partnerships with ants, express more than fifty proteins in the nectar (González-Teuber et al., 2009); the majority of which are pathogenesis-related proteins such as chitinases, glucanases and thaumatin-like proteins that standby in defense against phytopathogens (González-Teuber et al., 2010). Chitinases were also observed in the FN of *F. irroratum* suggesting that they might constitute a common class of nectarin (Heil, 2011). Whether expressed in the FN

or EFN, nectarins serve the defense of vulnerable tissues beyond that which may be achieved by secondary metabolites as few benefits of a toxic nectar environment have been observed to date (Adler and Irwin, 2005). But, most importantly for my research group, nectar has been clearly demonstrated as a protein-rich and enzymatically diverse environment and, theoretically, an effective medium for nectar-mediated protein delivery to an imbibing consumer.

In line with this reasoning, the Kearney lab group is attempting to develop *Impatiens walleriana* as a model plant for nectar delivery of transgenic mosquitocidal proteins (Chen and Kearney, 2015). However, the use of orally delivered pesticides is rather novel; most vector control tools currently in use to combat malaria and other such mosquito-borne diseases involve the use of long-lasting insecticide-treated nets (LLINs) and indoor residual spraying (IRS) that kill by external contact (2006, 2008a). LLNs and IRS constitute the bulk of the World Health Organization's (WHO) integrated vector management (IVM) approach (2004, 2008b) to combatting malaria in affected countries, but the program is far from its goals of complete control and then elimination of the disease (Beier et al., 2008). These methods are effective in the short term by reducing malaria parasite transmission by over 90% in some cases and thereby reducing the frequency of new infection and malaria-related deaths. However, in the long term, these methods allow for low levels of transmission, barely detectable amounts of infective bites per person per year, that can correlate with prevalence rates of over 20% (Beier et al., 1999). Clearly, the incorporation of new vector management tools into the IVM framework and coordination with existing methods will be necessary in order to achieve the lofty goal of eliminating malaria and other vector-borne diseases.

Orally ingested pesticides such as attractive toxic sugar baits (ATSBs) offer an exciting new avenue in vector control in pursuance of this objective. The ATSB method, stemming from the "attract and kill" principle, uses a flower or fruit scent as an attractant, sugar solution as a feeding stimulant and an oral toxin to kill (Müller et al., 2010). As such, it exclusively affects outdoor adult mosquito populations, which rely heavily upon plant sugars as an energy source; in fact, male mosquitoes, which do not drink blood, rely exclusively on them (Foster, 1995). Even females of highly blood-dependent species such as *Anopheles sergentii (Gu et al., 2011), Anopheles gambiae (Fernandes and Briegel, 2005),* and *Aedes Aegypti* (Mostowy and Foster, 2004) reap great benefits from sugar sources which manifest in increased egg laying and longevity. ATSB field trials in the arid climate of Israel proved the method extremely effective in decimating various local mosquito populations and demonstrated its effectiveness over a variety of species (Müller and Schlein, 2006, 2008a, 2008b; Müller et al., 2008). Likewise, trials conducted in the semi-arid climate of Mali were equally as effective especially in reducing the relative abundance of *An. gambiae*. Within a week of spraying, ATSBs reduced adult female and male population densities up to 90% and kept levels low for the remainder of the observation period (Müller et al., 2010). The ATSB method's capacity to control local populations and reduce the numbers of recently matured female anophelines to enter houses and feed on humans thoroughly qualifies ATSBs as a viable vector control tool in the semi-arid regions of Africa. Surely, the successes with this method provide sufficient rationale to justify investigation of nectar-mediated delivery of mosquitocidal proteins in efforts to expand the repertoire of effective vector control tools.

Impatiens walleriana (**Figure 1**), a flowering plant native to Eastern Africa, is extraordinarily well-suited to serve as the basis for a model system for the nectar delivery of transgenic proteins. Chen and Kearney arrived at this conclusion as the result of a study designed to evaluate the attractiveness of and the survivability conferred by five easily grown garden plants (*I. walleriana* included) to *Culex pipiens and Aedes aegypti* mosquitoes. Chen and Kearney initially selected *I. walleriana* for their evaluation by virtue of its known association with ants due to abundant secretion of sugar- and amino acid-rich EFN (Lanza et al., 1993). The other four candidate plants involved in the study (*Ricinus communis, Campsis radicans, Asclepias syriaca and Passiflora edulis*) were chosen on a similar basis whether for reported associations with ants or mosquitoes, respectively (Elias and Gelband, 1975; Foster, 2008; Gary and Foster, 2004; Xu and Chen, 2010). Through both solo and competitive attraction assays, impatiens demonstrated superior attractiveness to both *C. pipiens* and *A. aegypti*. In the solo attraction assays measuring immediate readiness to imbibe, 80—95% of the mosquitoes imbibed nectar from a single *I. walleriana* plant by the end of day 1 of enclosure surpassing the other nectar-secreting plants by 40—90% (Chen and Kearney, 2015). The competitive attraction assays yielded similar results: in the presence of five other plants and a tube of sucrose solution, the mosquitoes showed a strong preference for *I. walleriana*, which attracted 70—80% of *A. aegypti* and 80—95% of *C. pipiens* (Chen and Kearney, 2015). *C. radicans* and sucrose solution followed next trailing by about 30% in both cases (Chen and Kearney, 2015). In survival assays, which assessed the longevity conferred to the mosquitoes via consistent nectar imbibition, *I. walleriana* nectar kept nearly 80% of the initial mosquito population alive after 20 days surpassing the 10%

(w/v) sucrose positive control and other nectar-secreting plants by a significant margin (Chen and Kearney, 2015). In addition to attractiveness and conferred survivability, Chen and Kearney also evaluated levels of EFN protein expression in the candidates looking forward to the possibility of utilizing a native nectarin promoter for transgenic protein expression. SDS PAGE results of precipitated nectar proteins revealed vigorous protein expression in *I. walleriana* EFN several orders beyond the other candidates (Chen and Kearney, 2015). So, in every respect, this study identified *I. walleriana* as an exceptional candidate to serve as the basis of a model plant system for the nectar delivery of transgenic proteins (Chen and Kearney, 2015).

Figure 1. Impatiens walleriana produces an abundance of nectar from extrafloral nectary tissues located on the petioles and leaf edges.

As protein sequence databases are continually filled and improved upon, most serious proteomics studies are now accomplished in part with mass spectrometry coupled

with various analysis methods. In fact, González-Teuber et al. identified or annotated 52 different pathogenesis-related proteins in *Acacia comigera* with the help of mass spectrometry techniques (González-Teuber et al., 2009). In an amazing feat, Shah et al. performed a full time-course proteomic analysis of developing extrafloral nectaries of *Ricin communis* where they identified 882 unique proteins throughout four stages of development (Shah et al., 2016). Additionally, mass spectrometry methods were instrumental in the identification of nectarin enzymes in the FN of ornamental tobacco (Carter and Thornburg, 2004a, 2000). And while these examples are confined to the realm of plant proteomics, mass spectrometry has found applications in every field associated with the identification and analysis of molecules. This is primarily due to the diverse array of available technologies that suit the analysis of many different types of molecules for a host of varying purposes. However, even as technologies and techniques diversify, the fundamental principles of mass spectrometry remain the same. Each and every mass spectrometer measures the mass-to-charge (m/z) ratios and the abundance of gas-phase ions of an ionized analyte. This being the case, there are four main components of a mass spectrometer: a sample introduction mechanism, an ion source, a mass analyzer and a detector (Harris, 2010). The characteristics of each component greatly affect how analyte ions travel and the manner in which *m/z* data is collected from their response.

Mass spectrometry is an analytical technique used for studying the masses of atoms, molecules or fragments of molecules (Watson and Sparkman, 2007). In order to acquire mass spectral data, condensed analytes are ionized and desorbed into the gaseous phase and accelerated along a path by an electric field where they are separated according

to their mass-to-charge ratio (*m/z*) (Harris, 2010). However, if molecules or fragments are not successfully ionized, they will not be affected by the electric and/or magnetic field applied by the mass spectrometer resulting in loss of signal.

Ionization methods are characterized in one of two ways: hard ionization and soft ionization. The most common hard ionization technique is electron ionization, or electron impact (EI), where the analyte is bombarded with accelerated electrons with kinetic energy of 70 eV (Harris, 2010). 70 eV is much more energy than necessary to ionize molecules; however, it is considered standard practice in EI because its gives highly reproducible fragmentation spectra that can be compared with library spectra (Harris, 2010). Usually, nonbonding or pi electrons are first to dissociate leaving a radical cation, denoted the molecular ion ($M^{+\bullet}$), which may then fragment further; however, EI is suitable primarily for volatile compounds <1 kDa in size and may completely eliminate $M^{+\bullet}$ or produce unusable spectra due to excessive fragmentation (Harris, 2010). Hard ionization techniques such as EI are classified by their ability to produce such fragment peaks. Due to these limitations, scientists use so-called soft ionization techniques in order to generate gas-phase ions directly from solution-phase molecules when working with larger biomacromolecules such as proteins, most notably of which are matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). These soft ionization techniques are able to produce gas-phase molecular ions from a few daltons to megadaltons while preventing excessive fragmentation and preserving structure (Mehmood et al., 2015).

MALDI involves a rapid photovolatilization of biomolecules embedded in an UV-absorbing matrix (Mehmood et al., 2015). The mechanism of ionization has not been

definitively determined but is generally understood to occur in two phases (**Figure 2**): laser activation of the matrix causes a plume, or desorption, of highly ionized matrix ions which then "envelops" and ionizes the analyte (Zenobi and Knochenmuss, 1998). MALDI allows for identification of peptides in complex mixtures given that the ions produced in the gas-phase are predominantly singly charged $[M+H]^+$ ions (Yates et al., 2009). But while a few examples may be cited where MALDI has been used in tandem with liquid chromatography (Stevenson et al., 1998), most choose ESI for liquid chromatography and tandem mass spectrometry, which will be discussed at length later.

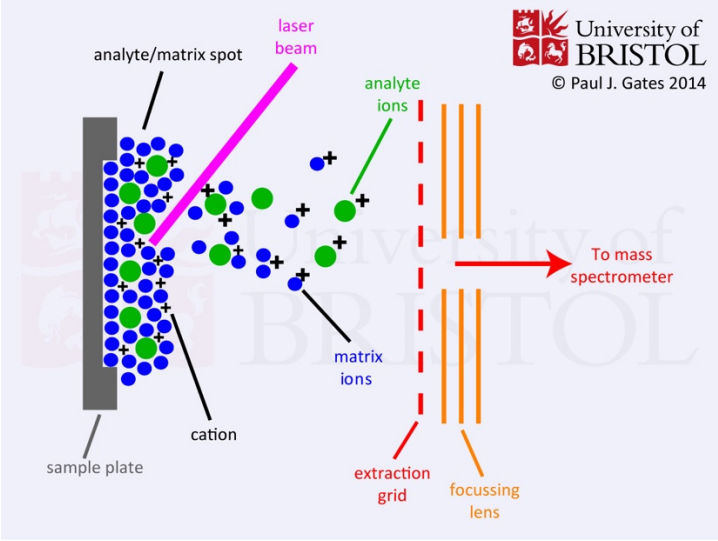

Figure 2. Schematic showing the process of analyte ionization by the laser desorbed MALDI matrix. Charge is transferred from the ionized matrix molecules to the analyte prior to entry to the mass analyzer (From University of Bristol School of Chemistry).

As opposed to the ionizing energy transfer involved in MALDI, ESI uses a high electric field to produce ions from solution (**Figure 3**). Voltages of 2—6 kV are applied to the capillary containing the analyte solution driving the accumulation of positive ions at the tip (Mehmood et al., 2015). This process yields an electrically charged spray, the Taylor Cone (Taylor, 1964), consisting of charged droplets containing the analyte;

however, strong electrostatic repulsions cause droplet fission and yield gas-phase ions upon entry into the mass analyzer (Mehmood et al., 2015). Peptide ions produced by ESI are most commonly found as $[M+nH]^{n+}$ but can also be seen paired with sodium, $[M+Na]^{+}$, and potassium, $[M+K]^{+}$ (Harris, 2010). A notable aspect of ESI that differentiates its mass spectra from other soft ionization sources such as MALDI is the generation of multiply charged protein/peptide ions, as implied by the ion denotation $[M+nH]^{n+}$ where n equals the number of charges. The additional *m/z* introduced by a multiply charged protein may convolute the spectra somewhat, especially if other unique proteins are present, but researchers have developed tools which can readily identify and group multiply charged species (Ferguson and Smith, 2003). ESI is preferred in most protein studies because it allows for coupling with a constant ion source such as liquid chromatography where as MALDI performs pulsed analysis (Benesch et al., 2007; Konermann et al., 2013). This distinction generally informs how these ionization technologies are paired with various mass analyzers.

