2018

IDENTIFYING MECHANISMS OF HOST PLANT SPECIALIZATION IN *APHIS CRACCIVORA* AND ITS BACTERIAL SYMBIONTS

Thorsten Hansen
University of Kentucky, hanse125@purdue.edu
Digital Object Identifier: https://doi.org/10.13023/ETD.2018.018

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Thorsten Hansen, Student

Dr. Jennifer A. White, Major Professor

Dr. Charles W. Fox, Director of Graduate Studies

IDENTIFYING MECHANISMS OF HOST PLANT SPECIALIZATION IN *APHIS CRACCIVORA* AND ITS BACTERIAL SYMBIONTS

THESIS

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in the College of Agriculture, Food and Environment at the University of Kentucky

By

Thorsten Hansen

Lexington, Kentucky

Director: Dr. Jennifer A. White, Professor of Entomology

Lexington, Kentucky

2017

ABSTRACT OF THESIS

IDENTIFYING MECHANISMS OF HOST PLANT SPECIALIZATION IN *APHIS CRACCIVORA* AND ITS BACTERIAL SYMBIONTS

Many insects form close relationships with microbial symbionts. Insect symbionts can provide novel phenotypes to their hosts, including influencing dietary breadth. In the polyphagous cowpea aphid, *Aphis craccivora*, the facultative symbiont *Arsenophonus* improves aphid performance on one host plant (locust), but decreases performance on other plants. The goal of my thesis was to investigate the mechanism by which *Arsenophonus* facilitates use of locust. First, I assembled an *Aphis craccivora-Arsenophonus-Buchnera* reference transcriptome to conduct RNAseq analysis, comparing gene expression in aphids feeding on locust and fava, with and without *Arsenophonus* infection. Overall, few transcripts were differentially expressed. However, genes that were differentially expressed mapped to a variety of processes, including metabolism of glucose, cytoskeleton regulation, cold and drought regulation, and B-vitamin synthesis. These results imply that *Arsenophonus* is producing B-vitamins, which might be deficient in locust. In a second set of experiments, I used qPCR to test whether symbiont function across host plants might be mediated by bacterial titer. I measured relative *Arsenophonus* abundance across plants, and found *Arsenophonus* titer was variable, but generally greater on locust than fava. In summary, my results suggest that *Arsenophonus* synthesis of B-vitamins should be further investigated and may be mediated by bacterial titer.

KEYWORDS: *Aphis craccivora*, *Arsenophonus*, B-vitamins, facultative symbionts, host plant facilitation, relative symbiont abundance

Thorsten Hansen

01/11/2018

IDENTIFYING MECHANISMS OF HOST PLANT SPECIALIZATION IN *APHIS CRACCIVORA* AND ITS BACTERIAL SYMBIONTS

By

Thorsten Hansen

Dr. Jennifer A. White

Director of Thesis

Dr. Charles W. Fox

Director of Graduate Studies

01/11/2018

Date

ACKNOWLEDGEMENTS

I would like to thank my advisor and mentor, Jen White, for guiding me and helping me develop my projects. My time within her lab has helped me grow professionally and feel confident in my development as a scientist. I would also like to thank my two committee members: Dr. Chuck Fox and Dr. Clare Rittschof for technical advice.

I want to thank my lab members: Dr. Paul Lenhart, Laura Rosenwald, Sarah Ferrell, Allison Lindsey, and Kyla O'Hearn. Dr. Lenhart has been an additional mentor giving me good advice throughout the course of my master's degree. Our lab manager Laura Rosenwald and our small army of undergrads, Sarah Ferrell, Allison Lindsey, and Kyla O'Hearn, have been a joy to work with a helped me frequently with experimental setup.

I also want to thank our department chair Dr. Reddy Palli for graciously allowing me access to equipment in his lab. I also want to thank several members of the Palli lab that I received technical guidance from: Smitha George, Megha Kalsi, and June-Sun Yoon.

Finally, I want to thank my girlfriend, Kimberly Baugh, who has given me unconditional love, support, and encouragement during my master's degree. Along with friends who have helped me keep sane when I was feeling stressed: Josiah Ritchey, Boris Sauterey, Kaitlyn Staton and Collin Ray.

TABLE OF CONTENTS

Acknowledgements iii

List of Tables v

List of Figures vii

Chapter 1: Background 1

Chapter 2: Differential expression of the *Aphis craccivora-Arsenophonus-Buchnera* transcriptome in response to two different host plants 7

Introduction 7

Methods 11

Results 14

Discussion 17

Chapter 3: Investigation of relative *Arsenophonus* abundance changes across two host plants 33

Introduction 33

Methods 35

Results 39

Discussion 40

Chapter 4: Conclusions and future directions 54

References 58

Vita 65

LIST OF TABLES

Table 2.1: Aphid clonal lines used for the RNAseq experiment 22

Table 2.2: Samples used to generate transcriptome including 12 experimental samples and 2 additional samples 23

Table 2.3: Transcriptome contig statistics for *A. craccivora-Buchnera-Arsenophonus* joint transcriptome 24

Table 2.4: Mapping statistics of 12 experimental sample reads to *A. craccivora-Buchnera-Arsenophonus* joint transcriptome 25

Table 2.5: Differentially expressed genes between *Arsenophonus*-infected aphids feeding on locust versus fava 26

Table 2.6: Differentially expressed genes between uninfected cowpea aphids on locust versus fava 27

Table 2.7: Summary table for transcriptional response comparisons 28

Table 2.8: Differentially expressed genes between *Arsenophonus*-infected and uninfected cowpea aphids on locust 29

Table 2.9: Differentially expressed genes between *Arsenophonus*-infected and uninfected cowpea aphids on fava 31

Table 3.1: The initial collection dates and locations of the aphid clonal colonies 49

Table 3.2: Primers, PCR and qPCR cycling conditions, and references used for symbiont diagnostics, symbiont housekeeping and aphid housekeeping genes 50

Table 3.3: The collection location, number of samples, and host plant of origin for the field collected adult cowpea aphids 51

Table 3.4: ANOVA table Experiment 1 52

Table 3.5: Statistical table for GLIMMIX output in Experiment 2 53

LIST OF FIGURES

Figure 3.1: Relative *Arsenophonus* abundance in experiment 1.....46

Figure 3.2: Relative *Arsenophonus* abundance in experiment 2.....47

Figure 3.3: Relative *Arsenophonus* abundance in aphids collected from different host plants in the field for experiment 3 48

Chapter 1: Background

The last century has yielded a renaissance of understanding for insect-plant interactions. It has been recognized that both plant amino acids (Kennedy, 1965) and secondary metabolites (Fraenkel, 1959) have been driving forces in the coevolution of insect herbivores with plants (Ehrlich and Raven, 1964). In accordance with the Red Queen hypothesis (Van Valen, 1977), for millions of years insects and plants have been locked in evolutionary battle, causing plants to develop physical armaments and complex secondary chemicals to deter herbivores, and insects to develop countermeasures against plant defenses. Insect measures include behavioral mechanisms (Berenbaum, 1983; Dussourd and Eisner, 1987), sequestration (Conner et al., 2000; Eisner et al., 1974), detoxification (Brattsten et al., 1977; Krieger et al., 1971; Li et al., 2002), inhibition (Musser et al., 2002), and microbial symbiosis (Douglas, 2015; Hammer and Bowers, 2015).

From the perspective of the insect host, microbial symbionts can be classified as obligate or facultative. Obligate symbionts are essential to their hosts' survival, due to their nutritional provisioning abilities, being maternally transmitted to each new generation of offspring (Hansen and Moran, 2014). In contrast, facultative symbionts are more dynamic, occasionally being horizontally transmitted to new hosts, along with maternal transmission, and can provide novel phenotypes such as: heat tolerance, parasitoid defense, fungal defense, reproductive manipulation and host plant usage (Ferrari and Vavre, 2011; Hansen and Moran, 2014; Oliver et al., 2010; Oliver and Martinez, 2014). These novel phenotypes, in turn, potentially give the host access to new ecological niches (Ferrari and Vavre, 2011; Hansen and Moran, 2014; Oliver et al., 2010;

Oliver and Martinez, 2014). Due to their role in nutritional supplementation, obligate symbionts have a long coevolutionary history with sap feeding insects, being associated with the origin of major clades at least 270 million years ago (Baumann, 2005; Downie and Gullan, 2005; Moran, 2007; Moran et al., 2005; Munson et al., 1991; Thao, Clark, et al., 2000; Thao, Moran, et al., 2000). Obligate symbionts have opened access for their hosts to new niches through the ability to use nutritionally unbalanced diets (Bennett and Moran, 2015; McLean et al., 2016), but their intensive level of coadaptation has made hosts dependent on their symbionts, and they cannot survive on their imbalanced diets without their microbes (Fisher et al., 2017).

This intense coadaptation also presents risks to the obligate symbiont. For the symbiont, the stable internal environment relaxes selection pressures on genes that code for products that would be critical for survival outside the host. Consequently, deleterious mutations accumulate in the symbiont genome causing pseudogenization and eventually purging of genes (Hansen and Moran, 2014). The gene loss leads to a reduced genome, where only minimal housekeeping gene are kept, along with genes that code for the nutrient provisioning that the host requires. Eventually, even the nutrient provision genes can be lost. Over evolutionary time, if the diet of the insect host provides a sufficient supply of the same amino acids that the obligate also produces, it can cause reduced selection pressure on the symbiont to produce these amino acids and lead to the loss of these nutritional supplementation genes; this leaves the insect host unable to rely on their obligate symbiont to produce these amino acids, forcing increased specialization onto host plants that have the required nutrients, narrowing the spectrum of potential hosts the

insect can feed on and in some cases leading to possible extinction (Bennett and Moran, 2015; McLean et al., 2016).

Like obligate symbionts, facultative symbionts can also affect their hosts' ability to use host plants. Facultative symbionts have been correlated with, and in some cases shown to promote, a shift in host plant utilization (Ferrari et al., 2012; Frago et al. 2012; Henry et al., 2013; Leonardo and Muiru, 2003; Tsuchida et al, 2011; Wagner et al., 2015). Facultative symbionts may influence insect-plant interactions through a few processes like secondary compound metabolism or nutritional supplementation. Research suggests facultative symbionts may facilitate metabolic detoxification of secondary chemicals (Ghanim and Kontsedalov, 2009; Pan et al., 2013) and/or nutritional supplementation of cofactors, vitamins, and nucleotides (Lamelas et al., 2011) may be involved. A facultative symbiont known to influence many insect-plant interactions, and possibly facilitate detoxification and nutritional supplementation, is *Arsenophonus*.

Many strains of *Arsenophonus* are arthropod-associated and show a diversity of different biological interactions with their various hosts. *Arsenophonus* is a reasonably prevalent symbiont, infecting of 4-7% of arthropods (Duron et al., 2008). The aphid genus *Aphis* is particularly prone to *Arsenophonus* infection, with 31% of tested species infected (Jousselin et al., 2013). In the parasitoid wasp, *Nasonia vitripennis*, *Arsenophonus nasoniae* manipulates host reproduction, and genomic sequencing of this strain of *Arsenophonus* reveals genes for virulence and symbiosis (Darby et al., 2010; Wilkes et al., 2010). In obligate blood feeding arthropods, *Arsenophonus* has been found as an obligate or co-obligate symbiont that supplies B vitamins to its hosts (Dale et al., 2006; Nováková et al., 2015). As a facultative symbiont in the sap feeder *Nilaparvata*

lugens, recent genomic studies have shown that *Arsenophonus* has the same capacity for B vitamin production (Xue et al., 2014). It is possible that the nutritional supplementation may be a key element in the facultative status of *Arsenophonus* across arthropods and in some cases, may facilitate use of host plants that might otherwise be challenging for the herbivore host alone.

In the cowpea aphid, *Aphis craccivora*, *Arsenophonus* affects aphid performance across different host plants. Aphids infected with *Arsenophonus* have greater fitness when feeding on locust, *Robinia pseudoacacia*, than uninfected aphids, but when feeding on *Vicia faba*, fava, and *Medicago sativa*, alfalfa, infected aphids perform more poorly than uninfected aphids (Wagner et al., 2015). The mechanism behind increased aphid performance on locust when infected with *Arsenophonus* is currently unknown. I hypothesize that *Arsenophonus* may either be providing nutritional supplementation or detoxification of locust-associated secondary plant chemicals.

The first possibility is that *Arsenophonus* could be helping the cowpea aphid cope with nutritional deficiency on locust. Host plant quality had been shown to affect aphid morph production, development rate, and fecundity (Nevo and Coll, 2001). Previous research has shown that aphid performance can vary greatly depending on differences in amino acid composition among different plant species (Douglas, 1993; Sandstrom and Pettersson, 1994). Differences in amino acid profiles can have a potentially profound effect on aphids. If locust phloem is nutritionally deficient, *Arsenophonus* may provide a way to produce metabolic components like amino acids, vitamins, or cofactors. The production of B vitamins is a likely possibility as *Arsenophonus* in the Hippoboscidae fly

(Nováková et al., 2015) and *N. lugens* (Xue et al., 2014) have the capabilities to produce B vitamins for their hosts.