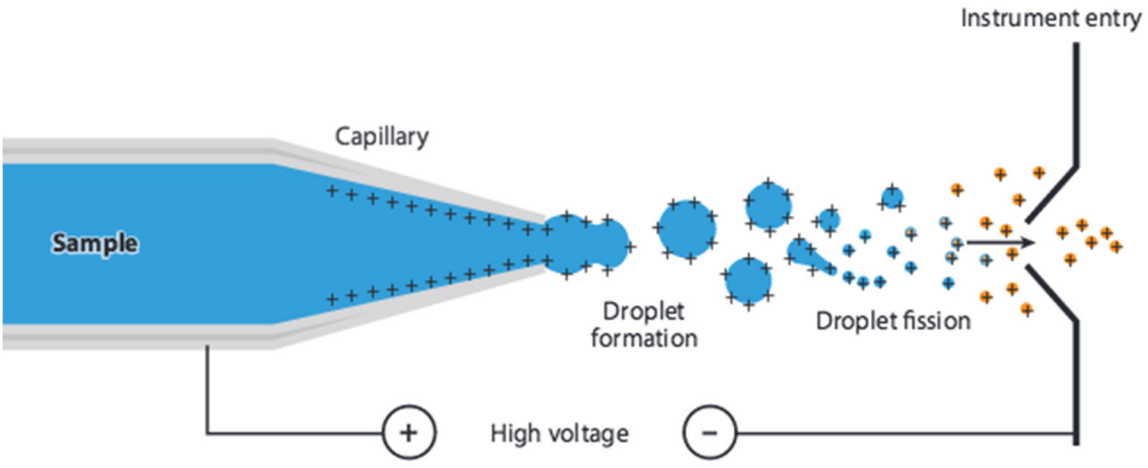

Figure 3. Charged droplets containing gas-phase ion precursors evolve from the capillary and undergo fission prior to instrument entry due to electrostatic repulsions prior to instrument entry (From Mehmood et al., 2015).

Once generated by MALDI, ESI or other ions sources, gas-phase ions are separated according to their *m/z* by the mass analyzer before they hit the detector. In the field of proteomics, time-of-flight (TOF) mass analyzers are commonly used due to their ability to separate ions of virtually any size (Konijnenberg et al., 2013; Lössl et al., 2014). TOF analyzers function on the principal that the "time of flight" of an ion over a fixed distance under high vacuum is directly proportional to its *m/z* (Chernushevich et al., 2001). Often times, TOF mass analyzers are hybridized with quadrapole mass analyzers, which act as mass filters or ion guides. In a basic quadrapole, both a constant voltage and radio-frequency oscillating voltage are applied to four parallel metal rods (Harris, 2010). The electric field created causes ions to fly between the conducting rods with complex trajectories; ultimately, only ions with a specific *m/z* can pass from the mass separator (Harris, 2010). In a standard transmission quadrapole (TQ) mass spectrometry, these ions would then reach the detector; however, in tandem mass spectrometry (MS/MS) using a hybridized system (**Figure 4**) such as Q-TOF (Quadrapole-TOF) MS, *m/z* selected ions encounter a collision cell prior to TOF analysis.

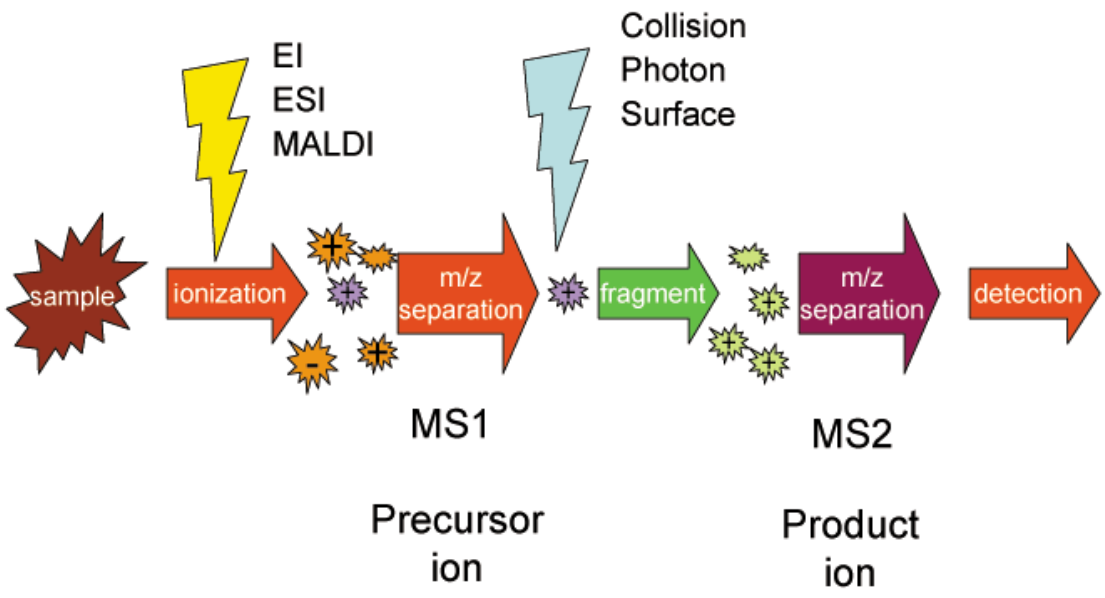

Figure 4. Simple schematic of MS/MS process showing the selection and successive fragmentation of a certain precursor peptide ion for further analysis and sequencing. By K. Murray (Kkmurray) - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1943319.

In a bottom-up proteomics approach, the likes of which used in this study, in which the protein(s) of interest has not been previously sequenced, proteins are identified via a peptide-to-protein inference logic (Gillet et al., 2015). In most cases, protein MS samples are enzymatically digested into peptides with the help of a specific protease. In this study, samples were digested with trypsin for its robust activity and general specificity. Additionally, trypsin digestion tends to yield fragments from 500 to 3,000 Da in size, which is an optimal mass range for chromatographic separation techniques, and yields fragments that are easily ionized due to the presence of C-terminal arginine and lysine residues, which readily protonate at lower pH's (Vandermarliere et al., 2013). Necessarily for bottom-up analyses, additional fragmentation steps are introduced to a MS/MS hybridized system in order to enable amino acid sequencing. As seen in Figure 4, fragmentation steps are added between mass analyzer components. For example, in

the Waters Synapt G2 MS used in this study, which may be generally described as a Q-TOF system, peptide ions filtered through the first quadrapole mass analyzer are subject to additional fragmentation in a collision cell, re-ionized and accelerated into the TOF MS component that collects *m/z* data.

Peptide fragmentation may be intentionally increased by methods such as collision-induced dissociation (CID) and electron-transfer dissociation (ETD) (Harris, 2010). In CID, which was utilized in this study, molecular ions that have passed through the quadrapole mass filter collide with residual N_2, or another inert gas, under pressure with enough energy to cleave amide bonds – the lowest energy bond in the peptide backbone. The produced fragments, denoted b and y fragments, are then analyzed via TOF (Harris, 2010). On the other hand, ETD provides a selective mechanism with which to cleave peptide chains in a MS/MS system. Upon ionization and *m/z* selection for the peptide fragment ions, a gas-phase anion, such as 9-anthracenecarboxylate, introduced via a separate ion spray transfers an electron to a gas-phase cationic polypeptide causing peptide cleavage between the α-carbon and amine group (McAlister et al., 2007). The fragments produced by ETD are denoted c and z fragments. These methods individually produce 2*n* possible ionized fragments where *n* equals the total number of amino acid residues. In the b and c fragments, the charge resides on the N-terminal side of the cleave and, in the y and z fragments, the charge resides on the C-terminal side (**Figure 5**). If greater fragmentation coverage is necessary, CID and ETD may be used in the same collision cell. While the fragment ion spectra may be more visually convoluted, the additional fragmentation gives the software more information with which to assemble the precursor peptide ion. The high resolution inherent to TOF analysis can detect these

minute differences in *m/z* and computer software can use this information to determine which amino acids are in each fragment and thereby identify their sequence (Coon, 2009).

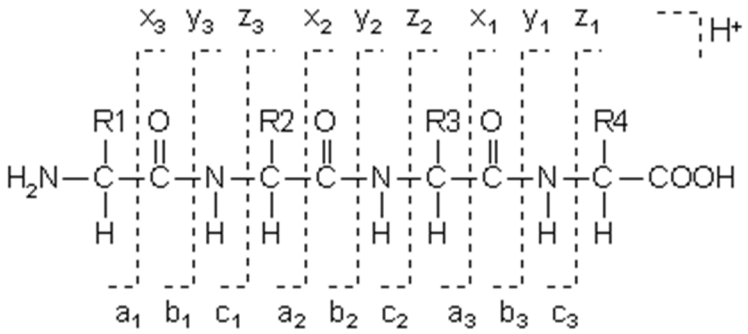

Figure 5. Observed peptide cleave sites for peptide fragmentation methods in mass spectrometry. CID generally produces b and y fragment ions and ETD selectively produces c and z fragment ions (From Johnson et al., 1987).

At this point, it is evident that many different criteria must be considered in making decisions regarding proper MS instrumentation. Certain analyzers or combinations of analyzers can yield *m/z* information of varying amounts of usefulness for any particular sample. In modern proteomics, we must also consider the mode of data acquisition and adjust the instrument configuration accordingly. Three primary operating modes are in use today: targeted data acquisition via selected reaction monitoring (SRM), data-independent acquisition (DIA), and data dependent acquisition (DDA) (Gillet et al., 2015).

SRM, as a targeted proteomic acquisition method, is unique in that no complete mass spectra are acquired (Picotti and Aebersold, 2012). Instead, select precursor ions are targeted and their respective fragment ion spectra are repeatedly collected; therefore, prior knowledge of the precursor ion's identity and fragmentation characteristics is

necessary (Picotti and Aebersold, 2012). SRM acquisition is thus best suited to confirming and/or quantifying the presence of the targeted peptides rather than discovery. This workflow is best suited for and most commonly carried out with a triple quadrapole system, which can be fine-tuned for very specific mass-to-charge ratios (Picotti and Aebersold, 2012).

DIA, in contrast with SRM and DDA as we will see, is aimed towards systematically acquiring MS/MS spectra for any possible precursor ion (Chapman et al., 2014). This workflow is not altered by actual detection (DDA) or the supposed presence (SRM) of fragment ions. However, instrumentation currently in use cannot sample fast enough to cover the entire peptide mass range produced by trypsin digestion with the small precursor isolation ranges used in DDA or SRM, about 1-3 Da (Chapman et al., 2014). Therefore, when using DIA, the precursor mass isolation range must be much larger, thus isolating a mixture of precursor peptide ions and yielding a multiplexed MS/MS spectra containing fragment signal from several different peptides (Gillet et al., 2015). While this data may be more difficult to process, DIA workflows allow for high throughput and high coverage analysis of complex protein mixtures.

DDA, also referred to as shotgun, is defined by iterative acquisition of intact precursor spectra (MS, or MS^1) and fragmented ion spectra (MS/MS, or MS^2) (Zhang et al., 2013). The MS software makes decisions in real time based on a predetermined set of criteria on which precursor ions to select for fragmentation (Gillet et al., 2015). As a general rule, only those precursors with the strongest signals are selected as they yield the highest chance of successful identification via subsequent analysis of their respective MS/MS spectra (Zhang et al., 2013). The identity of the precursor peptide can be derived

from the information of the precursor mass and the related fragment ion masses in the MS/MS spectra. DDA MS/MS workflows were used exclusively in the identification of peptide sequences in this study.

The peptide sequence data generated from the MS/MS workflow is then instrumental in the identification and characterization the parent protein(s) through the use of protein sequence databases. However, some strategies for proteome profiling can be more effective than others on a case-by-case basis. The Kearney group's decisions concerning this matter will be discussed at length in the Discussion and Conclusions section of this thesis.

In lieu of the results found by Christopher Kearney and Zhongyuan Chen, the purpose of this research is to further characterize the *I. walleriana* extrafloral nectar proteome in order to develop *I. walleriana* as a model system for nectar delivery of transgenic proteins. As nectar protein activity and expression are often linked to microbial defense and pathogenesis (González-Teuber et al., 2010), this study also seeks to evaluate the *in vitro* antimicrobial capabilities of *I. walleriana* EFN against gram-negative *Escherichia coli*, gram-positive *Staphylococcus aureus* and *Saccharomyces cerevisiae* in order to reveal protein function in this area. Indeed, statistically significant antimicrobial activity was observed against *S. cerevisiae*, commonly known as baker's yeast, which has been demonstrated to be the product of pathogenesis-related proteins in the EFN of *Acacia* myrmecophytes. Following these results, enzymatic assays for the presence of peroxidase family enzymes were positive affirming the expression of soluble PR proteins in the nectar, the activity of which was subsequently quantified and then directly visualized in an in-gel activity stain relative to the total protein content of the

EFN. Additionally, MS/MS data of trypsin digested EFN nectarins was analyzed with Waters MassLynx software to generate peptide sequence data in preparation for analysis with an *I. walleriana* tissue-specific transcriptome-derived protein sequence database.

CHAPTER TWO

Materials and Methods

Nectar Collection

Nectar was collected from impatiens plants by adhering the viscous nectar to autoclaved pipette tips or filter paper. The nectar was then dissolved in an aqueous solution of a predetermined volume in capped microcentrifuge tubes to prevent excessive evaporation. After collection, the nectar solution volume was measured to determine nectar concentration (v/v%). For antimicrobial assays, nectar was collected to concentrations above 15% v/v and then diluted to a uniform concentration of 15% v/v for assays. Nectar samples were kept frozen at -20°C prior to use. For enzymatic assays, nectar was used in the same day as collection.

SDS-PAGE Gel Electrophoresis

SDS-PAGE was conducted with 15% bis-acrylamide SDS-PAGE gels according to the methods of Laemmli (1970). Lane and band analysis was performed with a Gel Doc EZ Imager (BioRad).

Total EFN Protein Quantification

Prior to quantification, nectar concentration (v/v%) was calculated post-collection. Using a BSA standard, a Bradford Assay (Thermo Scientific) using a Biotek EL-800 Plate Reader at 600 nm was used in order to determine total protein content. Three replicates of two dilutions were analyzed. Dilution calculations were used work

from the nectar concentration noted at collection to determine the initial nectar protein concentration as produced by the nectaries using the. The initial nectar protein concentrations as produced by the nectaries were determined by dilution calculations using the collected nectar concentration noted at collection and the protein concentration determined in the Bradford Assay.

Anti-Bacterial Activity Assay

Anti-bacterial activity was assayed according to the antimicrobial screen protocol developed by the Kearney lab. Individual *E. coli* and *S. aureus* strains were cultured in Luria broth overnight at 37°C and then diluted. Turbidity was visually checked against a #0.5 McFarland suspended latex bead standard. EFN sample and control solutions were pipetted into the first column of wells of a 96-well mictrotiter plate. Bacterial solution was added in equal volumes to the rows in use. Serial dilutions of sample and control solutions were then made down each row. Samples treated with proteinase K were incubated at 37°C for 30 minutes prior to plating. The plate was incubated at 37°C for *E. coli* and room temperature for *S. aureus*. The presence or absence of microbial growth was then visually assayed for each well and the lowest dilution associated with antimicrobial activity was recorded.

Yeast Assay of Anti-Fungal Activity

An assay with *Saccharomyces cerevisiae* was performed to evaluate the potential effects of EFN enzymes on fungal growth. A single colony of *S. cerevisiae* was cultivated on potato dextrose agar plates. This single clone was cultured in YPD liquid growth media at room temperature for 30 hours, which was then centrifuged (250 g for

10 min at room temperature), resuspended in 1X phosphate-buffered saline (PBS) solution and stored at 4°C. *I. walleriana* EFN 12% (v/v) and 5% sugar (1:1 fructose:glucose) control solution were used in this assay. 20 µl aliquots of EFN or sugar solution were mixed with 20 µl of yeast suspension and incubated for 1 hour at 30°C. 20 µl of a dilution 1:1000 in 1X PBS was plated on potato dextrose agar plates. Colony-forming units were determined after 48 hours. Statistical difference between the two were determined by one-way ANOVA.

Detection of Peroxidase Activity in Nectar

Peroxidase activity in nectar was detected as follows. The protein content of 18% aqueous nectar solution was concentrated about 4.5X using the Vivaspin 4 centrifugal concentrator (Vivaproducts). 80 µl of concentrated nectar solution was added to a 96-well microtiter plate well followed by 40 µl of developing solution. The developing solution contained 80 mg/ml of 4-aminoantipyrine, 20% (v/v) N,N-dimethylanaline in 0.1 M sodium phosphate buffer, pH 5.5. Peroxidase presence was visually detected with the development of purple color in solution.

Peroxidase Activity Assay

Peroxidase activity was evaluated as follows. The protein content of EFN 10% (v/v) solution was concentrated 4.5X with the Vivaspin 4 concentrator. The reaction solution (total volume was 1 mL) contained 24 µL nectar solution, 5 µL guaiacol (99%), 42 µL H_2O_2 (3% v/v) and 930 µL 50 mM $NaPO_4$ buffer, pH 6.0. The H_2O_2 dependent oxidation of guaiacol was monitored at 470 nm, using an extinction coefficient of 26.6 x

10^3 μmol^{-1} cm^{-1}, with a Thermo Scientific BioMate 3S Spectrophotometer according to the method of Hammerschidt (Hammerschmidt et al., 1982).

Native PAGE and In-Gel Peroxidase Activity Stain

Native PAGE gels were prepared to 12% bis-acrylamide concentration with a Tris-HCl, pH 8.8 buffer system, and proteins were let to migrate at low voltage (~100V) for two hours. The peroxidase activity and Coomassie staining were done according to the method of Fieldes (Fieldes, 1992). Gels were submerged in peroxidase staining reagent (consisting of 50 mL of 10 mM H_2O_2, 75 mL of 20 mM guaiacol (Thermo Scientific) and 75 mL of 0.1 M tris-HCl buffer, pH 4.0) for 10-40 minutes before taking pictures. Gels were then left in Coomassie-fixing solution (with a final concentration of 0.018 mg/mL Coomassie Brilliant Blue R-250) for at least 16 hours.

2D Gel Electrophoresis

2D gel electrophoresis was performed with the Ettan IPGphor 3 IEF system (GE) according to the GE Healthcare operating instructions. Prior to sample loading, Immobiline DryStrips for IEF were rehydrated in prepared DeStreak rehydration solution face-down in the IEF strip holder for 10 hours to overnight. Mineral oil was also applied on top of the strip to prevent the strip from drying out. Nectar protein samples were prepared in sample buffer containing DTT reducing agent and then pipetted into the lateral wells at either end of the strip holder below the mineral oil. The strip holders were then properly positioned on the Ettan IPGphor 3 platform. IEF and the second dimension (SDS-PAGE) was run according to the method of Giri (Giri et al., 2006).

In-Gel Tryptic Digestion and MS Sample Preparation

In-gel tryptic digestion was performed according to the protocol provided with the Thermo Scientific in-gel tryptic digestion kit. Gel bands were excised and subjected to two 30 minute rounds of destaining at 37°C. Gel slices were then subject to reduction by TCEP for 10 minutes for 60°C and then akylation by iodoacetamide for 1 hour in the dark. Gel slices were then subject to two wash rounds in destain solution for 15 minutes at 37°C. Gel slices were then shrunk and dried in 100% acetonitrile and then digested in an activated trypsin solution. Digested samples were dried down with an Eppendorf Vacufuge plus and resuspended in 1% formic acid.

Tandem Mass Spectrometry and Nectarin Peptide Sequencing

MS/MS peptide sequence data was acquired using a Waters Synapt G-2 HDMS, utilizing a UPLC sample inlet system. Peptide sequence data was acquired with Waters MassLynx MS Software using a DDA workflow. The DDA workflow selects high intensity LC peaks for iterative cycles of MS^1 and MS^2, in which precursor ion and fragment ion spectra are acquired in the same time frame. MassLynx and BioLynx PepSeq utilize the MaxEnt3 charge deconvolution function to determine the $[M+H]^+$precursor peptide mass and mass differences between fragment ion peaks in order to assemble the peptide sequence from the terminal lysine/arginine residue. Peptide sequences were scored based on how closely each fragment mass difference matched with the theoretical mass of the assigned amino acid.

CHAPTER THREE

Results

SDS-PAGE and Quantification of I. walleriana Nectarins

SDS-PAGE analysis of *I. walleriana* EFN proteins was conducted in order to affirm the banding pattern observed by Chen and Kearney as well as to isolate peptides for MS/MS analysis. The banding pattern observed (**Figure 6a**) shows six major bands. In particular, the protein in the band at 22.4 kDa is expressed at a relative abundance several times greater than those of other observed protein bands as shown by the integration in the band intensity plot (**Figure 6b**). Table 1 provides a comprehensive data set describing the SDS—PAGE banding pattern of the *I. walleriana* EFN proteome.

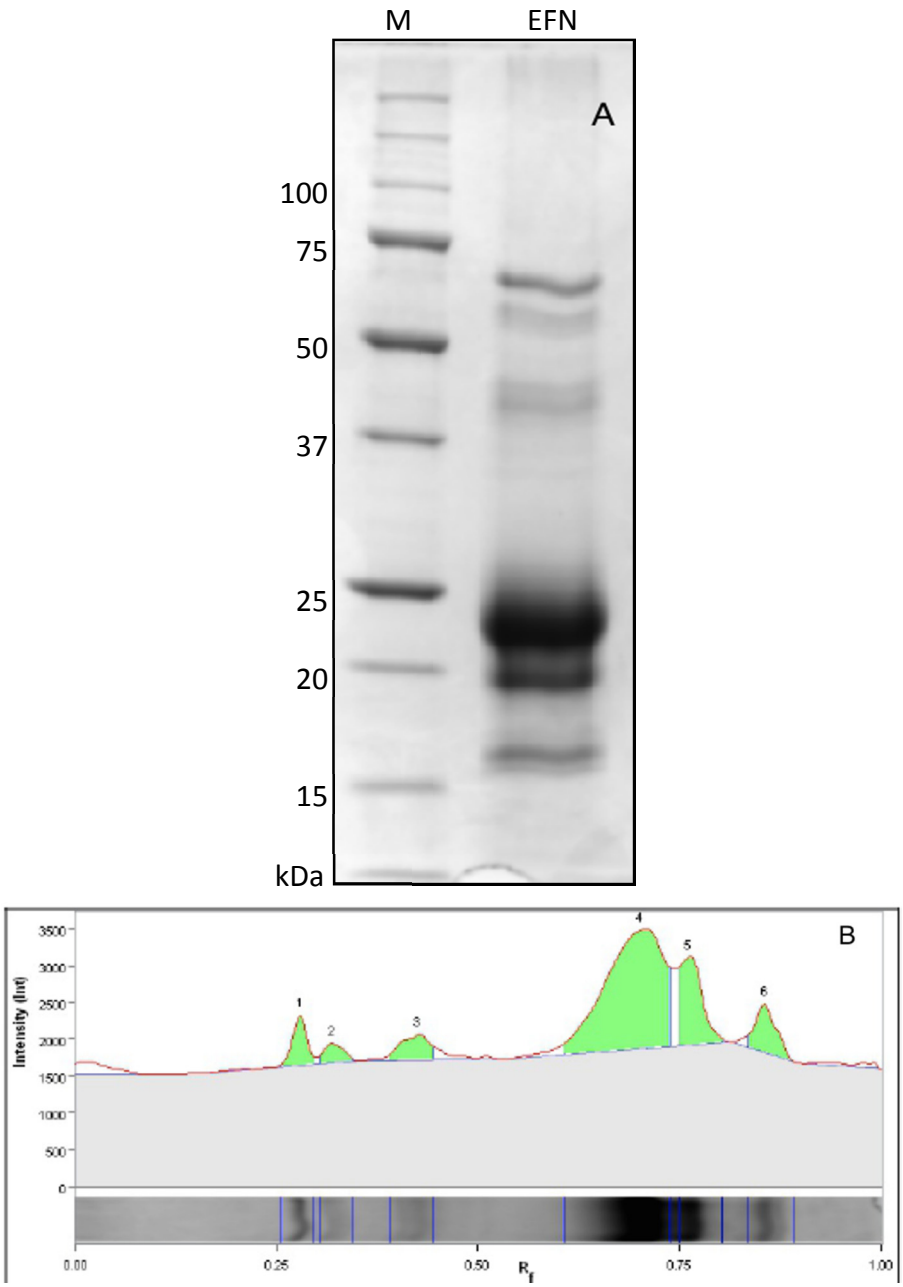

Figure 6. SDS-PAGE of *I. walleriana* extrafloral nectar proteins. Nectar was chloroform/methanol precipitated in order to concentrate proteins. The EFN lane (panel A) represents the equivalent of 20µl pure nectar. Gel Doc™ EZ (Bio-Rad) lane analysis (panel B) shows the relative migration of nectars (data in **Table 1**) and a visual representation of the integration used to determine quantity.