Alternatively, *Arsenophonus* may help the aphid overcome secondary plant defense chemicals. Secondary chemicals in plants, which were once considered waste products, are now widely understood to be defensive compounds (Becerra et al., 2009; Ehrlich and Raven, 1964; Richards et al., 2015). Alfalfa, fava and locust are all a part of the plant family Fabaceae (Doyle and Luckow, 2003). There are a variety of secondary metabolites present within Fabaceae, and there is differential secondary chemistry among different groups within the family (Wink, 2013). Microbial symbionts across many different insect groups have been found to detoxify potentially toxic chemicals (Hammer and Bowers, 2015). If *Arsenophonus* can produce enzymes linked to detoxification, it could explain how *Arsenophonus* facilitates the use of locust, particularly if locust has some secondary chemical compounds that fava or alfalfa do not.

The main of objective of my thesis is to understand the role *Arsenophonus* plays in improving cowpea aphid performance on locust. I addressed this topic through two different methods: RNAseq analysis and qPCR. In Chapter 2, I used RNAseq analysis of the joint transcriptome of *Aphis craccivora*, its facultative symbiont *Arsenophonus*, and its obligate symbiont *Buchnera aphidicola.* In this exploratory process, I looked for differential gene expression across host plants, to generate mechanistic hypotheses for *Arsenophonus* function. Following my transcriptomic results, in Chapter 3 I used qPCR to assess how the *Arsenophonus* bacterial population size (titer) responds to the selective environments of locust and fava. I hypothesized that bacterial titer would be higher on locust than fava because aphids experience increased performance when infected and

feeding on locust compared to those infected and feeding on fava. More generally, the results from my thesis may give insight into the role facultative symbionts may play in the evolution of dietary breadth and host plant usage in herbivores.

Chapter 2: Differential expression of the *Aphis craccivora-Arsenophonus-Buchnera* transcriptome in response to two different host plants

Introduction

Insect herbivores have been closely tied to plants for millions of years. In response to the herbivore threat, plants have developed physiological, molecular, and chemical defenses to deter the feeding of insects, evolving into many divergent taxa in the process (Agrawal and Weber, 2015; Futuyma and Agrawal, 2009). Selection pressures, biotic and abiotic, have prompted insect herbivore evolution of behavioral avoidance mechanisms (Berenbaum, 1983; Dussourd and Eisner, 1987), sequestration (Conner et al., 2000; Eisner et al., 1974), detoxification (Brattsten et al., 1977; Krieger et al., 1971; Li et al., 2002), inhibition (Musser et al., 2002), and microbial symbiosis (Douglas, 2015; Hansen and Moran, 2014). Particularly critical to many insect clades, bacterial symbiosis has influenced host plant utilization and the evolution of many insect herbivores (Dussourd and Eisner, 1987; Schoonhoven et al., 2005).

Many insect herbivores are associated with symbionts. From the perspective of the insect hosts, symbionts can be split into two categories; primary obligate symbionts, which are essential to the function of their hosts and provide nutritional supplementation, and facultative symbionts, which are non-essential, but can provide a range of functions (Douglas, 2015; Hansen and Moran, 2014). Obligate symbionts helped their hosts to initially colonize niches where key nutrients are missing, such as plant phloem, which in turn facilitated the adaptive radiation of their insect hosts into diverse clades (Bennett and Moran, 2015; McLean et al., 2016). Obligate symbionts are non-free living, with small population sizes that experience regular bottlenecks, and thus have no opportunity for

genetic recombination (Hansen and Moran, 2014). The internal cellular environment of their hosts is benign, meaning that selection pressures on any traits not necessary for symbiosis are reduced or absent, thus allowing the accumulation, and subsequent fixation, of deleterious mutations and erosion of the genome, losing elements that would be necessary for free-living microbes (Hansen and Moran, 2014). Even genes important to nutritional supplementation can eventually be lost, restricting ecological host range and possibly even leading to extinction of both host and symbiont (Bennett and Moran, 2015; McLean et al., 2016). These processes lead to obligate symbionts having a few characteristic traits, including A+T bias in their DNA makeup, reduced genome size, and lack of mobile DNA elements (Hansen and Moran, 2014). However, many host lineages with failing obligate symbionts have acquired another microbial partner to augment or replace the original. There are many examples of insect clades with co-symbionts or replacement symbionts (Koga et al., 2013; McCutcheon et al., 2009; McCutcheon and von Dohlen, 2011; Urban and Cryan, 2012). But where do these replacement symbionts come from? One possibility would be facultative symbionts that co-occur in the same host.

Facultative symbionts are non-essential to their hosts, but some do hold the capacity to provide nutritional functions like obligate symbionts. In the right circumstances, if an obligate symbiont loses the ability synthesize nutritional components and the facultative symbiont can produce the nutritional components lost in the obligate, the insect host may be able to continue with the facultative compensating for the obligate. In the aphid species *Cinara cedri*, functional annotation of the genomes of both the obligate symbiont, *Buchnera*, and the facultative symbiont, *Serratia*, indicate that

Serratia may be transitioning from a facultative symbiont to a co-obligate. *Buchnera* has lost the ability to synthesis some nutritional components, but *Serratia* has retained these missing nutritional synthesis pathways (Lamelas et al., 2011). The transition of facultative symbiont to obligate does not seem to be isolated to *Serratia* only, as the same process may be happening in *Hamiltonella*, in the aphid genus *Uroleucon* (Degnan and Moran, 2008) and whitefly *Bemisia tabaci* (Sagot et al., 2015), along with facultative symbionts in weevils (Toju et al., 2013). However, most facultative symbionts are likely not transitioning to a co-obligate or replacement status, as the loss of obligate symbiont function is a rare event over evolutionary time.

Unlike obligate symbionts, many facultative symbionts provide their hosts with novel phenotypes that can give access to a wider array of ecological niches (Douglas, 2015; Hansen and Moran, 2014; Oliver et al., 2010; Su et al., 2013). In aphids, there is evidence that colonization of novel host plants has been facilitated by the presence of facultative symbionts (Frago et al., 2012; Henry et al., 2013; Wagner et al., 2015). Aphids with similar ecologies across different regions have been found harboring the same facultative symbiont (Henry et al., 2013), along with a pattern of higher infection frequency in host-alternating or polyphagous aphid species suggesting a specific metabolic function (McLean et al., 2016). Akin to the co-obligate or replacement symbioses, it has been suggested that facultative symbionts may be providing nutritional supplementation (Oliver et al., 2010). Alternatively, it is possible that facultative symbionts may provide plant secondary chemical detoxification abilities, similar to many gut symbionts (Hammer and Bowers, 2015). Understanding the molecular underpinnings of facultative symbionts will help define the metabolic function(s) they may contribute in

facilitation of novel host plants. A molecular method that can be used to discover the metabolic functions of facultative symbionts, transcriptomics, can act as a hypothesis-generating procedure, allowing annotation and identification of differentially expressed symbiont genes between different host plants, which can provide clues to symbiont function.

Aphis craccivora, the cowpea aphid, is a global polyphagous pest that can be infected with a variety of facultative symbionts (Brady and White, 2013). In Eastern North America, the cowpea aphid is mainly found on two host plants, alfalfa and locust, with aphids on locust almost always associated with the infection by the facultative symbiont *Arsenophonus* (Wagner et al., 2015). *Arsenophonus* infected cowpea aphids have increased performance on locust when compared to uninfected aphids. In contrast, *Arsenophonus* infected cowpea aphids perform worse on alfalfa and fava when compared to their uninfected counterparts (Wagner et al., 2015).

Knowing that *Arsenophonus* infected aphids perform better on locust, we can manipulate the host plants the aphids are feeding on and explore the effects on the transcriptome of the aphid and its facultative symbiont. I used next generation sequencing technologies to explore the molecular mechanisms behind this aphid-symbiont interaction. Using RNA-seq, I created an *Aphis craccivora-Arsenophonus-Buchnera* reference transcriptome. Then, using the reference transcriptome, I compared expression of *Arsenophonus* infected and uninfected cowpea aphids on both locust and fava to identify molecular mechanisms associated with *Arsenophonus* infection and locust utilization, with a specific focus on bacterial biological pathways linked to host plant facilitation.

Methods

My experiment aimed to generate plausible hypotheses to explain how *Arsenophonus* facilitates the use of locust as a host plant by the cowpea aphid, *Aphis craccivora*. My experiment consisted of a 2 × 2 factorial design: aphid infection status (Infected with *Arsenophonus* vs Uninfected) by host plant (*Robinia pseudoacacia*, locust vs *Vicia faba*, fava). I used three genetically distinct aphid clones in the experiment, which came from colonies that had been maintained in the lab on fava multiple years prior to the experiment (Table 2.1). Each aphid colony had originally been initiated with a single aphid individual, which parthenogenetically reproduced to produce genetically identical aphid populations. Two of the clones, LE+ and LW+, were naturally infected with *Arsenophonus*. Sub-colonies of each of these clones were subsequently cured of infection via antibiotic diet (LE- and LW-; Wagner et al., 2015). The third clone, AL-, was naturally uninfected with *Arsenophonus*, and was experimentally transfected with *Arsenophonus* via hemolymph microinjection (AL+; Wagner et al., 2015). Thus, for all three clones we had paired, genetically identical colonies, differing only in *Arsenophonus* infection status. For the experiment, each of these six colonies was subdivided and reared on both fava and locust, for a total of 12 experimental units. Each colony was maintained in a 3.78-liter plastic jar with mesh panels for ventilation fitted over either fava or locust seedlings growing in 10cm pots. Aphids fed on the host plants for approximately 2 weeks under ambient laboratory temperature (22° $\pm$ 4°C) and supplemental full spectrum lights (16 L:8 D).

For each of the 12 colonies, nymphal aphids were collected for RNA extraction. All aphid nymphs collected from fava were 4th instar, but limited supply of aphids on

locust required a broader collection range, corresponding to all 1st-4th instar nymphs. For each sample approximately 50 aphids were collected (corresponding to 3.00-18.0 mg fresh weight aphid/sample). RNA was isolated from each sample by homogenizing specimens in Trizol reagent, and purified using QIAGEN RNeasy Plus Universal RNA extraction kit (Hilden, Germany) according to manufacturer protocols. RNA concentration was assessed using a Thermo Fisher Scientific nanodrop 2000 spectrophotometer (Waltham, MA, USA).

RNA sample library preparation and sequencing were performed at the Beckman Coulter Genomics sequencing facility (Danvers, MA, USA). In total, 14 *A. craccivora* RNA samples were sequenced, two of which were not part of the experiment described above (Table 2.2). Each library was prepared using Illumina TruSeq Stranded Total RNA Library Prep Kit, which includes Ribo-Zero chemistry to reduce ribosomal representation in the library. Libraries were multiplexed into a single sample that was sequenced across 2 lanes on an Illumina HiSeq 2500 platform with 2 x 100bp read lengths. Samples were demultiplexed and adapters were removed from the reads by the sequencing facility. In total, 463,982,406 reads were produced across the 14 samples, all of which I used to construct the reference transcriptome, which I subsequently used to evaluate differential expression of the 12 experimental samples.

To construct the reference transcriptome, I pre-processed the samples using the programs FastQC (Andrews, 2010) to check quality and Trimmomatic (Bolger et al., 2014) to remove low quality sequences. Standard Trimmomatic settings were used, an internally calculated quality score that accounts for sequence quality was used to remove leading low-quality nucleotide bases below 3, trailing low-quality nucleotide bases below

3, entire sequences where a 4-base sliding window indicated average quality per a base dropped below 15, and sequences that were shorter than 36 bases. Next, I created a *de novo* assembly of the Illumina reads using Trinity software on the University of Kentucky high performance computing cluster. Reads were normalized through *in silico* read normalization processes, which produced a combined transcriptome for *Aphis craccivora*, its obligate symbiont *Buchnera*, and the facultative symbiont *Arsenophonus*. A *de novo* assembly was needed as all three organisms' have congeneric genomes available, but not conspecific. Transcriptome mapping and coverage statistics were generated in CLC Genomics workbench 10.1.1 (CLC Bio, Aarhus, Denmark, http://www.clcbio.com/) using default read alignment options.

Within CLC genomics workbench, I used the RNA-Seq Analysis with default read alignment options to map the reads from each individual sample set back to the Trinity reference to produce transcript expression tracks. Differential expression analysis was also run within CLC genomics workbench. For expression analysis options, paired reads were counted as two (two reads in an intact pair are each counted as one mapped read and mapped members of broken pairs each get a count of one). The differential expression was used to compare aphid infection status (Infected with *Arsenophonus* vs Uninfected) by host plant (locust vs fava). I then had four differential expression contrasts: 1) infected with *Arsenophonus* on locust vs infected with *Arsenophonus* on fava, 2) uninfected on locust vs uninfected on fava, 3) infected with *Arsenophonus* on locust vs uninfected on locust and 4) infected with *Arsenophonus* on fava vs uninfected on fava. Comparison 1 was used to identify candidate genes associated with improved aphid performance on locust in the presence of *Arsenophonus*. Comparison 2 identified

transcripts that change expression across hostplants even in the absence of *Arsenophonus*. By comparing this group of transcripts to the first group, I could then identify candidates that were uniquely associated with change of hostplant use in the presence of *Arsenophonus*. Comparisons 3 and 4 were used to look for transcripts associated with symbiont infection on each host plant. Aphid clone was not incorporated into any of these contrasts. All comparisons were made using expression values measured in Reads Per Kilobase of transcript per Million mapped reads (RPKM) that were transformed (log 2) and normalized according to quantile values. P-values for differential gene expression were adjusted with a False Discovery Rate (FDR) correction; values at $\alpha = 0.05$ following correction were considered significant.