Table 1. Gel Doc™ EZ Lane and Band Analysis

Band No.	Mol. Wt. (KDa)	Rel. Quant.
1	62.9	0.635
2	55.4	0.326
3	41.1	0.614
4	22.4	6.00
5	19.6	1.71
6	16.2	0.872

Total EFN nectar protein was quantified using a Bradford assay (Thermo Scientific) of a nectar dilution series. The standard BSA curve used in this determination is shown below (**Figure 7**). In an 18% aq. nectar solution, the protein concentration was determined to be 84.275 ± 8.069 μg/mL. Using dilution calculations, the total protein concentration in the nectar was determined to be 468 ± 45 μg/mL (see Table 2). This result is consistent with that produced by Chen and Kearney of 430 μg/mL (Chen and Kearney, 2015).

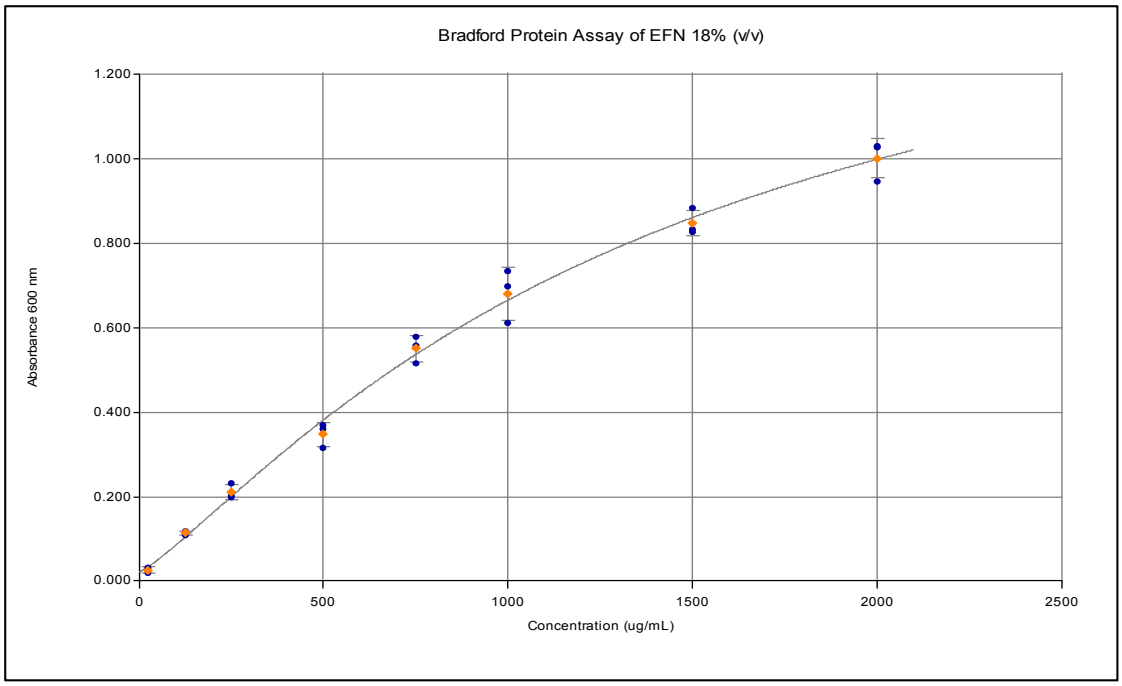

Figure 7. BSA standard absorbance curve used in Bradford assay to determine total protein content of an 18% (v/v) aq. nectar solution. The polynomial regression is described as follows: Y = (0.0215-1.64)/(1+(X/1.41E+03)^1.21) + 1.64; R^2 = 0.998.

Table 2. Bradford Assay of Protein Content of *I. walleriana* Nectar

	[Protein] in 18% solution (µg/mL)
Replicate #1	88.092
	95.524
	80.594
Replicate #2	79.580
	87.874
	73.986
Mean ± CI[a]	84.275 ± 8.069
	[Protein] in nectar (µg/mL)
Mean * DF	468.14 ± 44.83

[a]Confidence interval at a 95% confidence level

Evaluation of EFN in vitro Antimicrobial Activity

For several different microbes including *Escherichia coli, Staphylococcus aureus* and *Saccharomyces cerevisiae*, the antimicrobial activity of the *I. walleriana* EFN was assayed. *E. coli* and *S. aureus* were chosen to represent gram-negative and gram-positive bacteria, respectively. Of the three microbes selected for assay, *E. coli* was the most extensively assayed due to safety and material constraints. As shown in Figures 8 and 9, 15% aqueous nectar solution inhibited *E. coli* growth consistently up to 2X dilutions given that the nectar solution was diluted once already in the first well. Furthermore, in order to establish unique properties of the EFN, it was screened alongside a similar solution of *I. walleriana* floral nectar (**Figure 8**). The EFN exhibited activity up to 2X dilutions while the floral nectar showed no activity whatsoever. In efforts to identify the source of the exhibited activity in the EFN, 15% aqueous nectar solution was treated with proteinase K with an incubation period prior to sample loading. However, the proteinase K-treated sample exhibited similar mild antimicrobial activity to the untreated EFN solution (**Figure 9**).

With regards to *S. aureus*, the EFN did not appear to have any antimicrobial activity (data not shown). Due to safety constraints, pictures could not be taken of the *S. aureus* microtiter plate.

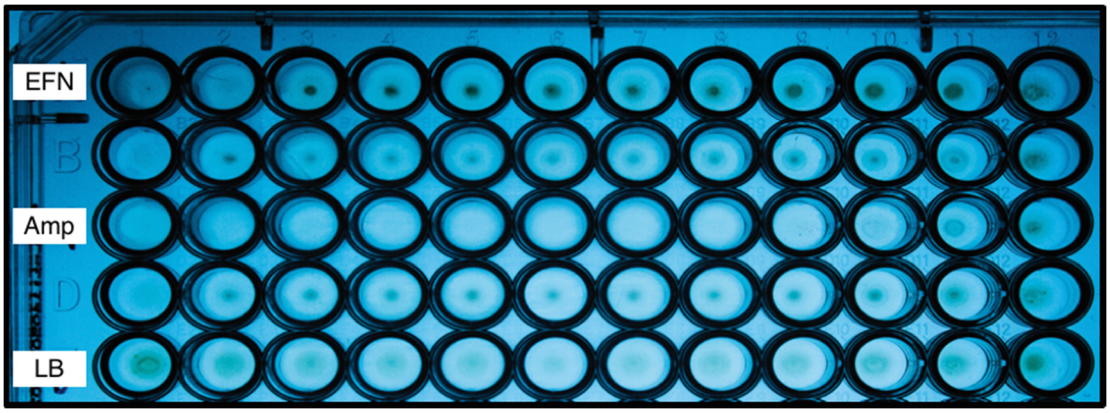

Figure 8. Antimicrobial activity assay conducted with *E. coli*. *Nec* lane contains serial dilutions of 15% EFN aq. solution. Positive control: *Amp* lane contains serial dilutions of ampicillin solution. Negative control: *LB* lane contains lysogeny broth. Unmarked lanes were not used.

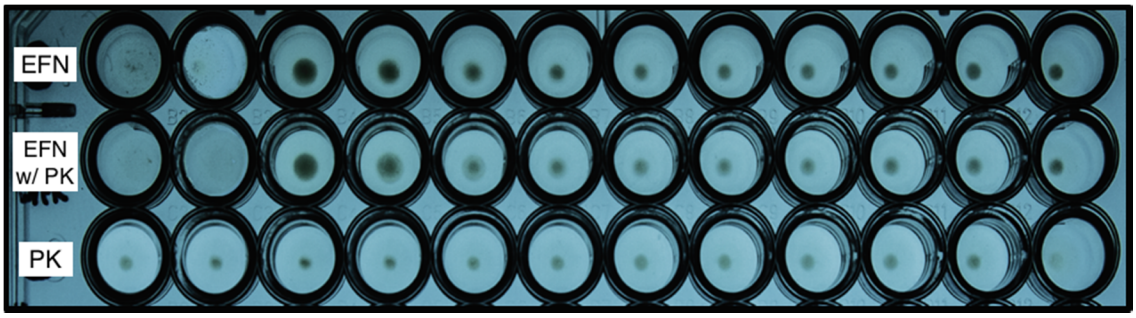

Figure 9. Antimicrobial assay conducted with *E. coli*. *EFN w/ PK* lane contains 15% EFN aqueous solution treated with proteinase K. Positive control: proteinase K-untreated 15% EFN aq. solution. Negative control: lysogeny broth with proteinase K.

Among other microbes, yeasts are most likely to be found in plant nectars; however, some plants such as *Acacia* myrmecophytes produce nectar that inhibits yeast growth (González-Teuber et al., 2009). I therefore used a bioassay to investigate the effects of *I. walleriana* EFN on the growth of yeast cells. First, a suspension of an isolated strain of yeast cells (*Saccharomyces cerevisiae*) was incubated with freshly collected EFN, which was then cultivated on potato dextrose agar plates to quantify the growing yeast in terms of colony-forming units. *I. walleriana* EFN inhibited the growth of yeasts as significantly fewer colony-forming units were found in the EFN than in the

control 5% sugar solution (**Figure 10**). Univariate ANOVA confirmed that the difference between the two was statistically significant ($F(1,4) = 10.816$, $p = 0.03$). It should be noted that other contaminating colonies were present on the plates that were not *S. cerevisiae*, but they were not counted.

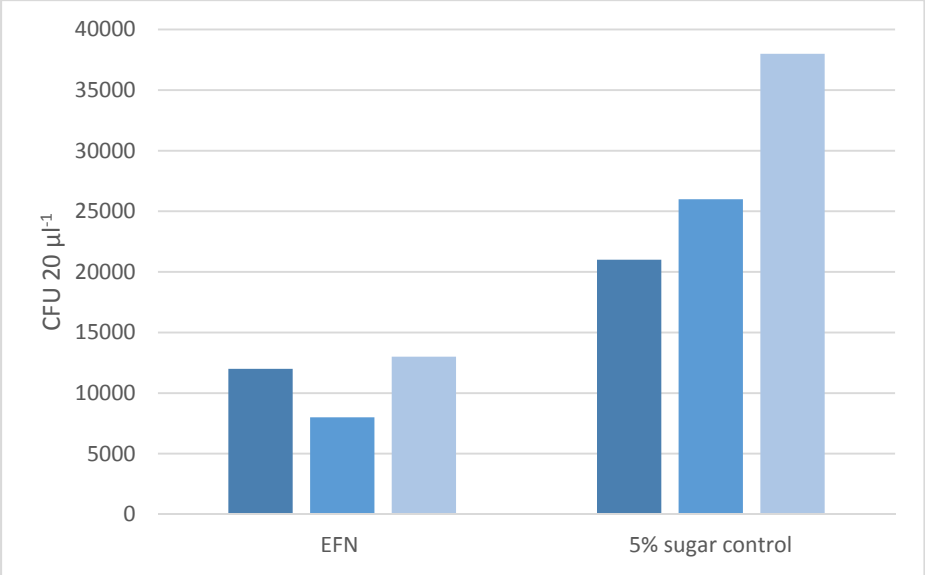

Figure 10. Effect of EFN activities on yeast growth. Yeast growth (number of colony-forming units per 20 µl) in EFN of *I. walleriana* compared to yeast growth in 5% sugar solution (1:1 glucose/fructose) over three different replicate experiments. Statistical analysis by univariate ANOVA confirmed the difference between the two groups to be statistically significant ($p = 0.03 < 0.05$).

Detection and Kinetics of Peroxidase Activity in I. walleriana EFN

In other plants such as *Acacia* myrmecophytes, EFN has been found to contain an abundance of plant pathogenesis-related proteins such as peroxidases, glucanases and chitinases (González-Teuber et al., 2010). As such, a hydrogen peroxide detection protocol was altered to perform a colorimetric assay for the detection of peroxidase activity in the nectar. In all replicates, the *I. walleriana* EFN exhibited a strong colorimetric response in the presence of hydrogen peroxide and a developing solution containing an oxidative indicator, which produces the colorimetric change when oxidized

in the peroxidase reaction (**Figure 11)**. The colorimetric response was rapid and was visually detectable within seconds.

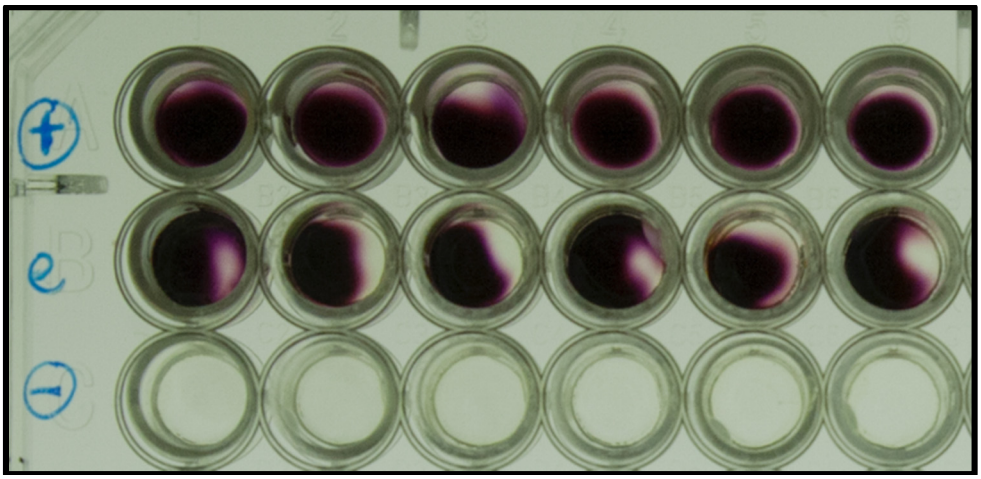

Figure 11. Assay for the colorimetric detection of peroxidase activity. All wells contain a working solution composed of 3% H_2O_2 solution and a developing solution for the colorimetric assay. The solution in *E* lane was introduced to 80 μL of 2X concentrated EFN protein solution and produced the colorimetric change seen above. Positive control: dilute horseradish peroxidase solution (Amresco) was introduced to the working solution. Negative control: working solution with an additional volume of DI water.

The peroxidase activity of the EFN was determined using a standard colorimetric assay with guaiacol as an oxidative indicator. As shown in Table 3, activity was measured in units of Abs min^{-1} μg $protein^{-1}$ at 470 nm using a set of nectar serial dilutions and was measured to be 12.52 ± 1.82 abs min^{-1} μg^{-1}. With an extinction coefficient of 26.6 mM^{-1} cm^{-1} for the oxidation of guaiacol at 470 nm, the specific activity was determined to be 479 μmol min^{-1} μg^{-1}.

Table 3. Peroxidase Activity Assay of Nectar Protein Content

Nectar Dilution Series	abs min^{-1} μg^{-1}	$\mu mol\ min^{-1}\ \mu g^{-1}$
16X	11.18	420
8X	11.10	417
4X	14.01	527
2X	13.79	518
1X	13.61	512
Mean ± CI	12.52 ± 1.82	479 ± 68

In order to determine the protein conferring the peroxidase activity to the nectar, three concentrated EFN samples of decreasing amounts were run through native PAGE gel and subjected to an in-gel peroxidase activity stain. Though in a smear, the activity stain using guaiacol as an oxidative indicator reveals two intense bands of peroxidase activity and, possibly, a faint band towards the bottom of the smear (**Figure 12a**). The subsequence nonspecific Coomassie stain shows the two intense bands as well as confirms the presence of a third major band below (**Figure 12b**).

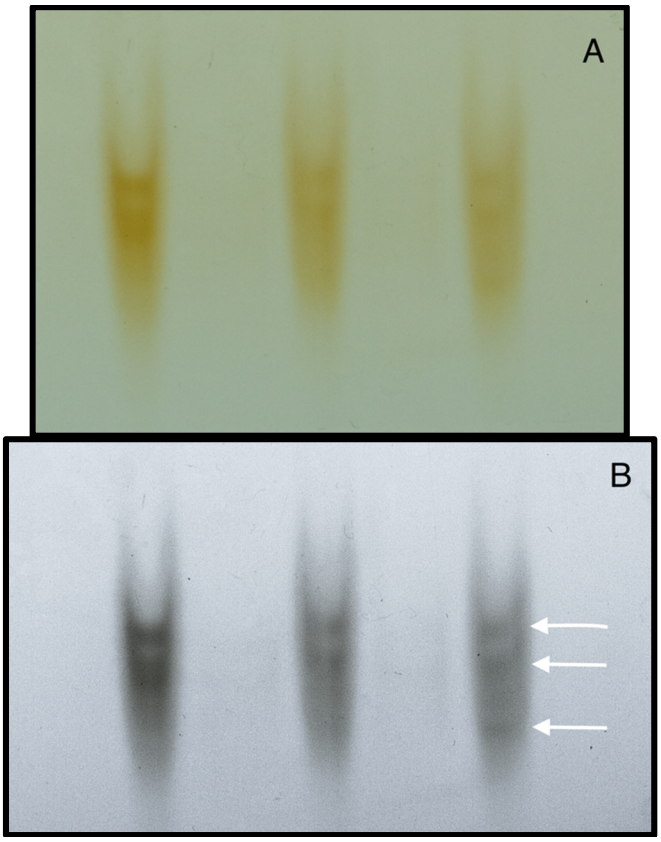

Figure 12. In-gel activity stain (panel A) and subsequent Coomassie Brilliant Blue R-250 stain of native PAGE of *I. walleriana* nectarins. In-gel activity stain using guaiacol (A) as an oxidative indicator revealed several bands of elevated peroxidase activity which were preserved in the Coomassie stain (B).

Peptide Sequencing of I. walleriana Nectarins

In order to isolate peptides for MS/MS and sequencing, we separated the nectarins with 2D gel electrophoresis. Five major protein bands were isolated, excised and prepared for MS/MS analysis (**Figure 13**). Isoelectric focusing prior to the MW separation revealed that the 23 kDa band, which was so heavily expressed as noted in the SDS-PAGE analysis, is a combination of at least four different peptides with varying isoelectric points. After Waters MS/MS, peptide fragment spectra were analyzed with Waters Bio-Lynx software to do bottom-up proteomics *de novo* sequencing. The DDA workflow actively selected precursor peptide ions and analyzed them in iterative cycles of MS/MS analysis where the precursor ions were fragmented via CID and accelerated

through a TOF system. In order for BioLynx to be able to successfully determine the amino acid composition of a peptide, it requires the charge state of the precursor ion and the molecular weight of the singly-charged precursor, or $[M+H]^+$. To gain both pieces of information usually requires the precursor charge state to be deconvoluted as most MS systems lack the resolving power to distinguish isotope peaks of highly charged protein species. With this information in hand and all modifications (e.g. alkylation steps involved in sample preparation) accounted for, BioLynx identifies the C-terminal lysine or arginine residue produced by tryptic digestion and builds the peptide towards the N-terminus by identifying the amino acid mass differences in sequential c and y fragment ions such that the resulting molecular weight matches that of the precursor peptide (**Figure 14**). An MS/MS spectra of a peptide from a tryptic digestion of alcohol dehydrogenase (ADH) is also included to show how a peptide from a positive control looks and scores (**Figure 15**). Comparison of the two spectra reveal the effects of low precursor MS signal; ion charges tend to accumulate on lower molecular weight fragments, often rendering insufficient signal to detect amino acid additions on longer fragment ions. Upwards of thirty sequences were collected from the *I. walleriana* nectar protein fragmentation spectra. Table 4 lists the highest probability sequences from the most promising fragmentation spectra acquired. Though I obtained similar results, the sequences and spectrum displayed below were the best ones to show and collected by Andrew Cox.

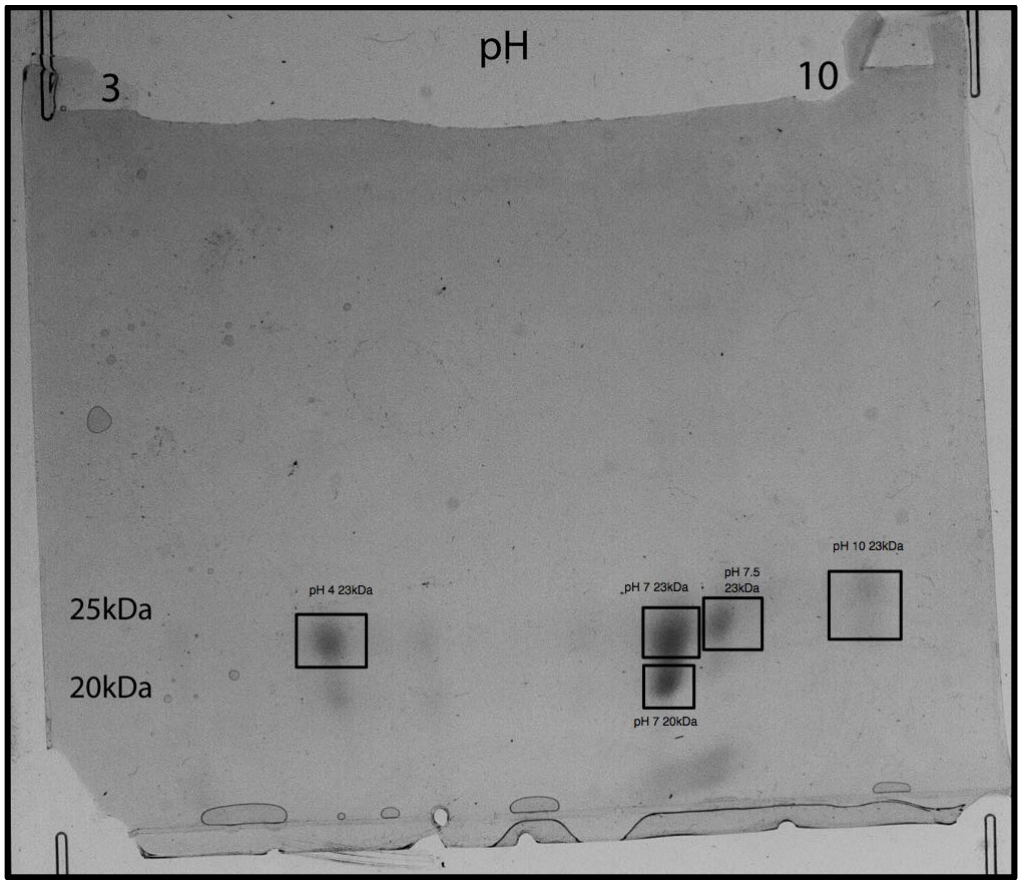

Figure 13. 2D gel electrophoresis separation of *I. walleriana* nectarins. Four bands of 23 kDa in size were isolated with pI's of 4.0, 7.0, 7.5 and 10.0. One band of 20 kDa in size was isolated with a pI of 7.0.