To identify differentially expressed genes, I blasted significantly differentially expressed transcripts using the BLAST2GO pipeline within CLC genomics workbench. In "CloudBlast," I used blastx against the NCBI non-redundant protein sequences database. Separately, I also used both blastx and blastn to verify the origin of my transcripts by blasting them against Aphididae, *Buchnera*, and *Arsenophonus* specific NCBI non-redundant protein sequences and nucleotide collection databases respectively.

Results

I generated a *de novo* assembled *Aphis craccivora-Arsenophonus-Buchnera* transcriptome from Illumina paired reads of 14 different clonal samples. The transcriptome nucleotide distribution was AT biased at 68.2%. In total, 326,591 contigs were generated with lengths 200bp and longer (Table 2.3). The total length of assembled contigs was 335.819 Mb and the N50 was 2023bp (Table 2.3). On average, 99% of reads mapped back to the assembled transcriptome (Table 2.4)

The differential expression comparisons looked at different combinations of aphid infection status and host plant, with three aphid clones as replicates for each treatment. The four comparisons were: 1) infected with *Arsenophonus* on locust vs infected with *Arsenophonus* on fava, 2) uninfected on locust vs uninfected on fava, 3) infected with *Arsenophonus* on locust vs uninfected on locust and 4) infected with *Arsenophonus* on fava vs uninfected on fava.

The first comparison of interest was of *Arsenophonus* infected aphids on locust versus fava. Only 5 transcripts were significantly differentially expressed, after correcting for the false discovery rate (Table 2.5). All five were upregulated on fava relative to locust, and none of them originated from *Arsenophonus.* Four of the transcripts were all isoforms of the same gene, corresponding to an uncharacterized protein from the *Acyrthosiphum pisum* genome (LOC103310381; Table 2.5). The fifth DE transcript was the outer membrane porin, OmpA-like protein, from the obligate symbiont *Buchnera*.

The second comparison was of uninfected cowpea aphid on locust versus uninfected cowpea aphids on fava. Overall, there were 44 DE genes in the absence of *Arsenophonus* (Table 2.6). When comparing these genes to those identified in the first comparison, 3 of the 5 genes listed for the first contrast are also found in the second contrast (Table 2.7). This indicates that most of the differential expression from the primary comparison is not specifically associated with *Arsenophonus* infection. Of the 44 DE genes in the uninfected aphid contrast, three originated from *Buchnera* and the rest from the aphid (Table 2.6). Most of the DE genes were upregulated on fava, but 12 of them were upregulated on locust. Of the genes upregulated on locust, 10 were associated with *Aphis craccivora* and 2 were associated with *Buchnera* (Table 2.6). The 10 DE

aphid genes correspond to: three isoforms of maltase A3-like which is involved in carbohydrate metabolism, tropomyosin isoform X10, which affects cytoskeleton regulation, cold and drought-regulated CORA-like protein, an *A pisum* uncharacterized protein (LOC107171435), and 4 genes that had no match in blastx. The 2 DE *Buchnera* genes are serine acetyltransferase, which participates in the pathways for cysteine and sulfur metabolism, and a putative protein (ECO:0000313).

When looking at the third comparison, of *Arsenophonus* infected aphids on locust vs uninfected aphids on locust, there were 75 DE genes (Table 2.8). However, of these 75, 71 were from *Arsenophonus*, which obviously correspond to the *Arsenophonus* infected treatment (Table 2.8). Of these *Arsenophonus* genes, 2 were connected to B-vitamin metabolism, beta-ketoacyl-[acyl-carrier-] synthase I and serine acetyltransferase (Strauss et al., 2001; Toomey and Wakil, 1966). The 4 non-*Arsenophonus* genes were all from *Buchnera*, with 3 being upregulated in the *Arsenophonus* infected treatment and 1 being upregulated in the uninfected treatment (Table 2.8). The 3 *Buchnera* genes upregulated in infected aphids were:1) a *Buchnera* gene that has homology with the hypothetical protein ALO39_101110 from the bacterium *Pseudomonas syringae*, 2) ORF16-lacZ fusion which metabolizes lactose, and 3) an operon that codes for shikimate dehydrogenase (aroE) and cysteinyl-tRNA synthetase (cysS), which are involved with the production of aromatic amino acids and cysteine metabolism respectively. The 1 DE *Buchnera* gene upregulated in uninfected aphids, which is also upregulated on fava in comparison 1, is OmpA-like protein.

The fourth and final comparison was of *Arsenophonus* infected aphids on fava vs uninfected aphids on fava. The *Arsenophonus* infected to uninfected comparison on fava

had 56 DE genes (Table 2.9). As expected, many of these genes (50), were of *Arsenophonus* origin, corresponding to the presence or absence of the facultative symbiont (Table 2.9). When contrasting comparison 3 and comparison 4, 49 of the *Arsenophonus* genes were the same (Table 2.7), and all represented bacterial maintenance genes (Table 2.8; Table 2.9)**.** Of the DE genes unique to comparison 3, only the 2 B-vitamin synthesis genes, beta-ketoacyl-[acyl-carrier-] synthase I and serine acetyltransferase (Strauss et al., 2001; Toomey and Wakil, 1966), were not linked to bacterial maintenance. For the 6 non-*Arsenophonus* genes in comparison 4, 4 are from *Buchnera* and 2 are of aphid origin (Table 2.9). From *Buchnera*, 3 genes were upregulated in the *Arsenophonus* infected treatment and 1 was upregulated in the uninfected treatment (Table 2.9). These 3 genes were hypothetical protein ALO39_101110, ORF16-lacZ fusion, and the operon containing aroE along with cysS (all 3 of which were also found in comparison 3). The 1 DE *Buchnera* gene that was upregulated in the uninfected treatment had no matches in blastx. Both DE aphid genes were upregulated in the *Arsenophonus* treatment (Table 2.9). Of the 2 upregulated aphid genes, one codes for glutathione S-transferase isoform D-like and the other has no match in blastx.

Discussion

I expected a significant transcriptomic response when I compared gene expression of *Arsenophonus*-infected aphids feeding on locust (an environment where the symbiont has been shown to have beneficial fitness effects) versus *Arsenophonus*-infected aphids feeding on fava. However, the overall differential gene expression was of only 5 genes in total, all of which were upregulated on fava. None of these genes were from

Arsenophonus and instead came from the aphid host and the primary symbiont, *Buchnera*. The aphid genes are all related to an uncharacterized protein from *A. pisum* (LOC103310381), so not much functional information can be learned from them. The single *Buchnera* gene, with 41-fold upregulation on fava, was the porin membrane protein OmpA.

OmpA, as found in *Escherichia coli*, provides structural integrity to the outer cellular membrane, can serve as a receptor to phages, and permits the diffusion of small solutes, which could possibly include amino acids (Manning et al., 1977; Sugawara and Nikaido, 1992; Wang, 2002). Considering that both *Buchnera* and *Escherichia coli* are gammaproteobacteria (Nováková et al., 2013), the proteins probably function similarly in *Buchnera*, albeit with a few differences. In *Buchnera*, the protein is used for host recognition purposes (Tamas et al., 2001), has been shown to have host interaction in hampering symbiont cell division (Login and Heddi, 2013), and is involved in amino acid metabolism (Sabater-Muñoz et al., 2017). For both comparison 1 and 3, OmpA was down regulated in conjunction with *Arsenophonus* on locust. Speculatively, *Arsenophonus* may obviate some part of the *Buchnera* OmpA function when aphids feed on locust.

To understand if the DE differences seen in our *Arsenophonus* infected treatments were a response to infection or host plant, we also compared gene expression of uninfected aphids on locust and fava as a control comparison. The same aphid isoforms of *A. pisum* protein LOC103310381 were upregulated on fava, even in the absence of *Arsenophonus*, indicating their expression is a general response to host plant changes, and not directly related to the symbiont. In the uninfected aphids, we detected several other genes that were differentially expressed as well, including three isoforms of maltase A3-

like, tropomyosin isoform X10, which affects cytoskeleton regulation, and cold and drought regulated CORA-like protein, all of which were upregulated on locust. The presence of maltase A3-like leads me to speculate that locust may have a higher starch content than fava, which maltase A3-like metabolizes into glucose (Stafford-Banks et al., 2014). Tropomyosin is a thin filament associated protein involved in muscle contraction (Lee et al., 2016; Meng et al., 2014). Within lepidopteran herbivores, changes in host plant nutrition can be linked to expression of tropomyosin isoform involved in flight muscle metabolism (Portman, 2013). This may be flight related for aphids as well, as not all aphids become winged alates, and the developmental trigger for an aphid to become an alate is often related to nutritional stress (Nevo and Coll, 2001). It is possible that there were more nymphs destined to become alates on locust versus fava, because we were not able to control for this factor during sample collection. Serine acetyltransferase in *Buchnera* was also upregulated on locust. This protein catalyzes serine, which is subsequently used in the pathway of cysteine amino acid synthesis (Kredich and Tomkins, 1966; Shigenobu et al., 2000) and can also be used in production of the B-vitamin, Pantothenate (Strauss et al., 2001). If the phloem from locust had high serine levels, which has been shown under non-drought conditions (Liu et al., 2013), *Buchnera* could be overexpressing serine acetyltransferase. The transcriptional response of aphids lacking *Arsenophonus* shows that host plant quality may have some effect on the aphid host and its primary symbiont. However, it provides no clues to the function of *Arsenophonus*.

The final set of contrasts compared the transcriptome of aphids with and without *Arsenophonus* on each host plant. It was expected that the majority of DE of genes in

these dual comparisons would likely correspond to *Arsenophonus* itself, which was present in the infected treatment but absent in the uninfected treatment. This expectation was validated with 75 of the DE genes deriving from *Arsenophonus*, and only 4 DE genes originating from *Buchnera.* Forty-nine of the DE *Arsenophonus* genes were detected from both aphids on locust and on fava. The bulk of these genes represent bacterial maintenance genes. However, shared amongst *Arsenophonus* infected aphids on both plants, are two standout *Arsenophonus* genes, lipoyl synthase and 3,4-dihydroxy-2-butanone-4-phosphate. Both genes produce intermediate products in B-vitamin pathways and are also found in the *Arsenophonus* of the brown planthopper, *Nilaparvata lugens* (Xue et al., 2014). The DE for only the *Arsenophonus* infected versus uninfected locust comparison showed yet two more *Arsenophonus* genes connected to B-vitamin metabolism, beta-ketoacyl-[acyl-carrier-] synthase I and serine acetyltransferase (Strauss et al., 2001; Toomey and Wakil, 1966). Interestingly, in the blood feeding family of flies, Hippoboscidae, *Arsenophonus* exists as an obligate symbiont that most likely produces B-vitamins for its hosts (Nováková et al., 2015). For one Hippoboscidae species, phylogenetic analysis shows that its obligate *Arsenophonus* symbiont is closely related to facultative *Arsenophonus* symbionts within sap feeders (Dale et al., 2006). As suggested with *Arsenophonus* in the Hippoboscidae fly (Nováková et al., 2015) and *N. lugens* (Xue et al., 2014), *Arsenophonus* might be providing the *Aphis craccivora* host B-vitamins. This is just speculative and needs experimental manipulation to test this hypothetical functional relationship.

From a transcriptomics perspective, there are a few experimental shortcomings that could be improved in future efforts. First, the age of aphid nymphs could be more

strictly controlled. Use of a broader range of aphid instars on locust was unavoidable in the present experiment, due to a limited supply of aphids. The life stage used for an expression study can make a difference in the DE seen (Brisson et al., 2007; Brisson et al., 2010) and using several instars may have increased the variability in our results. Additionally, it is possible we did not have enough read depth, or replicate clones to reliably detect differential expression. RNAseq experiments in general have lots of variation, since there is only so much that can be controlled for in experiments and changing conditions can drastically affect expression. Transcriptomic studies function effectively as exploratory tools, allowing the production of hypothesis driven questions. A future step should be qPCR validation on the DE genes seen here, with an emphasis on those related to vitamin B synthesis. Further transcriptomic studies could increase read depth, allowing for identification of more DE genes. An expanded repertoire of DE genes might reveal genes related to host plant utilization that we did not find in the present experiment.

Table 2.1: Aphid clonal lines used for the RNAseq experiment.

Clonal Line	Host plant of origin	Infection status	Means of symbiont infection	Collection location	Date of collection
LE+	Locust	Ars +	Naturally infected	37º57'N 84º43'W	Sept 2011
LE-	Locust	Uninfected	Cured	37º57'N 84º43'W	Sept 2011
LW+	Locust	Ars +	Naturally infected	37º57'N 84º23'W	Sept 2011
LW-	Locust	Uninfected	Cured	37º57'N 84º23'W	Sept 2011
AL+	Alfalfa	Ars +	Transinfected	38º04'N 84º39'W	Aug 2011
AL-	Alfalfa	Uninfected	Naturally uninfected	38º04'N 84º39'W	Aug 2011

Table 2.2: Samples used to generate the transcriptome, including 12 experimental samples and 2 additional samples.

Clonal line (Sample)	Origin host plant	Feeding host plant	*Arsenophonus* presence	Symbiont manipulation
Experimental Samples				
LE+	Locust	Locust	Present	Naturally infected
LE+	Locust	Fava	Present	Naturally infected
LE-	Locust	Locust	Absent	Cured of *Arsenophonus*
LE-	Locust	Fava	Absent	Cured of *Arsenophonus*
LW+	Locust	Locust	Present	Naturally infected
LW+	Locust	Fava	Present	Naturally infected
LW-	Locust	Locust	Absent	Cured of *Arsenophonus*
LW-	Locust	Fava	Absent	Cured of *Arsenophonus*
AL+	Alfalfa	Locust	Present	Transinfected
AL+	Alfalfa	Fava	Present	Transinfected
AL-	Alfalfa	Locust	Absent	Naturally uninfected
AL-	Alfalfa	Fava	Absent	Naturally uninfected
Additional Samples				
AC1AB-	Alfalfa	Fava	Absent	Cured of *Hamiltonella*
SHP-	Locust	Fava	Absent	Naturally uninfected

Table 2.3: Transcriptome contig statistics for *A. craccivora-Buchnera-Arsenophonus* joint transcriptome.

Contig measurements	Length (bp)
N50	2,023
N20	4,328
Median length	495
Average	1028
Total	335,819,155

Table 2.4: Mapping statistics of 12 experimental sample reads to *A. craccivora-Buchnera-Arsenophonus* joint transcriptome.

Measurements	Count	Average length	Total bases (bp)	Percentage mapped (%)
Total reads	432,223,538	100	43,222,353,800	
Mapped	429,572,673	100	42,957,267,300	99.39
Not mapped	2,650,865	100	265,086,500	0.61
Contigs	326,591	1,028.26	335,819,155	
Reads in pairs	394,909,800	162.70	39,490,980,000	91.37
Broken paired reads	34,662,873	100	3,466,287,300	8.02

Table 2.5: Differentially expressed genes between *Arsenophonus*-infected aphids feeding on locust versus fava.

Transcript identifier	Gene length (bp)	Product description	Fold change	*P*-value	Upregulated on
Aphis craccivora					
Unique to PL vs PF*					
DN59507_c1_g1_i2	1676	*A. pisum* uncharacterized protein LOC103310381	-11.41	0.00013	Fava
Shared with NL vs NF*					
DN59507_c1_g1_i1	1304	*A. pisum* uncharacterized protein LOC103310381	-84.37	6.52E-09	Fava
DN59507_c1_g1_i3	2099	*A. pisum* uncharacterized protein LOC103310381	-10.42	1.63E-05	Fava
DN59507_c1_g1_i6	713	*A. pisum* uncharacterized protein LOC103310381	-8.68	0.00014	Fava
Buchnera					
Unique to PL vs PF*					
DN70885_c0_g1_i1	257	outer membrane A precursor	-41.11	0.00434	Fava

*Abbreviations: P stands for *Arsenophonus* infected, N stands for uninfected, L is locust host plant, and F is fava host plant.

Table 2.6: Differentially expressed genes between uninfected cowpea aphids on locust versus fava.

Transcript identifier	Gene length (bp)	Product description	Fold change	P-value	Upregulated on
Aphis craccivora					
Unique to NL vs NF*					
DN63093_c5_g3_i2	515	maltase A3-like	10.84	0.00122	Locust
DN58402_c4_g1_i1	649	maltase A3-like	10.06	0.0058	Locust
DN61795_c9_g1_i6	706	A. pisum uncharacterized protein LOC107171435	7.57	0.00304	Locust
DN63093_c2_g1_i1	501	maltase A3-like	5.62	4.05E-06	Locust
DN62929_c5_g6_i1	443	tropomyosin isoform X10	4.67	0.00299	Locust
DN61233_c0_g1_i3	324	cold and drought-regulated CORA-like	3.88	6.23E-05	Locust
DN61233_c2_g1_i2	1742	no match	3.24	1.33E-05	Locust
DN62618_c0_g1_i1	274	no match	2.6	0.00194	Locust
DN62618_c0_g6_i1	202	no match	2.55	0.0177	Locust
DN62618_c0_g8_i1	230	no match	2.14	0.00644	Locust
DN62100_c5_g1_i18	765	40S ribosomal S7	-70.45	0.00315	Fava
DN96811_c0_g1_i1	947	no match	-33.02	0.0239	Fava
DN55094_c0_g2_i1	434	no match	-30.08	0.00571	Fava
DN57421_c0_g1_i2	1456	aminoacylase-1-like isoform X1	-8.28	7.25E-16	Fava
DN65072_c2_g3_i5	2115	beta-retaining glycosyl hydrolase	-7.16	3.77E-06	Fava
DN58696_c2_g1_i3	242	transport Sec61 subunit gamma	-5.84	0.0428	Fava
DN65491_c6_g3_i11	405	unconventional myosin-XVIIIa isoform X4	-5.66	0.0408	Fava
DN58467_c2_g3_i1	2027	A. pisum uncharacterized protein LOC100573101	-5.42	0.00171	Fava
DN59433_c4_g1_i6	555	cuticle 7-like	-5.15	0.00083	Fava
DN59809_c5_g1_i2	1802	A. pisum uncharacterized protein LOC107173407	-4.78	0.00134	Fava
DN58778_c8_g2_i4	1454	TPA_inf: cathepsin B	-4.37	4.05E-06	Fava
DN58949_c8_g1_i2	1895	A. pisum uncharacterized protein LOC107165607	-4.09	0.00147	Fava
DN54183_c0_g1_i1	848	proteasome subunit alpha type-7-1	-3.89	0.00774	Fava
DN51361_c0_g1_i1	841	MD-2-related lipid-recognition -like	-3.86	0.00387	Fava
DN57987_c18_g1_i1	3193	sclerostin domain-containing 1	-3.56	9.97E-05	Fava
DN61102_c9_g4_i1	403	cuticular precursor	-3.5	0.0128	Fava
DN59067_c7_g2_i2	3198	agrin	-3.33	6.38E-06	Fava
DN60158_c1_g1_i2	1813	nuclear polyadenylated RNA-binding 3	-3.15	0.00083	Fava
DN62443_c7_g4_i1	1638	A. pisum uncharacterized protein LOC100167400	-3.14	0.0148	Fava
DN61710_c2_g1_i2	1083	soluble calcium-activated nucleotidase 1	-3.13	0.0486	Fava
DN57596_c0_g1_i1	229	no match	-3.05	0.0284	Fava
DN61610_c5_g1_i1	383	no match	-2.96	0.00299	Fava
DN58792_c1_g3_i1	3696	A. pisum uncharacterized protein LOC107168224	-2.36	0.0354	Fava
DN61437_c6_g2_i3	791	ATP synthase subunit mitochondrial	-2.19	2.03E-11	Fava
DN62950_c2_g1_i4	2617	A. pisum uncharacterized protein LOC107165083 isoform X2	-2.18	0.00644	Fava
DN64092_c4_g1_i1	1775	urease accessory -like	-2.17	0.00333	Fava
DN64007_c1_g1_i3	1599	muscle	-1.84	9.54E-05	Fava
DN58973_c5_g2_i2	785	mitochondrial-processing peptidase subunit beta	-1.72	0.00429	Fava
Shared with PL vs PF*					
DN59507_c1_g1_i1	1304	A. pisum uncharacterized protein LOC103310381	-64.49	5.19E-07	Fava
DN59507_c1_g1_i3	2099	A. pisum uncharacterized protein LOC103310381	-5.78	0.000111	Fava
DN59507_c1_g1_i6	713	A. pisum uncharacterized protein LOC103310381	-4.44	9.54E-05	Fava
Buchnera					
Unique to NL vs NF*					
DN45610_c0_g1_i1	233	serine acetyltransferase	28.49	0.0439	Locust
DN58726_c0_g2_i18	428	putative protein ECO:0000313	14.92	0.0355	Locust
DN57981_c0_g3_i1	327	hypothetical protein ESOG_04481, partial	-9.85	0.0354	Fava

*Abbreviations: P stands for *Arsenophonus* infected, N stands for uninfected, L is locust host plant, and F is fava host plant.

Table 2.7: Summary table for transcriptional response comparisons

Transcriptional response	*Ars*+ Locust vs Fava	*Ars*- Locust vs Fava	Overlap
Upregulated in Locust	0	12	0
Upregulated in Fava	5	32	3
Total	5	44	3
Transcription response	Locust *Ars*+ vs *Ars*-	Fava *Ars*+ vs *Ars*-	Overlap
Upregulated in *Ars*+	74	55	49
Upregulated in *Ars*-	1	1	0
Total	75	56	49

Table 2.8: Differentially expressed genes between *Arsenophonus*-infected and uninfected cowpea aphids on locust.

Transcript identifier	Gene length (bp)	Product description	Fold change	P-value	Upregulated on
Arsenophonus					
Unique to PL vs NL*					
DN63397_c5_g1_i1	240	Uncharacterised protein [Shigella sonnei]	635.43	1.29E-15	*Ars* +
DN57702_c0_g7_i1	875	serine acetyltransferase	582.55	1.45E-12	*Ars* +
DN56518_c2_g2_i1	343	cell wall-associated hydrolase	537.05	1.11E-13	*Ars* +
DN65502_c3_g4_i1	299	hypothetical protein BTY97_18675	531.81	3.60E-14	*Ars* +
DN44815_c0_g1_i1	243	Quinone oxidoreductase	449.99	5.92E-15	*Ars* +
DN20855_c0_g1_i1	294	Uncharacterised protein	256.27	2.57E-19	*Ars* +
DN58726_c0_g2_i2	374	secreted ECO:0000313	255.99	4.89E-12	*Ars* +
DN65067_c0_g2_i5	374	daphnid bacterial-ribosomal-RNA- possible HGT	232.95	0.00905	*Ars* +
DN65557_c5_g1_i1	1010	superoxide dismutase [Mn]	108.7	0.00441	*Ars* +
DN58081_c1_g3_i1	212	conserved hypothetical protein	97.32	1.15E-10	*Ars* +
DN56952_c0_g2_i1	7357	phenylalanine--tRNA ligase subunit beta	82.04	0.00303	*Ars* +
DN54507_c1_g1_i1	4723	beta-ketoacyl-[acyl-carrier-] synthase I	79.05	0.00235	*Ars* +
DN55867_c1_g2_i1	443	hypothetical protein XBKQ1_1410001	78.41	0.0254	*Ars* +
DN46776_c0_g1_i1	3953	molecular chaperone	74.86	0.0254	*Ars* +
DN51157_c0_g1_i1	4334	peptidase M23	69.15	0.00201	*Ars* +
DN48630_c0_g1_i1	2533	plasmid stabilization	67.94	0.00433	*Ars* +
DN65557_c3_g7_i1	1834	murein DD-endopeptidase	66.99	0.0116	*Ars* +
DN43079_c1_g2_i1	5772	transketolase	60.06	0.024	*Ars* +
DN40091_c0_g1_i1	5260	Cold-shock DEAD box A	53.45	0.0275	*Ars* +
DN97069_c0_g1_i1	648	30S ribosomal S20	40.85	0.0209	*Ars* +
DN67994_c0_g1_i1	443	50S ribosomal L25	39.56	0.0132	*Ars* +
DN58358_c0_g2_i1	665	SsrA-binding protein [*Arsenophonus* sp. ENCA]	26.47	1.79E-05	*Ars* +
DN61866_c2_g7_i1	214	daphnid bacterial-ribosomal-RNA- possible	22.65	0.0104	*Ars* +
DN64845_c6_g1_i1	235	hypothetical protein NTHI1209_00002 [*Haemophilus influenzae*]	7.33	0.0367	*Ars* +
DN56404_c3_g2_i1	498	conserved hypothetical protein	5.11	0.00314	*Ars* +
Shared with PF vs NF*					
DN58358_c0_g1_i6	941	SsrA-binding protein [*Arsenophonus* sp. ENCA]	698.29	4.41E-16	*Ars* +
DN39876_c1_g1_i1	2783	molecular chaperone	587.29	1.76E-09	*Ars* +
DN57009_c0_g1_i1	2209	16S rRNA (cytidine(1402)-2 -O)-methyltransferase	539.42	1.01E-34	*Ars* +
DN65557_c3_g6_i1	1212	transposase IS5 ssgr IS903 family	476.64	2.11E-27	*Ars* +
DN81462_c0_g1_i1	1386	porin	432.42	4.13E-17	*Ars* +
DN65502_c3_g3_i9	416	dehydration responsive	398.58	6.79E-11	*Ars* +
DN39355_c0_g1_i1	502	heat-shock	379.76	2.47E-05	*Ars* +
DN24426_c0_g1_i1	1562	tRNA dihydrouridine synthase	318.76	9.15E-14	*Ars* +
DN39355_c0_g2_i1	1045	heat-shock	307.95	7.10E-05	*Ars* +
DN81995_c0_g1_i1	451	murein lipo	266.69	4.29E-07	*Ars* +
DN30853_c1_g1_i1	1341	porin	216.01	8.67E-12	*Ars* +
DN29545_c0_g1_i1	3434	molecular chaperone	198.64	9.14E-06	*Ars* +
DN47916_c0_g1_i1	679	50S ribosomal L21	197.62	1.51E-07	*Ars* +
DN65557_c3_g2_i2	458	IS5 IS1182 family transposase	197.55	4.25E-12	*Ars* +
DN59119_c2_g1_i1	833	primosomal replication N	195.41	1.94E-09	*Ars* +
DN54389_c0_g2_i1	230	membrane [ECO:0000313]	190.76	4.17E-12	*Ars* +
DN89639_c0_g1_i1	1333	Z-ring-associated	181.79	5.88E-08	*Ars* +
DN59518_c3_g1_i2	1451	elongation factor Tu	179.06	7.92E-09	*Ars* +
DN59119_c5_g1_i1	543	50S ribosomal L9	172.42	1.58E-07	*Ars* +
DN65557_c3_g1_i1	548	hypothetical protein [*Arsenophonus nasoniae*]	171.8	1.93E-07	*Ars* +
DN58932_c3_g2_i1	1379	transcriptional regulator	166.72	6.85E-08	*Ars* +
DN65557_c3_g2_i1	621	IS5 IS1182 family transposase	163.64	3.96E-09	*Ars* +
DN65267_c1_g8_i8	14797	pre- translocase subunit	159.76	4.28E-08	*Ars* +
DN58538_c2_g5_i6	266	Quinone oxidoreductase [Escherichia coli O25b:H4]	155.65	4.12E-11	*Ars* +
DN65557_c20_g1_i1	657	50S ribosomal L28	155.15	3.34E-08	*Ars* +