Table 4. Predicted Peptide Sequences

Sequence[a]	Probability[b] (%)
KELSGPQEDVCAKLLSR	76.57
EKLSGPQEDVCAKLLSR	22.36
RYGGLQLVPVGLLTGSTK	54.03
RYGGLQLAPSPLLTGSTK	20.50
RYGGLQLAPPSLLTGSTK	9.06
KSGSGGGYNKDFPK	50.18
SKGSGGGYNKDFPK	49.82
RYNLQLLTSTGK	77.07
YRNLQLLTSTGK	22.93
NFLFGVNQFVPN	50.67
FNLFGVNQFVPN	49.33

[a]Sequences grouped by shading are from same 35recursor ion
[b]Probability of reported sequence being true w.r.t. its precursor ion

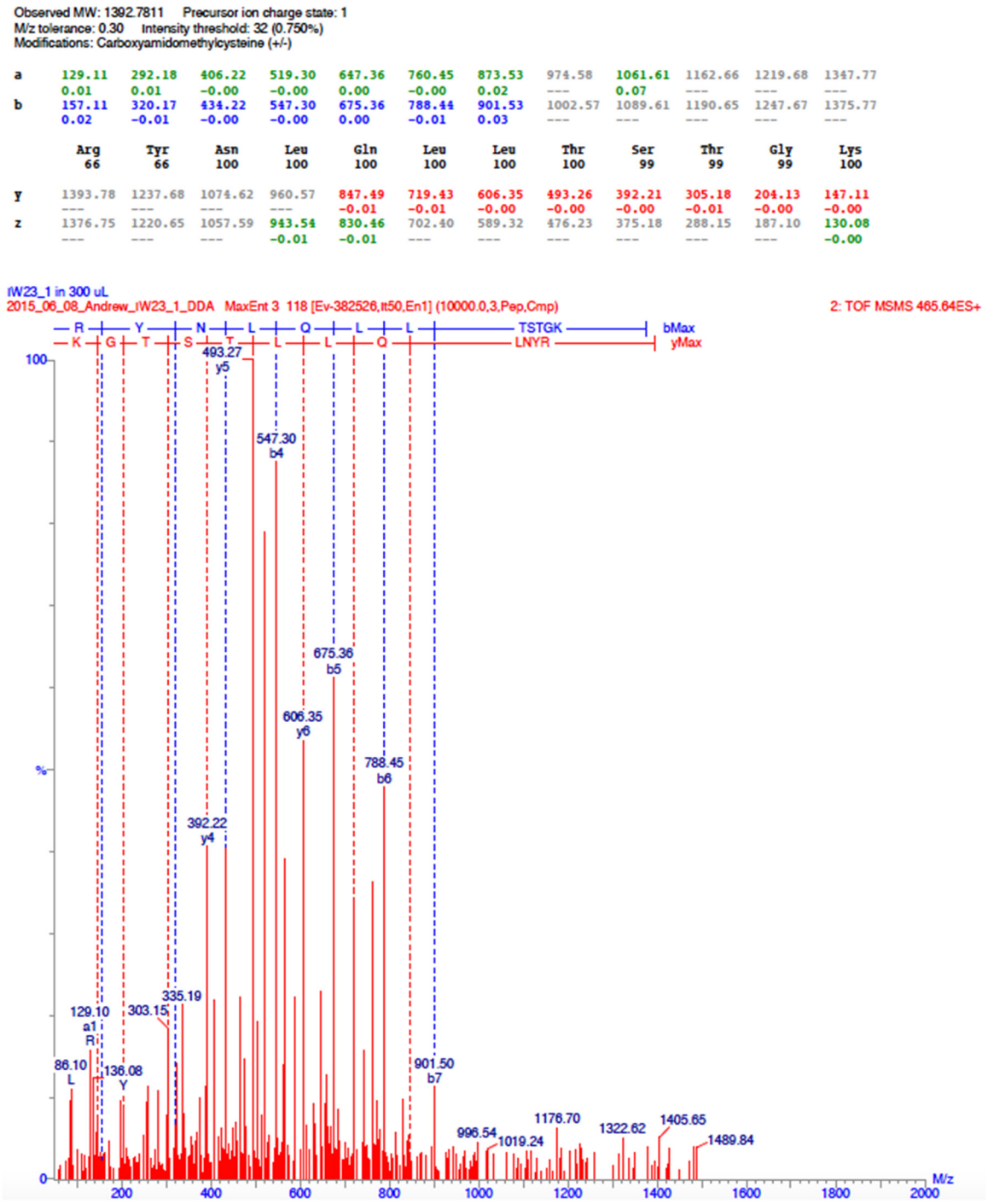

Figure 14. An example of the MS/MS spectra we collected using Waters BioLynx-PepSeq software. The software detected specific mass differences between b and y fragment ions, respectively, and assembles the peptide *de novo* using the information. Noted in the text at the top of the figure, BioLynx takes the precursor ion molecular weight and charge state into account. The red text on the right notes the precursor ion *m/z* peak from which the fragment spectrum was generated.

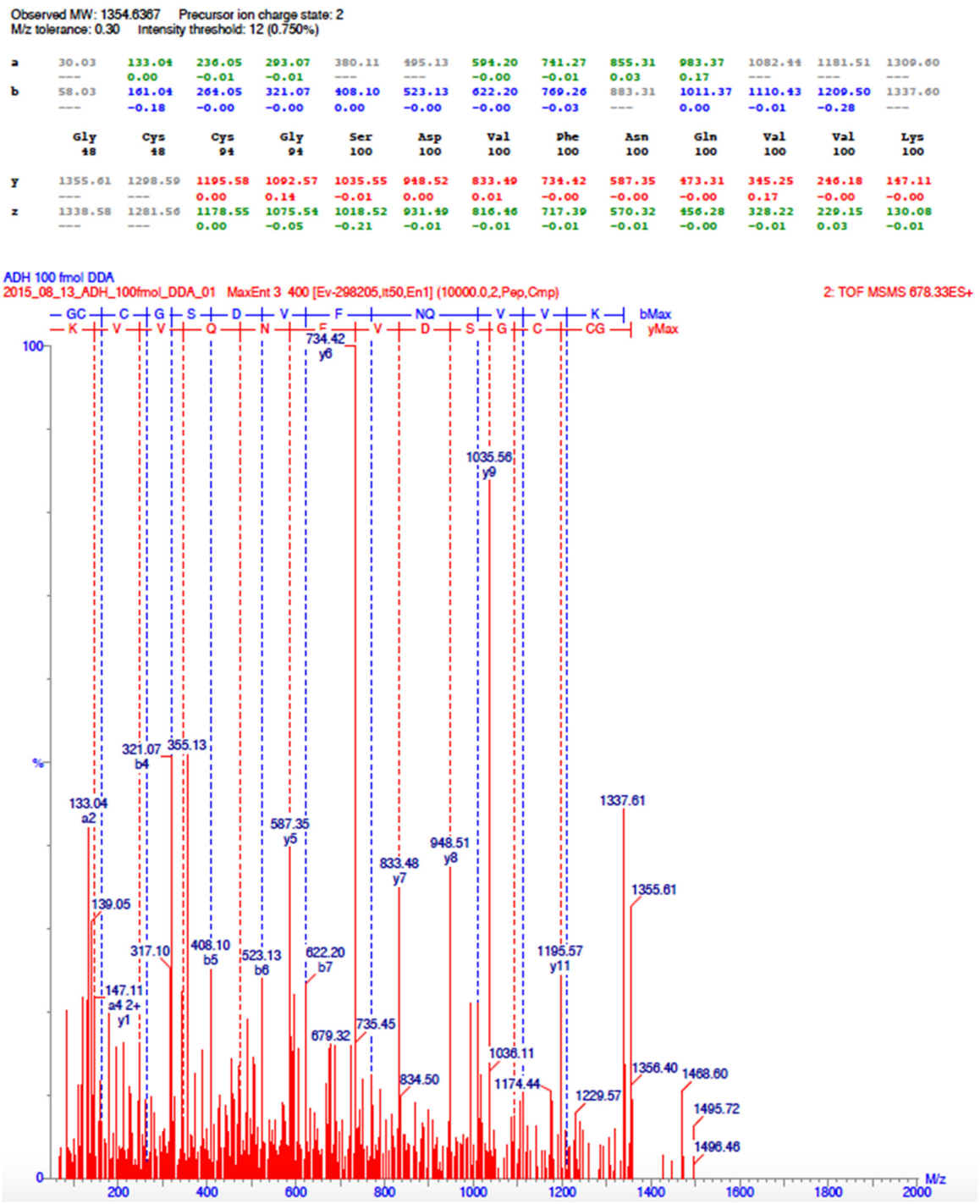

Figure 15. An MS/MS spectrum of a peptide from ADH. Its purposes here are to show a sequenced peptide with full fragment coverage and strong scores.

CHAPTER FOUR

Discussion and Conclusions

Extrafloral nectar-mediated interactions with aggressive defenders such as ants are well documented in regards to a variety of plant species (Heil, 2011). However, these defenders, and pollinators with regards to floral nectar, are not sterile and thus introduce respiring microbes into the metabolite-rich nectar environment (Park and Thornburg, 2009). As extrafloral nectary tissues lack any anatomical protections and offer openings for microbial infection (Keitt and Ivanoff, 1941), they require molecular defense mechanisms inherent to the nectar produced. *Acacia* myrmecophytes, which exhibit strong mutualist behavior with ants, constitutively express a large number of pathogenesis-related proteins such as chitinases, β-1,3-glucanases and peroxidases (González-Teuber et al., 2009). Chitinases and β-1,3-glucanases confer antimicrobial activity against fungus and bacteria via degradation of cell wall components (Sela-Buurlage et al., 1993). Peroxidases have been heavily implicated in plant cell defense, including but not limited to: cell wall reinforcement via oxidative crosslinking, lignin biosynthesis and production of hydrogen peroxide (Brisson et al., 1994; He et al., 2001). This study investigated the antimicrobial activity of *I. walleriana* EFN, such as that observed in the floral nectar of ornamental tobacco (Carter and Thornburg, 2004a, 2000, 2004b), and whether the activity of any pathogenesis-related proteins could be implicated in this defense. These conclusions as well as the acquired peptide sequence data

contribute the the development of *I. walleriana* as a model plant system for the nectar-mediated delivery of transgenic proteins.

The antimicrobial assays of EFN demonstrated a significant amount of activity against the growth of *E. coli in vitro* at concentrations down to 7.5% of secreted nectar concentrations. As referenced above, enzymes expressed in nectar solutions may confer antimicrobial activity indirectly via the production of antimicrobial agents such as hydrogen peroxide or directly by utilizing cell wall components (Carter and Thornburg, 2004b; Kram et al., 2008). In order to determine the manner in which the observed antimicrobial activity was conferred, the antimicrobial activity of EFN treated with Proteinase K was assayed. However, the EFN retained its mild antimicrobial activity after the nectarins were degraded. These results suggest that the agent directly conferring antimicrobial activity against *E. coli* is non-proteinaceous, resembling the mode of action preventing microbial infestation in ornamental tobacco via hydrogen peroxide production. However, it must be noted that the EFN did not have the same activity against gram-positive *S. aureus*. Instead, the EFN caused *S. aureus* to grow in a large, cloudy smear in the bottom of the well as opposed to the tight colonies observed with *E. coli* (data not shown). This is suspected be the effect of the high osmotic pressure environment created by high sugar concentrations in the nectar. This may have had a greater effect on *E. coli* resulting in the observed activity, but the specific mode of action remains to be investigated.

The EFN had much more pronounced effects on the yeast growth. Incubating yeast cells with the EFN solution prior to plating reduced the viable colony-forming units by 75% relative to the sugar control. In several other studies, antifungal activity is most

commonly attributed to pathogenesis-related proteins such as chitinases and β-1,3-glucanases (González-Teuber et al., 2009; Sela-Buurlage et al., 1993). As mentioned above, peroxidases are often found among that group as a family of enzymes nearly always expressed in plant tissues upon infection or wounding (Orozco-Cardenas and Ryan, 1999). The results of the assay of yeast growth implicate the presence and expression of such pathogenesis-related proteins in the EFN, which was proven true with the detection of peroxidase activity. After detection, the specific activity (per μg total protein) was quantified to provide a baseline for further study of enzymatic activity.

Activity staining experiments with native PAGE attempted to attribute peroxidase activity to a specific protein band. However, activity staining with guaiacol revealed three distinct bands of high intensity and non-specific Coomassie staining revealed no other peptides. This may be due to the presence of large tertiary or quaternary protein structures inhibiting migration by varying degrees depending on orientation in the gel matrix. In a separate experiment, a denaturing SDS-PAGE gel containing EFN nectarins was subject to a renaturing protocol using Triton X-100 (Thermo Fisher) and then to the same peroxidase staining protocol without positive results, which suggests quaternary structure is necessary for peroxidase activity (figure not shown). The separation of multiple protein bands at 23 kDa by isoelectric focusing in 2D gel electrophoresis further supports this claim. The high pI's, between 7.0 and 10.0, of the majority of the protein bands may have caused the non-uniform migration of the nectarins as the rate of migration in native PAGE is determined by the proteins inherent acidity in the PAGE buffer system. Furthermore, the pI of the functional protein complexes may be

influenced by varying degrees of protein glycosylation in the high sugar concentration environment.

The results of this study suggest that the majority, if not all, of the EFN soluble protein content contributes to peroxidase activity in the nectar, which is the likely source of the observed antifungal activity. However, several questions remain to be investigated. It is not clear as to what provided the source of antimicrobial activity against *E. coli*, except that it may not be attributed to direct enzymatic activity acting on bacterial components as substrate. However, the significant amount of disparity in colony-forming units produced by the EFN on yeast growth suggest that a more direct mode of action may be involved here, which may be facilitated by the high levels of peroxidase activity detected in the nectar.

Going forward, the detection of peroxidase activity yields the Kearney lab an analytical target as it searches for the promoter responsible for increased expression of the 23 kDa protein band in nectary tissues. Amino acid and, therefore, DNA sequences associated with enzyme active sites and substrate binding regions are much more likely to be conserved. If found, such sequences would provide the basis for promising PCR primers which, if successfully annealed to genomic DNA, provide a preferable starting position from which we may chromosome walk to the desired promoter sequence. The peptide sequences identified by the MS/MS DDA proteomics workflow provide the basis from which to search for these associations and others.