Table 2.8, continued

Transcript identifier	Gene length (bp)	Product description	Fold change	P-value	Upregulated on
Arsenophonus					
Shared with PF vs NF*					
DN55229_c0_g2_i1	1100	50S ribosomal L13	116.94	5.65E-06	*Ars* +
DN59518_c3_g1_i1	6816	elongation factor G	116.74	8.96E-06	*Ars* +
DN53747_c0_g2_i1	3313	Secretion system effector C ( ) like family	116.16	2.93E-05	*Ars* +
DN1534_c0_g1_i1	863	lipoyl synthase	102.21	0.000363	*Ars* +
DN89650_c0_g1_i1	633	peroxidase	100.7	0.00062	*Ars* +
DN55356_c0_g1_i1	6136	30S ribosomal S1	94.74	0.000451	*Ars* +
DN59518_c5_g1_i1	12763	DNA-directed RNA polymerase subunit beta	87.36	7.77E-05	*Ars* +
DN40593_c0_g1_i1	1448	flavodoxin	85.065	0.00289	*Ars* +
DN54817_c0_g2_i1	3605	signal recognition particle	75.11	0.000397	*Ars* +
DN73782_c0_g1_i1	1464	enoyl-[acyl-carrier-] reductase	72.53	0.000556	*Ars* +
DN59212_c1_g10_i1	266	conserved hypothetical protein	70.66	2.46E-06	*Ars* +
DN65557_c3_g2_i3	684	IS5 IS1182 family transposase	68.41	0.00667	*Ars* +
DN60979_c0_g3_i2	295	hypothetical protein ALO80_101181	67.92	1.03E-06	*Ars* +
DN90989_c0_g1_i1	909	conjugal transfer pilus assembly protein TraU [*Arsenophonus nasoniae*]	54.51	0.000361	*Ars* +
DN61866_c2_g7_i2	501	daphnid bacterial-ribosomal-RNA- possible	37.51	0.000948	*Ars* +
DN61866_c2_g7_i3	425	daphnid bacterial-ribosomal-RNA- possible	22.03	0.0132	*Ars* +
DN59212_c1_g7_i1	309	hypothetical conserved	18.66	0.00201	*Ars* +
DN59212_c1_g8_i2	320	3,4-dihydroxy-2-butanone-4-phosphate synthase	13.99	2.04E-05	*Ars* +
DN65067_c1_g4_i1	282	hypothetical protein Abol_046_003	11.96	3.55E-05	*Ars* +
DN35650_c0_g1_i1	206	Putative uncharacterized protein	11.92	0.0271	*Ars* +
DN63397_c5_g8_i2	356	PG1 homology to Homo sapiens	7.62	4.95E-06	*Ars* +
Buchnera					
Unique to PL vs NL*					
DN70885_c0_g1_i1	257	outer membrane A precursor	-13.69	0.0419	*Ars* -
Shared with PF vs NF*					
DN60979_c0_g3_i1	294	hypothetical protein ALO39_101110	109.61	3.06E-06	*Ars* +
DN63397_c5_g8_i5	2763	ORF16-lacZ fusion	16.25	2.15E-07	*Ars* +
DN65502_c3_g3_i2	444	shikimate dehydrogenase (aroE) gene, complete cds; 23S ribosomal RNA and 5S ribosomal RNA genes, complete sequence; and cysteinyl-tRNA synthetase (cysS) gene, partial cds	13.8	0.000279	*Ars* +

*Abbreviations: P stands for *Arsenophonus* infected, N stands for uninfected, L is locust host plant, and F is fava host plant.

Table 2.9: Differentially expressed genes between *Arsenophonus*-infected and uninfected cowpea aphids on fava.

Transcript identifier	Gene length (bp)	Product description	Fold change	P-value	Upregulated on
Aphis craccivora					
Unique to PF vs NF*					
DN60515_c4_g10_i1	442	no match	352.94	0.038	*Ars* +
DN63642_c3_g3_i1	581	glutathione S-transferase isoform D-like	18.88	5.87E-03	*Ars* +
Arsenophonus					
Unique to PF vs NF*					
DN58358_c0_g1_i8	397	transcriptional regulator	270.71	5.90E-14	*Ars* +
DN65067_c0_g2_i2	374	daphnid bacterial-ribosomal-RNA- possible HGT	24.27	3.25E-08	*Ars* +
DN59212_c1_g3_i6	280	conserved hypothetical protein [Brucella suis bv. 4 str. 40]	21.9	0.00153	*Ars* +
DN60179_c4_g9_i1	215	IS1 transposase	6.84	2.47E-05	*Ars* +
Shared with PL vs NL*					
DN65557_c3_g6_i1	1212	transposase IS5 ssgr IS903 family	410.84	3.48E-13	*Ars* +
DN81462_c0_g1_i1	1386	porin	404.19	2.16E-16	*Ars* +
DN39876_c1_g1_i1	2783	molecular chaperone	390.68	3.59E-13	*Ars* +
DN57009_c0_g1_i1	2209	16S rRNA (cytidine(1402)-2 -O)-methyltransferase	358.24	1.72E-15	*Ars* +
DN58358_c0_g1_i6	941	SsrA-binding protein [*Arsenophonus* sp. ENCA]	342.03	1.74E-16	*Ars* +
DN24426_c0_g1_i1	1562	tRNA dihydrouridine synthase	245.81	4.04E-08	*Ars* +
DN65557_c3_g2_i2	458	IS5 IS1182 family transposase	202.59	4.04E-08	*Ars* +
DN30853_c1_g1_i1	1341	porin	187.86	5.95E-08	*Ars* +
DN65502_c3_g3_i9	416	dehydration responsive	150.89	1.96E-19	*Ars* +
DN59119_c2_g1_i1	833	primosomal replication N	147.41	1.48E-06	*Ars* +
DN65557_c20_g1_i1	657	50S ribosomal L28	147.15	1.66E-05	*Ars* +
DN59212_c1_g10_i1	266	conserved hypothetical protein [Asaia platycodi SF2.1]	142.03	1.95E-11	*Ars* +
DN39355_c0_g1_i1	502	heat-shock	136.56	1.20E-03	*Ars* +
DN29545_c0_g1_i1	3434	molecular chaperone	132.99	2.91E-05	*Ars*+
DN39355_c0_g2_i1	1045	heat-shock	130.53	1.78E-03	*Ars* +
DN59518_c3_g1_i2	1451	elongation factor Tu	122.45	1.93E-05	*Ars*+
DN65557_c3_g1_i1	548	hypothetical protein	118.33	1.66E-05	*Ars* +
DN65267_c1_g8_i8	14797	pre translocase subunit	108.87	4.08E-05	*Ars* +
DN65557_c3_g2_i1	621	IS5 IS1182 family transposase	99.51	1.00E-03	*Ars* +
DN47916_c0_g1_i1	679	50S ribosomal L21	98.36	8.92E-05	*Ars* +
DN89639_c0_g1_i1	1333	Z-ring-associated	98.01	2.48E-05	*Ars* +
DN59119_c5_g1_i1	543	50S ribosomal protein L9 [*Arsenophonus* sp. ENCA]	95.01	5.28E-04	*Ars* +
DN55229_c0_g2_i1	1100	50S ribosomal L13	94.63	8.15E-04	*Ars* +
DN90989_c0_g1_i1	909	conjugal transfer pilus assembly protein TraU [*Arsenophonus nasoniae*	93.91	0.00645	*Ars* +
DN81995_c0_g1_i1	451	murein lipo	90.05	2.46E-06	*Ars* +
DN58932_c3_g2_i1	1379	transcriptional regulator	87.78	1.23E-05	*Ars* +
DN59518_c3_g1_i1	6816	elongation factor G	83.92	2.52E-03	*Ars* +
DN58538_c2_g5_i6	266	Quinone oxidoreductase [Escherichia coli O25b:H4]	83.39	3.67E-16	*Ars* +
DN61866_c2_g7_i3	425	daphnid bacterial-ribosomal-RNA- possible	74.05	2.00E-13	*Ars* +
DN59518_c5_g1_i1	12763	DNA-directed RNA polymerase subunit beta	73.46	0.00404	*Ars* +
DN89650_c0_g1_i1	633	peroxidase	72.07	0.00156	*Ars* +
DN53747_c0_g2_i1	3313	Secretion system effector C ( ) like family	69.57	2.11E-02	*Ars* +
DN73782_c0_g1_i1	1464	enoyl-[acyl-carrier-] reductase	69.42	0.005532	*Ars* +
DN55356_c0_g1_i1	6136	30S ribosomal S1	67.79	1.28E-02	*Ars* +
DN54817_c0_g2_i1	3605	signal recognition particle	66.85	4.96E-03	*Ars* +
DN59212_c1_g8_i2	320	3,4-dihydroxy-2-butanone-4-phosphate synthase	64.68	5.81E-12	*Ars* +
DN61866_c2_g7_i2	501	daphnid bacterial-ribosomal-RNA- possible	63.32	3.10E-13	*Ars* +
DN65557_c3_g2_i3	684	IS5 IS1182 family transposase	62.57	0.013421	*Ars* +
DN1534_c0_g1_i1	863	lipoyl synthase	57.64	0.0218	*Ars* +
DN40593_c0_g1_i1	1448	flavodoxin	56.99	2.33E-02	*Ars* +
DN60979_c0_g3_i2	295	hypothetical protein ALO80_101181	23.75	0.00407	*Ars* +

Table 2.9, continued

Transcript identifier	Gene length (bp)	Product description	Fold change	P-value	Upregulated on
Arsenophonus					
Shared with PL vs NL*					
DN54389_c0_g2_i1	230	membrane {ECO:0000313	19.68	1.02E-12	*Ars* +
DN63397_c5_g8_i2	356	PG1 homology to Homo sapiens	15.87	4.22E-09	*Ars* +
DN59212_c1_g7_i1	309	conserved hypothetical protein [Asaia platycodi SF2.1]	15.74	1.67E-06	*Ars* +
DN35650_c0_g1_i1	206	Putative uncharacterized protein	15.54	9.86E-05	*Ars* +
DN65067_c1_g4_i1	282	hypothetical protein Abol_046_003	4.36	0.0517	*Ars* +
Buchnera					
Unique to PF vs NF*					
DN55094_c0_g2_i1	434	no match	-25.62	4.08E-02	*Ars* -
Shared with PL vs NL*					
DN60979_c0_g3_i1	294	hypothetical protein ALO39_101110	291.56	1.67E-06	*Ars* +
DN63397_c5_g8_i5	2763	ORF16-lacZ fusion	14.65	1.30E-10	*Ars* +
DN65502_c3_g3_i2	444	shikimate dehydrogenase (aroE) gene, complete cds; 23S ribosomal RNA and 5S ribosomal RNA genes, complete sequence; and cysteinyl-tRNA synthetase (cysS) gene, partial cds	8.72	3.59E-11	*Ars* +

*Abbreviations: P stands for *Arsenophonus* infected, N stands for uninfected, L is locust host plant, and F is fava host plant.

Chapter 3: Investigation of relative *Arsenophonus* abundance changes across two host plants

Introduction

Microbial partners have played a key role in helping insects adapt to countless environments. From the insect host's perspective, symbionts can be split into primary obligate symbionts, which are essential to the function of their hosts and provide nutritional supplementation, and facultative symbionts, which are non-essential to host function and provide a range of possible functions, including defense against parasitoids, heat tolerance, or change in plant usage (Douglas, 2015; Hansen and Moran, 2014; Oliver et al., 2010; Su et al., 2013). While obligate symbionts come at a consistent metabolic cost to their insect host, the costs associated with facultative symbionts are less stable (Douglas, 2015; Hansen and Moran, 2014). Infected hosts can experience a net cost or net benefit relative to an uninfected host, dependent on the cost to benefit ratio. The ratio itself is not fixed, as costs and benefits shift according to external and internal factors.

Many facultative symbionts are vertically transmitted from mother to offspring, and this transmission process is likely tied to both the cost and benefit of symbiotic infection. Fidelity of vertical transmission is often dependent on the size of the bacterial population, or bacterial titer, within the host; lower bacterial titer can result in imperfect transmission leading to a loss of infection (Hosokawa et al., 2007; Serbus et al., 2011). In contrast, higher titer often improves transmission efficiency (Hosokawa et al., 2007), and can increase the beneficial phenotypes induced by the symbiont (Iturbe-Ormaetxe et al., 2011), but can also impose fitness penalties on the host if titer gets too high, potentially even resulting in premature host death (Serbus et al., 2011). Thus, the host-symbiont

relationship is a balancing act, with both the host and symbiont trying to optimize costs and benefits, potentially by mediating symbiont titer (Hansen and Moran, 2014; Oliver et al., 2014).

Balancing selection works to keep facultative symbionts at intermediate titers, but various environmental conditions can influence the cost to benefit ratio, changing the optimization point for both organisms and thus influencing titer (Oliver et al., 2014). The external environmental conditions of the insect host habitat and internal insect physiology can affect symbiont titer. The insect host can act as a "environmental conduit" in which host sex, life stage, and genotype can greatly affect symbiont titer (Leclair et al., 2016; Parkinson et al., 2017; Zhang et al., 2016). Therefore, it is reasonable to believe that the internal environment of an insect host can be affected the nutritional content received from its diet resulting in changes in symbiont titer. I hypothesize, from knowledge of previous studies investigating facultative symbiont titer, when the insect host feeds on different host plants, that there is selection for higher titer of facultative symbionts on certain host plants (Enders and Miller, 2016; Ghanim and Kontsedalov, 2009; Oliver et al., 2010; Pan et al., 2013).

The aim of this study is to investigate how the titer of a beneficial facultative symbiont changes across different environments (i.e. host plants). Specifically, the facultative symbiont *Arsenophonus* has been shown to improve cowpea aphid (*Aphis craccivora*) fitness on the host plant *Robinia pseudoacacia*, black locust, relative to uninfected aphids, but to decrease aphid performance on *Vicia faba*, fava, and *Medicago sativa*, alfalfa (Wagner et al., 2015). Thus, aphids receive a net benefit from *Arsenophonus* infection when feeding on locust, but a net cost when on the other plants.

Using this system, I hypothesize that *Arsenophonus* titer will be higher when the aphid is maintained on locust (where the symbiont is more beneficial), and lower when the host aphid is maintained on fava (where the symbiont is less beneficial). To address this hypothesis, I performed three separate experiments. In the first experiment, I evaluated symbiont titer in laboratory aphid colonies that were switched to locust after years of maintenance on fava, where the symbiont conferred no advantage to the aphid host. In the second experiment, I collected new aphid colonies off locust in the field and monitored symbiont titer over time when maintained long term on locust versus fava. Finally, in the third experiment I evaluated symbiont titer of aphids collected directly from different host plants in the field to test whether *Arsenophonus* titer differed across host plants under natural conditions.

Methods

Experiments were conducted with clonal colonies of the cowpea aphid, *Aphis craccivora*. Each genetically uniform colony had been initiated with a single individual collected in North America (Table 3.1). Insects were reared on either fava or locust as host plants. Plants were seeded at 2-4 seeds per 10 cm pot in Promix potting media, grown in a greenhouse with supplemental light to ensure 16L:8D daylength. Locust seeds were scarified in boiling water, then planted 24 h later (Aliero, 2004). All plants were watered 3 times a week, and grown for 2-3 weeks prior to adding aphids. Aphids were caged and maintained at ambient laboratory temperature (22 $\pm$ 4ºC) under full spectrum lights for 16L:8D. They were caged in 3.78 L plastic jars with mesh panels for ventilation, and subsets of aphids were transferred to fresh plants every 2 weeks.

Experiment 1: To evaluate whether host plant affected *Arsenophonus* titer, I first used laboratory colonies that had been maintained for up to 6 years on fava prior to experimentation, and evaluated whether titer increased in these colonies if the aphids were switched to locust. I used 3 genetically distinct aphid clones: LE, LW, and AL (Table 3.1). Clone nomenclature is consistent with previous studies (e.g. Wagner et al. 2015). Both LE and LW were naturally infected with *Arsenophonus*, whereas AL was experimentally transinfected with *Arsenophonus* via hemolymph microinjections (Wagner et al. 2015). Before the start of and during the experiment, aphids from each colony were tested for *Arsenophonus* infection using diagnostic PCR (Table 3.2; Wagner et al. 2015). I set up 6 replicate jars from each clone on fava and 6 on locust, for a total of 36 aphid colonies as experimental units. Each colony was maintained on its host plant for 4 months.

Five fourth instar aphids were collected and pooled from each colony for DNA extraction, using DNeasy® Blood and Tissue Kit (Qiagen, Valencia, CA). I extracted DNA according to manufacturer's protocols, except reagent volumes were halved for the cell lysis step, and I used WS Gencatch™ DNA Purification Buffer (Epoch Life Science Inc., Missouri City, Texas) as a supplemental buffer. DNA concentration for each sample was first assessed using a Thermo Fisher scientific nanodrop 2000 spectrophotometer (Waltham, MA, USA) and then normalized to 5-10 ng/ul using AE Buffer (DNeasy Blood and Tissue Kit, Qiagen) as done in previous symbiont studies (Enders and Miller 2016, Martinez et al. 2014). To quantify *Arsenophonus* abundance, I used quantitative PCR to estimate the relative abundance of the *Arsenophonus* MN cell division protein (*ftsK*) in relation to aphid elongation factor 1α (*Ef1α*), both of which are single copy

genes. Primer sequences and cycling conditions are detailed in Table 3.2. All qPCR reactions were performed in 10μL volume on the StepOnePlus™ Real-Time PCR System (Waltham, MA, USA) using QuantiTect® SYBR® Green PCR Kit, with 500 nmol/L of each primer and 5-10 ng input DNA. All primers produced a single unique melt peak. Individual samples were run in triplicate along with a negative control without DNA template (Enders and Miller 2016).

The aphid host gene *Ef1α* was used to standardize differences in endosymbiont concentrations among individual extractions on the same qPCR plate by multiplying each sample C_t by a correction factor (CF = maximum Ef1α C_t/sample Ef1α C_t) (Enders and Miller 2016, Martinez et al. 2014). Relative endosymbiont abundance (RA) was calculated through $2^{-\Delta Ct}$; where ΔC_t = C_t (endosymbiont gene) – C_t (Ef1α). RA was log (x+1) transformed, to better fit assumptions of normality and equal variance. All RA analyses were conducted using SAS software (v. 9.4, SAS Institute Inc., Cary, NC). The SAS command **PROC GLM** was used to compare RA in a fully factorial model with type III errors, with host plant and aphid clone as fixed factors. For clone, post hoc multiple comparisons were performed using Tukey HSD tests on least squared means.

Experiment 2: As a separate investigation of host plant effect on *Arsenophonus* titer, I conducted a yearlong timeseries experiment comparing *Arsenophonus* titer between aphids feeding on locust versus fava using 3 newly initiated aphid colonies. The three new aphid clones (LO, LC, and LP) were initially collected from infested locust trees (Table 3.1). Individual aphids from each location were set up on fava leaves embedded in 1% agar in 35 mm Petri dishes. Once progeny were produced, aphids from each dish were tested for *Arsenophonus* infection using diagnostic PCR (Table 3.2;

Wagner et al. 2015). Once *Arsenophonus* infection was validated, genetically identical sisters from each clone were subdivided to initiate three replicate sub-colonies on locust, and three sub-colonies on fava. Experimental design, rearing methods and *Arsenophonus* quantification methods were the same as the previous experiment. RA was log (x+1) transformed and was evaluated in the 1st, 2nd, 6th, 8th, 10th, and 12th months of rearing. I compared titer using a fully factorial fixed effects model in SAS (SAS software, Version 9.4 of the SAS System for Windows Copyright© 2002-2012 SAS Institute Inc., Cary, NC, USA.) with the command **PROC GLIMMIX**. My design was unstructured to account for the unequally spaced timepoints, with treatment (fava or locust), and aphid clone (LO, LC, and LP), also included as factors. Post hoc multiple comparisons were performed using Tukey HSD tests on least squared means.

Experiment 3: My final experiment evaluated *Arsenophonus* titer in cowpea aphid specimens collected from five separate host plants (*Robinia pseudoacacia*, *Vicia faba Acacia retinodes*, *Leucanthemum paludosum* and *Rosa hybrida*) in the field. The samples came from a previous survey evaluating facultative symbiont infection prevalence in world populations of cowpea aphids and only specimens that previously tested positive for *Arsenophonus*, via diagnostic PCR, were used in the present study (Brady et al., 2014) (Table 3.3). Up to 5 randomly chosen specimens were used per population; some host plants were represented by multiple aphid populations from different locations. Because these samples were older and possibly degraded, they were diagnostically screened for the aphid CO1 gene, to ensure the sample retained sufficient quality for qPCR (Table 3.2). All DNA, as with both experiment 1 and 2, was assessed for concentration and then diluted to 5-10 ng/µl. *Arsenophonus* quantification methods were the same as previous

experiments, except that each DNA extraction came from an individual adult aphid, rather than a pool of 5 aphids. As with both other experiments, RA was log (x+1) transformed. For statistical analysis, I used **PROC GLM** in SAS, with host plant as a factor comparing *Arsenophonus* titer across the five host plants.

Results

Experiment 1: When I evaluated differences in endosymbiont titer between host plants for the three lab clonal lines, I found that there was greater *Arsenophonus* abundance on locust compared to fava ($F_{1,30} = 8.12$, $P < 0.01$, Table 3.4, Fig. 3.1). *Arsenophonus* abundance was 40% higher on locust compared to fava. The three aphid clones also varied in *Arsenophonus* titer ($F_{2,30} = 22.70$, $P < 0.001$, Table 3.4, Fig. 3.1). Post hoc multiple comparisons indicated that the transinfected clone had much lower *Arsenophonus* titer compared to its naturally-infected counterparts (Fig. 3.1). *Arsenophonus* abundance was 12.5-fold greater in the naturally infected colonies compared to the transinfected colony. *Arsenophonus* titer was not affected by the interaction between host plant and clone ($F_{2,30} = 1.44$, $P = 0.2528$, Table 3.4, Fig. 3.1)

Experiment 2: *Arsenophonus* titer on locust was greater than on fava ($F_{1,70} = 8.39$, $P < 0.01$), and changed over time ($F_{5,70} = 2.46$, $P < 0.05$, Table 3.6, Fig. 3.2). The host plant by time interaction term was marginally significant, showing a trend that *Arsenophonus* titer might be affected by the interaction between host plant and time ($F_{5,70} = 2.25$, $P = 0.0589$, Table 3.6). There was no main effect of clonal line on *Arsenophonus* titer ($F_{2,70} = 2.54$, $P = 0.0865$, Table 3.6, Fig. 3.2), but there was an interaction of clone with time, indicating variation of titer among clones across timepoints ($F_{10,70} = 3.10$, $P < 0.01$, Table 3.6).

Experiment 3: Finally, when I compared *Arsenophonus* titer among populations of aphids collected from different host plants in the field, I found titer varied significantly ($F_{4, 37} = 3.47$, $P < 0.05$). Titer was significantly higher in aphids collected from *Acacia retinodes* than any other host plant (Fig. 3.3). Titer did not differ significantly among aphids collected from the other plants.

Discussion

Overall in both experiment 1 and 2, I found that *Arsenophonus* titer was higher on locust compared to fava. This effect was not completely consistent over time in experiment 2: I observed a marginally significant host plant by time effect, and visual inspection of the data shows that the effect of titer differences between plants seemed to diminish over time (Fig. 3.2). For some timepoints, titer on fava spiked compared to locust (Fig. 3.2). In contrast, in experiment 1, which consisted of only time point, titer was higher from aphids on locust, even though these lab colonies had spent many years on fava before I initiated the experiment. This result suggests that even if fava generally selects for lower *Arsenophonus* titer, titer may remain responsive to host plant, and able to increase when the environment changes. It is also worth noting that titer in experiment 1 was lower relative to experiment 2. In contrast to both experiments 1 and 2, for experiment 3 the symbiont titer in aphids collected from host plants in the field showed no difference in titer when comparing aphids from locust versus fava, only a trend for greater titer on locust compared to fava, and the absolute titer values for both plants were low (Fig. 3.3).