After MS/MS data is acquired and processed to yield peptide sequence data, there are several different means by which the precursor protein(s) may be identified. The most common method used to identify peptides following the acquisition of MS/MS

spectra is protein sequence database searching. To this end, protein database search engines have been designed, including but not limited to Mascot (Perkins et al., 1999), X!Tandem (Craig and Beavis, 2004), Sequest (Eng et al., 1994), and MyriMatch (Tabb et al., 2007). Public protein sequence databases such as UniProt, Swiss-Prot, and ENSEMBL are used as reference databases for these search engines. Though, however convenient for routine use, this method for protein/peptide identification often falls short and did so in work outside of this study; Mascot searches through UniProt and Swiss-Prot databases with our MS/MS-derived peptide sequences yielded inconclusive results. Given that protein sequence databases hold collections of all known protein sequences from any given organism, they may not closely represent the specific protein pool in a given sample. Large databases, such as those listed above, yield reduced sensitivity, lower signal-to-noise ratio, and more distraction when maintained under search criteria necessary to minimize false discovery rates (Nesvizhskii, 2010). Additionally, these protein databases may be incomplete in regards to sequence variation information, such as single nucleotide variations (SNVs) and RNA-splicing and -editing variants, which can account for distinct phenotypic variation (Wang et al., 2012). Without taking such variance into account, proteomics studies could easily miss novel protein forms. Some attempts have been made to account for sequence variation by integrating genomic variation information from databases such as dbSNP and COSMIC (Alves et al., 2008; Bunger et al., 2007; Schandorff et al., 2007) and inferring splice variants from the Expressed Sequence Tags (EST) database (Chang et al., 2010; Fermin et al., 2006). These incorporations enable detection of otherwise lost protein variants but at the cost of severely enlarging the database and higher risk of false positives (Bunger et al., 2007).

In order to gain insight into unique cellular systems at both the genomic and proteomic levels, researchers have recently turned towards utilizing RNA and protein profiling techniques in parallel (Adamidi et al., 2011; Desgagné-Penix et al., 2010). Emerging high-throughput RNA sequencing (RNA-Seq) technology promises an unbiased and possibly comprehensive look into sample-specific transcriptomes including transcript expression levels and sequence variations (Lundberg et al., 2010). Thus, a sample-specific protein sequence database derived from RNA-Seq data has the capability of much better representing the protein pool of a sample, as opposed to a public resource protein sequence database, without having to be excessively large. Additionally, RNA-Seq gives insight into the various levels of protein expression in different tissues, though protein abundance can only be partially described by mRNA concentration (Vogel et al., 2010).

Two alternate computational strategies for transcriptome reconstruction are in use (Haas and Zody, 2010). The first approach aligns all mRNA sequence reads to an unannotated reference genome and then merges sequences with overlapping alignment (Yassour et al., 2009). Mapping-first approaches include those found in Scripture and Cufflinks (Guttman et al., 2010; Trapnell et al., 2010). Alternately, assembly-first approaches (*de novo*) directly use the mRNA reads to assemble the transcriptome in the absence of a reference genome, which is useful when the genomic sequence is unavailable, gapped, fragmented or highly altered, as in the case of cancer cells (Grabherr et al., 2011). Between the two approaches, mapping-first approaches have undergone significantly more development than *de novo* approaches. This is partly due to the increased difficulty of assembling continuous sequence contigs as the number of reads

grow; it becomes increasingly difficult to determine which reads to connect to contigs as the number of short mRNA reads increases (Grabherr et al., 2011).

Impatiens walleriana is not a model system and its genome has not been sequenced. Thus, the most direct approach to assembling a RNA-Seq derived protein sequence library is a *de novo* approach. One method that has demonstrated efficient ability to reconstruct transcriptomes *de novo* is Trinity. In response to the difficulties mentioned above, Trinity takes advantage of a computational solution given by the de Bruijin graph (De Bruijin, 1946; Good, 1946), which already provides the foundation for several genome assembly methods (Butler et al., 2008; Zerbino and Birney, 2008). In this graph, contigs are seeded with a node defined by a sequence of a fixed length of *k* nucleotides (denoted '*k*-mer'), which are connected with other *k*-mer's on the edges provided they overlap by k-1 nucleotides, yielding an efficient presentation of all possible ways by which linear sequences might be reconstructed given k-1 overlap.

The Trinity method of *de novo* transcriptome reconstruction consists of three software modules: Inchworm, Chrysalis and Butterfly. In the first step, Inchworm first constructs and refines the mRNA *k*-mer 'dictionary', excluding likely error-containing *k*-mers and those that only appear once. Inchworm then seeds and extends in the manner described above, expanding one contig on both sides until no more k-1 overlaps are available. Next, Chrysalis pools related contigs that share at least one k-1 overlap and constructs a de Bruijin graph for each pool of contigs. These graphs display the overlaps between variants caused by alternative RNA splicing or unique portions of paralogous genes. Lastly, Butterfly uses the de Bruijin graphs constructed in Chrysalis and the original reads to construct full length transcripts, including those arising from alternative

splicing and paralogous genes. This is accomplished by a dynamic scoring procedure by which Butterfly identifies paths taken in de Bruijin graphs and scores them as they are supported by original reads. In the end, Trinity yields a highly accurate transcriptome reconstruction, ready for use as a sample-specific protein sequence database to by queried with MS/MS-derived peptide sequences. Upon receiving the RNA-seq libraries generated from *Impatiens walleriana* leaf, root and nectary tissues, we will use Trinity or another *de novo* transcriptome assembly method to elucidate transcripts and thus peptide expression specific to the nectary tissues. These discoveries might allow us to design successful RNA primers which will enable us to isolate a nectary specific promoter that might express a transgenic mosquitocidal proteins in lethal quantities.

BIBLIOGRAPHY

Adamidi, C., Wang, Y., Gruen, D., Mastrobuoni, G., You, X., Tolle, D., Dodt, M., Mackowiak, S.D., Gogol-Doering, A., Oenal, P., et al. (2011). De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics. Genome Res. *21*, 1193–1200.

Adler, L.S., and Irwin, R.E. (2005). Ecological Costs and Benefits of Defenses in Nectar. Ecology *86*, 2968–2978.

Adler, L.S., Wink, M., Distl, M., and Lentz, A.J. (2006). Leaf herbivory and nutrients increase nectar alkaloids. Ecol. Lett. *9*, 960–967.

Alves, G., Ogurtsov, A.Y., and Yu, Y.-K. (2008). RAId_DbS: mass-spectrometry based peptide identification web server with knowledge integration. BMC Genomics *9*, 1–12.

Baker, H.G., and Baker, I. (1973). Amino-acids in Nectar and their Evolutionary Significance. Nature *241*, 543–545.

Baker, H.G., and Baker, I. (1975). Studies of nectar-constitution and pollinator–plant coevolution. In Coevolution of Animals and Plants, L.E. Gilbert, and P.H. Raven, eds. (Austin: University of Texas), pp. 126–152.

Baker, I., and Baker, H.G. (1976). ANALYSES OF AMINO ACIDS IN FLOWER NECTARS OF HYBRIDS AND THEIR PARENTS, WITH PHYLOGENETIC IMPLICATIONS. New Phytol. *76*, 87–98.

Beier, J.C., Killeen, G.F., and Githure, J.I. (1999). Short report: entomologic inoculation rates and Plasmodium falciparum malaria prevalence in Africa. Am. J. Trop. Med. Hyg. *61*, 109–113.

Beier, J.C., Keating, J., Githure, J.I., Macdonald, M.B., Impoinvil, D.E., and Novak, R.J. (2008). Integrated vector management for malaria control. Malar J *7*.

Benesch, J.L.P., Ruotolo, B.T., Simmons, D.A., and Robinson, C.V. (2007). Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. Chem. Rev. *107*, 3544–3567.

Brisson, L.F., Tenhaken, R., and Lamb, C. (1994). Function of Oxidative Cross-Linking of Cell Wall Structural Proteins in Plant Disease Resistance. Plant Cell *6*, 1703–1712.

Bunger, M.K., Cargile, B.J., Sevinsky, J.R., Deyanova, E., Yates, N.A., Hendrickson, R.C., and Stephenson, J.L. (2007). Detection and Validation of Non-synonymous Coding SNPs from Orthogonal Analysis of Shotgun Proteomics Data. J. Proteome Res. *6*, 2331–2340.

Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I.A., Belmonte, M.K., Lander, E.S., Nusbaum, C., and Jaffe, D.B. (2008). ALLPATHS: De novo assembly of whole-genome shotgun microreads. Genome Res. *18*, 810–820.

Carter, C.J., and Thornburg, R.W. (2004a). Tobacco Nectarin III is a Bifunctional Enzyme with Monodehydroascorbate Reductase and Carbonic Anhydrase Activities. Plant Mol. Biol. *54*, 415–425.

Carter, C., and Thornburg, R.W. (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., and Thornburg, R.W. (2004b). Is the nectar redox cycle a floral defense against microbial attack? Trends Plant Sci. *9*, 320–324.

Carter, C., Healy, R., O'Tool, N.M., Naqvi, S.M.S., Ren, G., Park, S., Beattie, G.A., Horner, H.T., and Thornburg, R.W. (2007). Tobacco Nectaries Express a Novel NADPH Oxidase Implicated in the Defense of Floral Reproductive Tissues against Microorganisms. Plant Physiol. *143*, 389–399.

Chang, K.-Y., Georgianna, D.R., Heber, S., Payne, G.A., and Muddiman, D.C. (2010). Detection of Alternative Splice Variants at the Proteome Level in Aspergillus flavus. J. Proteome Res. *9*, 1209–1217.

Chapman, J.D., Goodlett, D.R., and Masselon, C.D. (2014). Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. Mass Spectrom. Rev. *33*, 452–470.

Chen, Z., and Kearney, C.M. (2015). Nectar protein content and attractiveness to Aedes aegypti and Culex pipiens in plants with nectar/insect associations. Acta Trop. *146*, 81–88.

Chernushevich, I.V., Loboda, A.V., and Thomson, B.A. (2001). An introduction to quadrupole–time-of-flight mass spectrometry. J. Mass Spectrom. *36*, 849–865.

Coon, J.J. (2009). Collisions or Electrons? Protein Sequence Analysis in the 21st Century. Anal. Chem. *81*, 3208–3215.

Craig, R., and Beavis, R.C. (2004). TANDEM: matching proteins with tandem mass spectra. Bioinformatics *20*, 1466–1467.

Deans, S.G., and Waterman, P.G. (1993). Biological activity of volatile oils. In Volatile Oil Crops: Their Biology, Biochemistry, and Production, R.K.M. Hay, and P.G. Waterman, eds. (Essex, England: Longman Scientific and Technical), pp. 97–111.

De Bruijin, A.G. (1946). A combinatorical problem. K. Ned. Akad. V Wet. 758–764.

Deinzer, M.L., Thomson, P.A., Burgett, D.M., and Isaacson, D.L. (1977). Pyrrolizidine Alkaloids: Their Occurrence in Honey from Tansy Ragwort (Senecio jacobaea L.). Science *195*, 497–499.

Desgagné-Penix, I., Khan, M.F., Schriemer, D.C., Cram, D., Nowak, J., and Facchini, P.J. (2010). Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures. BMC Plant Biol. *10*, 252.

Ecroyd, C.E., Franich, R.A., Kroese, H.W., and Steward, D. (1995). Volatile constituents of Dactylanthus taylorii flower nectar in relation to flower pollination and browsing by animals. Phytochemistry *40*, 1387–1389.

Elias, T.S., and Gelband, H. (1975). Nectar: Its production and functions in trumpet creeper. Science *189*, 289–291.

Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom. *5*, 976–989.

Escalante-Pérez, M., and Heil, M. (2013). The Production and Protection of Nectars. In Progress in Botany, U. Lüttge, W. Beyschlag, D. Francis, and J. Cushman, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 239–261.

Farkas, Á., Mihalik, E., Dorgai, L., and Bubán, T. (2011). Floral traits affecting fire blight infection and management. Trees *26*, 47–66.

Ferguson, P.L., and Smith, and R.D. (2003). Proteome Analysis by Mass Spectrometry. Annu. Rev. Biophys. Biomol. Struct. *32*, 399–424.

Fermin, D., Allen, B.B., Blackwell, T.W., Menon, R., Adamski, M., Xu, Y., Ulintz, P., Omenn, G.S., and States, D.J. (2006). Novel gene and gene model detection using a whole genome open reading frame analysis in proteomics. Genome Biol. *7*, 1–13.