Through amplification of single copy gene fragments with qPCR we can measure relative *Arsenophonus* abundance. This allows us to provide an estimate of endosymbiont

(*Arsenophonus*) gene copy number, in relation to a single copy host gene (Martinez et al. 2014). However, this method estimates relative genome abundance and not exact number of bacterial cells (Martinez et al. 2014). Every value is a comparison of the number of symbiont genome copies to a single host copy. This relative quantity is useful, however, as most endosymbionts cannot be cultured (Douglas, 2015; Hansen and Moran, 2014), and it allows us to compare microbial population sizes across treatments.

Using relative symbiont abundances, we can compare our *Arsenophonus* values to other microbial populations. In comparison to *Arsenophonus* abundances found in other experiments (Enders and Miller, 2016; Ghanim and Kontsedalov, 2009), some of my values fell into a comparable range of 0.1 to 0.5, many of my values were much lower than 0.1, and a few were much higher, ranging from 1.0 to 6.0. When relating my *Arsenophonus* relative titer values to other facultative symbiont, obligate symbiont, and pathogen titers there is a shift in scale. Other facultative symbionts often have RA values that range from 1 to 50 (Leclair et al., 2016; Pan et al., 2013). Amongst obligate symbionts, relative titer is even higher, ranging from 10 to 100 copies on average (Enders and Miller, 2016; Leclair et al., 2016; Vogel and Moran, 2011) and greater than 1000 in some cases (Parkinson et al., 2017). Pathogen titer varies over a wide range, from 1 to 800, averaging much higher than facultative symbionts (Blomquist and Kirkpatrick, 2002; Frost, Willis, and Groves, 2011; Glaser and Meola, 2010; Serbus et al., 2011; Serbus et al., 2015). These comparisons, amongst different microbial populations, show that the range of *Arsenophonus* titer values found among our experiments is inclusive of the range of *Arsenophonus* titers found in other studies, but more variable.

In experiment 1, two of the lab colonies, LW and LE, were originally collected from locust and were naturally infected with *Arsenophonus* whereas the third clone, AL, was originally collected from alfalfa and was not naturally infected with *Arsenophonus*. The AL infection had been generated by hemolymph transinfection (Wagner et al. 2015), and had been stably maintained in the lab for several years. The transinfected colony did show higher titer on locust and lower titer on fava, but titer on either plant was much lower (2-3 orders of magnitude) than the two naturally infected colonies (Fig. 3.1). Even though titer in the transinfected line was extremely low, it was still better on locust compared to fava suggesting that maintaining lines on locust may help promote retention of the symbiont. Unfortunately, this insight came a little too late for the transinfected AL line, from which the *Arsenophonus* infection was lost shortly after this experiment ended. The eventual loss of the transinfected colony is congruent with other transinfection studies, as host background can influence infection establishment and retention (Chang and Wade, 1994; Fujii et al., 2001; Russell and Moran, 2005).

In experiment 2, *Arsenophonus* titer was generally higher on locust, but not always. At some timepoints, the pattern would reverse because symbiont titer spiked on fava in comparison to its locust counterpart. The variability in *Arsenophonus* titer among clones and timepoints implies that there were uncontrolled factors affecting titer in our experiment. One possibility is that there was nutritional quality variation among both plants. Among pea aphid clones, it has been shown that there are fitness differences based on aphid clonal genotype along with differences in amino acid composition among host plants of different species (Sandstrom and Pettersson, 1994). The amino acid composition of fava may be providing a more ideal amino acid profile distribution at some timepoints

than others, allowing greater aphid growth (Douglas, 1993; Liu et al., 2013; Sandstrom and Pettersson, 1994).

A notable observation, in both the field and lab, is the preference of cowpea aphids for the new growth of locust plants. Younger plant tissues hold greater amino acid concentrations than older (Nevo and Coll, 2001), suggesting that the aphids are responding to nutritive differences between new growth and old. When collecting aphids for DNA sampling in my experiments, I collected them from all areas of both host plants, including old growth (low quality) and new growth (high quality) areas. Sometimes my samples may have differed in the proportion of aphids from old and new foliage. Additionally, locust plants grown in the greenhouse were sometimes stressed enough that even fresh growth may have been of low nutritional quality for the aphids. It is likely that some of the variation in titer I observed may have reflected the variation in quality among plants at different timepoints. However, it should be mentioned that aphids do find themselves on stressed host plants in nature, as colonies grow large and overtax the plant, so the plant quality used in the experiment may not have had an unrealistic effect on titer. Nevertheless, a more consistent plant quality, for both host plants, would be ideal for future experiments.

Sample collection methods from colonies may have also influenced titer. Equivalent to a colony life cycle on plants (Dixon, 1977, 1985), our colonies started off at low density, and the first generation on a new plant produces larger individuals that are less crowded than later generations. Sample collection was conducted on a calendar schedule, and usually coincided with uncrowded aphid colonies, but at times they would be collected from overcrowded cages. In future efforts, shorter time periods between

colony maintenance and sample collection would reduce the chances of collecting from overcrowded cages. The standard for aphids collected was non-winged and most samples collected were of only non-winged individuals. However, particularly for locust, there were sampling points in which I could not collect five 4th instar apterous aphids and would supplement with alatoid nymphs. To avoid collection of alatoid nymphs in future experiments, larger locust plants of better quality should be used.

For experiment 3, wild cowpea aphid populations also showed a difference in *Arsenophonus* titer across host plants, but not between locust and fava. There was a trend for greater titer in locust compared to fava populations (Fig. 3.3), but there was again quite substantial variation in titer in aphids from both plants, with some locust-collected aphids having very low titer (1.21E-7), and at least one collected from fava having relatively high titer for *Arsenophonus* (0.0776). Interestingly, the greatest *Arsenophonus* titer levels were found on the host plant *A. retinodes* (Fig. 3.3). This data would suggest that *A. retinodes*, particularly compared to *Leucanthemum paludosum* and *Rosa hybrida*, is a relatively good host plant for *Arsenophonus* infected aphids to colonize.

All three host plants, locust, fava and *A. retinodes*, are from different genera within the family Fabaceae (Doyle and Luckow, 2003). Black locust and fava are a part of two separate, but closely related, lineages within the larger clade of Hologalegina, while *A. retinodes* is distantly related within the clade of Mimosoideae (Doyle and Luckow, 2003). It is possible that the physiology, nutritional and chemical profiles are not extremely different among these plants, meaning that *Arsenophonus* might be providing relatively the same benefits across all three host plants. Field aphids disperse to new plants like "aerial plankton," moving away from the plant they are on and trying to

establish on whatever plants they land on. For future studies, I would want to address how *Arsenophonus* titer in field *Aphis craccivora* populations changes across host plants from Fabaceae, including these three plants. With a new study, I could factor in the nutritional change between plants (in the form of amino acid profiles and nitrogen content), and plant secondary phytochemical profiles.

Ultimately, *Arsenophonus* titer was affected by aphid host plant species, but also exhibited substantial variation over time, presumably in response to uncontrolled variables such as host plant quality and aphid crowding conditions. My experiments show that *Arsenophonus* titer is dynamic, and may contribute to the phenotypic plasticity of cowpea aphid across host plants. A clearer understanding of the mechanistic benefit *Arsenophonus* provides might yield insight into the role facultative symbionts may play in the evolution of dietary breadth and host plant usage in herbivores. Future experiments addressing *Arsenophonus* titer should control for aphid overcrowding, host plant quality, more frequent time point collections, and the addition of uninfected and transinfected colonies.

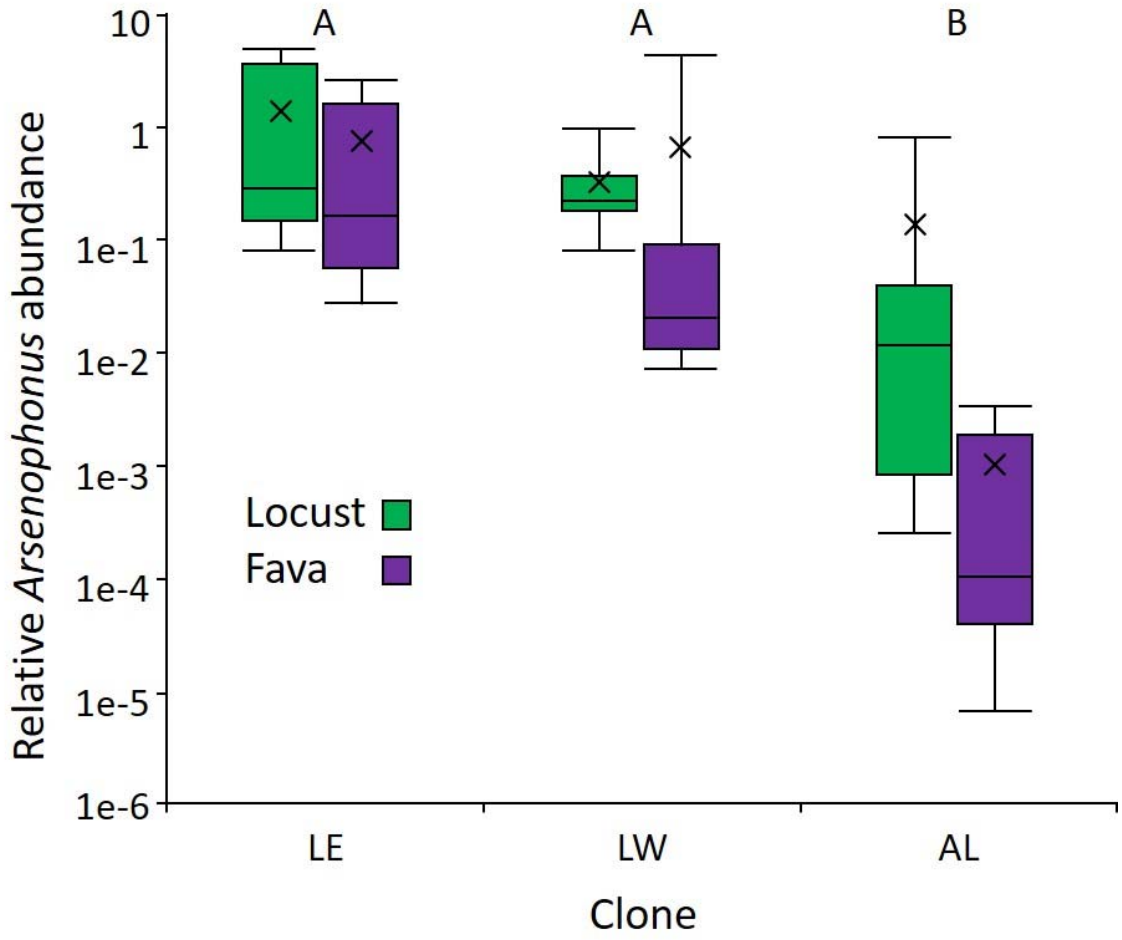

Figure 3.1: Relative *Arsenophonus* abundance in experiment 1, of three clonal aphid lines maintained on either fava or locust, as measured using qPCR of the *Arsenophonus* gene *ftsK* relative to the aphid EF1-alpha gene. The x in the figures represents the mean, while the middle line in each box is the median. The top and bottom of each box are the 1st and 3rd quartile values respectively. The whiskers of the plot represent maximum and minimum values.

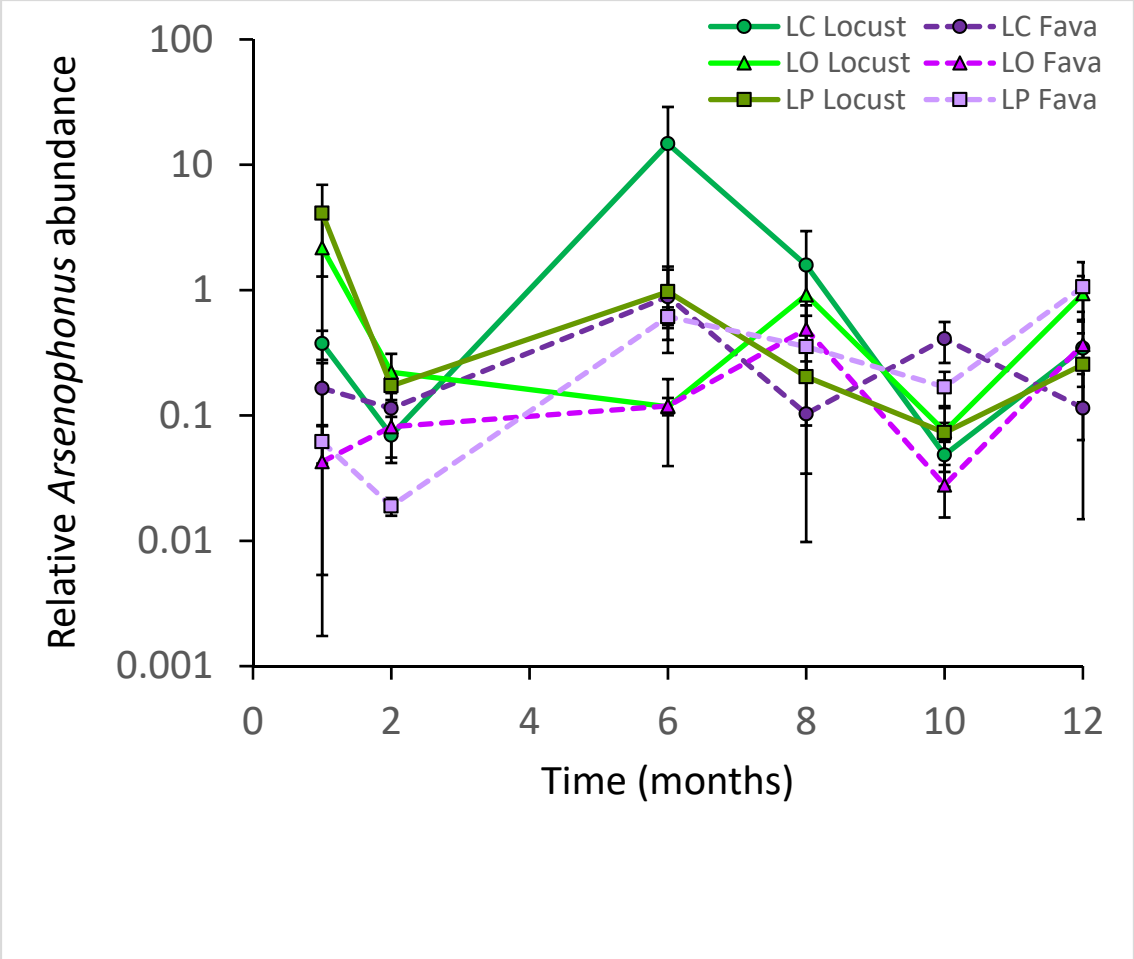

Figure 3.2: Relative *Arsenophonus* abundance in experiment 2. Mean ± SE relative *Arsenophonus* abundance in three clonal aphid lines (LC, LO, LP) maintained over a 12-month period on either locust or fava host plants. *Arsenophonus* abundance was measured using qPCR, comparing the bacterial gene *ftsK* relative to the aphid EF1-alpha gene.