Fernandes, L., and Briegel, H. (2005). Reproductive physiology of Anopheles gambiae and Anopheles atroparvus. J. Vector Ecol. *30*, 11–26.

Ferreres, F., Andrade, P., Gil, M.I., and Tomàs-Barberàn, F.A. (1996). Floral nectar phenolics as biochemical markers for the botanical origin of heather honey. Z. Für Lebensm.-Unters. -Forsch. *202*, 40–44.

Fieldes, A. (1992). Using coomassie blue to stabilize H2O2-guaiacol stained peroxidases on polyacrylamide gels. Electrophoresis *13*, 454–455.

Foster, W. (2008). Phytochemicals as population sampling lures. J Am Mosq Contr Assoc *24*.

Foster, W.A. (1995). Mosquito sugar feeding and reproductive energetics. Ann Rev Entomol *40*.

Gary, R.E., and Foster, W.A. (2004). Anopheles gambiae feeding and survival on honeydew and extra-floral nectar of peridomestic plants. Med. Vet. Entomol. *18*, 102–107.

Gillet, L.C., Leitner, A., and Aebersold, R. (2015). Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis-Testing. Annu. Rev. Anal. Chem.

Giri, A.P., Wünsche, H., Mitra, S., Zavala, J.A., Muck, A., Svatoš, A., and Baldwin, I.T. (2006). Molecular Interactions between the Specialist Herbivore Manduca sexta (Lepidoptera, Sphingidae) and Its Natural Host Nicotiana attenuata. VII. Changes in the Plant's Proteome. Plant Physiol. *142*, 1621–1641.

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

González-Teuber, M., Pozo, M.J., Muck, A., Svatos, A., Adame-Álvarez, R.M., and Heil, M. (2010). Glucanases and Chitinases as Causal Agents in the Protection of Acacia Extrafloral Nectar from Infestation by Phytopathogens. Plant Physiol. *152*, 1705–1715.

Good, I.J. (1946). Normal Recurring Decimals. J. Lond. Math. Soc. *s1-21*, 167–169.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

Griebel, C., and Heß, G. (1940). The vitamin C content of flower nectar of certain Labiatae. Z. Für Unters. Lebensm. *79*, 168–171.

Gu, W., Müller, G., Schlein, Y., Novak, R.J., and Beier, J.C. (2011). Natural Plant Sugar Sources of Anopheles Mosquitoes Strongly Impact Malaria Transmission Potential. PLoS ONE *6*, e15996.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., et al. (2010). Ab initio reconstruction of

cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat. Biotechnol. *28*, 503–510.

Haas, B.J., and Zody, M.C. (2010). Advancing RNA-Seq analysis. Nat. Biotechnol. *28*, 421–423.

Hammerschmidt, R., Nuckles, E.M., and Kuć, J. (1982). Association of enhanced peroxidase activity with induced systemic resistance of cucumber to Colletotrichum lagenarium. Physiol. Plant Pathol. *20*, 73–82.

Harris, D.C. (2010). Quantitative Chemical Analysis (New York: W.H. Freeman and Co).

He, C., Hsiang, T., and Wolyn, D.J. (2001). Activation of defense responses to Fusarium infection in Asparagus densiflorus. Eur. J. Plant Pathol. 473–483.

Heil, M. (2011). Nectar: generation, regulation and ecological functions. Trends Plant Sci. *16*, 191–200.

Heinrich, G. (1989). Analysis of cations in nectars by means of a laser microprobe mass analyser (LAMMA). Beitr Biol Pflanz 293–308.

Ivanoff, S.S., and Keitt, G.W. (1941). Relations of nectar concentration to growth of Erwinia amylovora and fire blight infection of apple and pear blossoms. J. Agric. Res. *62*, 733–743.

Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., and Watson, J.T. (1987). Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. Anal. Chem. *59*, 2621–2625.

Keitt, G.W., and Ivanoff, S.S. (1941). Transmission of fire blight by bees and its relation to nectar concentration of apple and pear blossoms. J. Agric. Res. *62*, 745–753.

Konermann, L., Ahadi, E., Rodriguez, A.D., and Vahidi, S. (2013). Unraveling the Mechanism of Electrospray Ionization. Anal. Chem. *85*, 2–9.

Konijnenberg, A., Butterer, A., and Sobott, F. (2013). Native ion mobility-mass spectrometry and related methods in structural biology. Biochim. Biophys. Acta BBA - Proteins Proteomics *1834*, 1239–1256.

Kram, B.W., Bainbridge, E.A., Perera, M.A.D.N., and Carter, C. (2008). Identification, cloning and characterization of a GDSL lipase secreted into the nectar of Jacaranda mimosifolia. Plant Mol. Biol. *68*, 173–183.

Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature *227*, 680–685.

Lanza, J., Vargo, E.L., Pulim, S., and Chang, Y.Z. (1993). Preferences of the Fire Ants Solenopsis invicta and S. geminata (Hymenoptera: Formicidae) for Amino Acid and Sugar Components of Extrafloral Nectars. Environ. Entomol. *22*, 411–417.

Lössl, P., Snijder, J., and Heck, A.J.R. (2014). Boundaries of Mass Resolution in Native Mass Spectrometry. J. Am. Soc. Mass Spectrom. *25*, 906–917.

Lundberg, E., Fagerberg, L., Klevebring, D., Matic, I., Geiger, T., Cox, J., Älgenäs, C., Lundeberg, J., Mann, M., and Uhlen, M. (2010). Defining the transcriptome and proteome in three functionally different human cell lines. Mol. Syst. Biol. *6*.

McAlister, G.C., Phanstiel, D., Good, D.M., Berggren, W.T., and Coon, J.J. (2007). Implementation of Electron-Transfer Dissociation on a Hybrid Linear Ion Trap–Orbitrap Mass Spectrometer. Anal. Chem. *79*, 3525–3534.

Mehmood, S., Allison, T.M., and Robinson, C.V. (2015). Mass Spectrometry of Protein Complexes: From Origins to Applications. Annu. Rev. Phys. Chem. *66*, 453–474.

Mostowy, W.M., and Foster, W.A. (2004). Antagonistic effects of energy status on meal size and egg-batch size of Aedes aegypti (Diptera: Culicidae). J. Vector Ecol. *29*, 84–93.

Müller, G.C., and Schlein, Y. (2006). Sugar questing mosquitoes in arid areas gather on scarce blossoms that can be used for control. Int J Parasitol *36*.

Müller, G.C., and Schlein, Y. (2008a). Efficacy of toxic sugar baits against adult cistern-dwelling Anopheles claviger. Trans. R. Soc. Trop. Med. Hyg. *102*, 480–484.

Müller, G.C., and Schlein, Y. (2008b). Efficacy of toxic sugar baits against adult cistern-dwelling Anopheles claviger. Trans R Soc Trop Med Hyg *102*.

Müller, G.C., Kravchenko, V.D., and 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 Mosqu Contr Assoc *24*.

Müller, G.C., Beier, J.C., Traore, S.F., Toure, M.B., Traore, M.M., Bah, S., Doumbia, S., and Schlein, Y. (2010). Successful field trial of attractive toxic sugar bait (ATSB) plant-spraying methods against malaria vectors in the Anopheles gambiae complex in Mali, West Africa. Malar. J. *9*, 1–7.

Nesvizhskii, A.I. (2010). A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. J. Proteomics *73*, 2092–2123.

Orozco-Cardenas, M., and Ryan, C.A. (1999). Hydrogen peroxide is generated systemically in plant leaves by wounding and systemin via the octadecanoid pathway. Proc. Natl. Acad. Sci. *96*, 6553–6557.

Park, S., and Thornburg, R.W. (2009). Biochemistry of Nectar Proteins. J. Plant Biol. *52*, 27–34.

Perkins, D.N., Pappin, D.J.C., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551–3567.

Peter Kevan, D.E. (1988). Yeast-Contaminated Nectar and its Effects on Bee Foraging. J. Apic. Res. *27*, 26–29.

Picotti, P., and Aebersold, R. (2012). Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat. Methods *9*, 555–566.

Roshchina, V.V., and Roshchina, V.D. (1993). The Excretory Function of Higher Plants (Berlin, Heidelberg: Springer Berlin Heidelberg).

Schandorff, S., Olsen, J.V., Bunkenborg, J., Blagoev, B., Zhang, Y., Andersen, J.S., and Mann, M. (2007). A mass spectrometry-friendly database for cSNP identification. Nat. Methods *4*, 465+.

Sela-Buurlage, M.B., Ponstein, A.S., Bres-Vloemans, S.A., Melchers, L.S., Elzen, P. van den, and Cornelissen, B. (1993). Only Specific Tobacco (Nicotiana tabacum) Chitinases and [beta]-1,3-Glucanases Exhibit Antifungal Activity. Plant Physiol. *101*, 857–863.

Shah, M., Teixeira, F.M., Soares, E.L., Soares, A.A., Carvalho, P.C., Domont, G.B., Thornburg, R.W., Nogueira, F.C.S., and Campos, F.A.P. (2016). Time-course proteome analysis of developing extrafloral nectaries of Ricinus communis. PROTEOMICS *16*, 629–633.

Stevenson, T.I., Loo, J.A., and Greis, K.D. (1998). Coupling Capillary High-Performance Liquid Chromatography to Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry and N-Terminal Sequencing of Peptides via Automated Microblotting onto Membrane Substrates. Anal. Biochem. *262*, 99–109.

Tabb, D.L., Fernando, C.G., and Chambers, M.C. (2007). MyriMatch: Highly Accurate Tandem Mass Spectral Peptide Identification by Multivariate Hypergeometric Analysis. J. Proteome Res. *6*, 654–661.

Taylor, G. (1964). Disintegration of Water Drops in an Electric Field. Proc. R. Soc. Math. Phys. Eng. Sci. *280*, 383–397.

T. Buban, and Zs. Orosz-Kovacs (2003). The nectary as the primary site of infection by Erwinia amylovora(Burr.) Winslow et al.: a mini review. Plant Syst. Evol. *238*, 183–194.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and

quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Vandermarliere, E., Mueller, M., and Martens, L. (2013). Getting intimate with trypsin, the leading protease in proteomics. Mass Spectrom. Rev. *32*, 453–465.

Vogel, S. Flowers offering fatty oil instead of nectar. In Abstracts XIth International Botany Congress, (Seattle, WA),.

Vogel, C., de Sousa Abreu, R., Ko, D., Le, S.-Y., Shapiro, B.A., Burns, S.C., Sandhu, D., Boutz, D.R., Marcotte, E.M., and Penalva, L.O. (2010). Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol. Syst. Biol. *6*.

Wang, X., Slebos, R.J.C., Wang, D., Halvey, P.J., Tabb, D.L., Liebler, D.C., and Zhang, B. (2012). Protein Identification Using Customized Protein Sequence Databases Derived from RNA-Seq Data. J. Proteome Res. *11*, 1009–1017.

Watson, J.T., and Sparkman, O.D. (2007). Introduction to Mass Spectrometry (Chichester, UK: John Wiley & Sons, Ltd).

Xu, F.F., and Chen, J. (2010). Competition hierarchy and plant defense in a guild of ants on tropical Passiflora. Insectes Sociaux *57*, 343–349.

Yassour, M., Kaplan, T., Fraser, H.B., Levin, J.Z., Pfiffner, J., Adiconis, X., Schroth, G., Luo, S., Khrebtukova, I., Gnirke, A., et al. (2009). Ab initio construction of a eukaryotic transcriptome by massively parallel mRNA sequencing. Proc. Natl. Acad. Sci. *106*, 3264–3269.

Yates, J.R., Ruse, C.I., and Nakorchevsky, A. (2009). Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. Annu. Rev. Biomed. Eng. *11*, 49–79.

Zenobi, R., and Knochenmuss, R. (1998). Ion formation in MALDI mass spectrometry. Mass Spectrom. Rev. *17*, 337–366.

Zerbino, D.R., and Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. *18*, 821–829.

Zhang, Y., Fonslow, B.R., Shan, B., Baek, M.-C., and Yates, J.R. (2013). Protein Analysis by Shotgun/Bottom-up Proteomics. Chem. Rev. *113*, 2343–2394.

(2004). Global Strategic Framework for Integrated Vector Management (Geneva, Switzerland: World Health Organization).

(2006). Malaria vector control and personal protection: report of a WHO Study Group [on Malaria Vector Control and Personal Protection ; Geneva, 12 - 14 March 2004] (Geneva, Switzerland).

(2008a). Global malaria control and elimination : report of a technical review (Geneva, Switzerland).

(2008b). WHO position statement on integrated vector management.