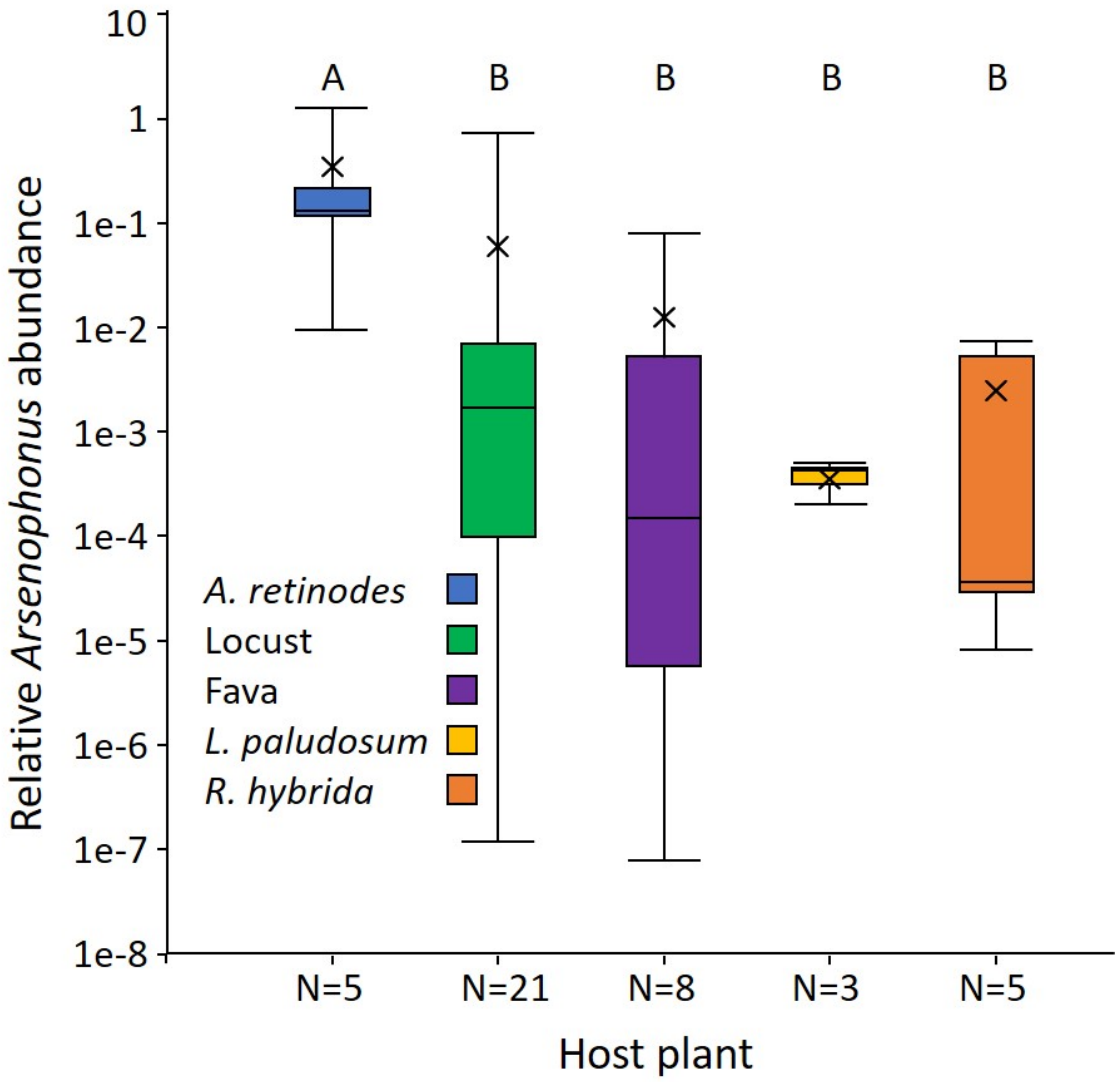

Figure 3.3: Relative *Arsenophonus* abundance in aphids collected from different host plants in the field, for experiment 3 measured using qPCR, comparing the bacterial gene *ftsK* relative to the aphid EF1-alpha gene. The x in the figures represents the mean, while the middle line in each box is the median. The top and bottom of each box are the 1st and 3rd quartile values respectively. The whiskers of the plot represent maximum and minimum values. Sample size per host plant indicated along the x-axis. Columns with different letters differed significantly at $\alpha = 0.05$.

Table 3.1: Initial collection dates and locations of the aphid clonal colonies.

Clonal Line	Host plant of origin	Infection status	Means of symbiont infection	Collection location	Date of collection
Experiment	1				
LW	Locust	Ars +	Naturally infected	37º57'N 84º23'W	Sept 2011
LE	Locust	Ars +	Naturally infected	37º57'N 84º43'W	Sept 2011
AL	Alfalfa	Ars +	Transinfected	38º04'N 84º39'W	Aug 2011
Experiment	2				
LO	Locust	Ars +	Naturally infected	39º07'N 84º29'W	July 2016
LC	Locust	Ars +	Naturally infected	45º02'N 93º30'W	Aug 2016
LP	Locust	Ars +	Naturally infected	46º36'N 94º18'W	Aug 2016

Table 3.2: Primers, PCR and qPCR cycling conditions, and references used for symbiont diagnostics, symbiont housekeeping and aphid housekeeping genes.

Target Organism	Target Gene	Primer Name	Primer Sequence 5' to 3'	References	PCR/qPCR Cycling Conditions
PCR					
Arsenophonus	23s	Ars23sF Ars23sR	CGTTTGATGAATTCATAGTCAAA GGTCCTCCAGTTAGTGTTACCCAAC	Brady and White 2013	95°C for 2 min, then 35 cycles consisting of 92°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec
qPCR					
Arsenophonus	*ftsK*	ftskF ftskR	TCAAGGTGGCGCTGAATCTT CGGGCTTACCTCTAGCTTTCC	Enders and Miller 2016	95°C for 2 min, then 35 cycles consisting of 92°C for 30 sec, 52°C for 30 sec, and 72°C for 30 sec
PCR					
Aphis craccivora	CO1	LCO1490 HCO700	GGTCAACAAATCATAAAGATATTGG TCAGGGTGACCAAAAAATCA	Breton et al. 2006	95°C for 2 min, then 35 cycles consisting of 92°C for 30 sec, 50°C for 30 sec, and 72°C for 1.5 min
qPCR					
Aphis craccivora	EF1α	EF1aF EF1R	CGCACCTGGTCACAGAGATT TGCTCACGGGTTTGTCCATT	Enders and Miller 2016	95°C for 2 min, then 35 cycles consisting of 92°C for 30 sec, 52°C for 30 sec, and 72°C for 1 min

Table 3.3: The collection location, number of samples, and host plant of origin for the field collected adult cowpea aphids.

Location, Nationality	Location, Region	# Samples	Host plant of origin
Greece	Poligono	5	*A. retinodes*
China	Beijing	3	*R. pseudoacacia* (locust)
Iran	Mashhad	5	*R. pseudoacacia* (locust)
Serbia	Mt. Dukat	5	*R. pseudoacacia* (locust)
Spain	Astorga	5	*R. pseudoacacia* (locust)
USA	Reno	3	*R. pseudoacacia* (locust)
Algeria	Ghardaia	3	*V. faba* (fava)
Algeria	Biskra	5	*V. faba* (fava)
China	Langfang	3	*L. paludosum*
France	Antibes	5	*R. hybrida*

Table 3.4: ANOVA table experiment 1.

Factor (s)	ss	d.f.	ms	*F* Value	*P*
Host plant	7.483	1	7.483	8.12	0.0078
Clone	41.82	2	20.91	22.70	<0.0001
Host plant x Clone	2.654	2	1.327	1.44	0.2528
Error	27.64	30	0.921		

Table 3.5: Statistical table for GLIMMIX output in experiment 2.

Factor (s)	d.f.	Total d.f.	*F* value	*P*
Host plant	1	70	8.39	0.0050
Clone	2	70	2.54	0.0865
Time	5	70	2.46	0.0409
Host plant x Clone	2	70	1.13	0.3289
Host plant x Time	5	70	2.25	0.0589
Clone x Time	10	70	3.10	0.0026
Host plant x Clone x Time	10	70	1.81	0.0741

Chapter 4: Conclusions and future directions

My thesis has explored the mechanistic basis for how the facultative symbiont *Arsenophonus* facilitates the use of the host plant *Robinia pseudoacacia* (locust) in the cowpea aphid, *Aphis craccivora* (cowpea aphid). Using RNA-seq of the de novo assembled *Aphis craccivora-Arsenophonus-Buchnera* reference transcriptome, I looked at differential expression of genes relative both host plant (locust and fava) and infection (*Arsenophonus* infected and uninfected aphids). In comparing *Arsenophonus* infected aphids on locust to *Arsenophonus* infected aphids on fava, I only found five differentially expressed genes, none of which were from *Arsenophonus*. Perhaps the only gene of consequence from this comparison was ompA from the obligate symbiont *Buchnera*, a membrane bound porin protein that allows passage of small molecules like amino acids. This gene was upregulated in infected aphids feeding on fava, suggesting that there might be increased nutritional need on fava, if ompA is being used for increased amino acid transport. As a control, uninfected aphids on locust were compared to uninfected aphids on fava, to identify differentially expressed genes related to host plant use but not related to *Arsenophonus*. I found differential expression of 44 genes. Two upregulated aphid genes on locust were maltase A3-like, which metabolizes glucose, and tropomyosin isoform X10, a filament protein involved in muscle contractions. In this control comparison, there was also one differentially expressed *Buchnera* gene, serine acetyltransferase, which is a part of the pathway that catalyzes the amino acid serine to cysteine and can be involved in B-vitamin production. The differential expression of these genes suggests that *Aphis craccivora* and its primary symbiont may experience some nutritional differences while feeding on locust versus fava. Finally, with an

expectation of over representation of *Arsenophonus* genes, I compared *Arsenophonus* infected aphids to uninfected aphids across each host plant. I found 75 and 56 differentially expressed genes for locust and fava respectively. Most of these genes were from *Arsenophonus*, 71 and 50 for locust and fava respectively. There were 49 genes found in common between aphids on both host plants, most of which were from *Arsenophonus* and were bacterial maintenance genes. However, two standout *Arsenophonus* genes were lipoyl synthase and 3,4-dihydroxy-2-butanone-4-phosphate. Both were highly upregulated, and connected to production of intermediate products used in B-vitamin pathways. Unique to the locust comparison, beta-ketoacyl-[acyl-carrier-] synthase I was another highly upregulated *Arsenophonus* gene that is involved in B-vitamin production pathways. Suggestively, an *Arsenophonus* version of the gene serine acetyltransferase was found highly upregulated on the locust comparison. These expressed *Arsenophonus* genes would suggest that the facultative symbiont may play a role in B-vitamin production. Future studies should explore the transcriptional response of *Aphis craccivora*, with *Arsenophonus*, on locust and fava under different conditions, like nitrogen fertilizer in the soil, drought stress, and heat stress, to understand how the *Arsenophonus* infected cowpea aphid responds to each host plant respectively. The concept of *Arsenophonus* B-vitamin production should also be explored, through experimental manipulation, to identify if it is indeed the mechanism linked to host plant facilitation of locust.

With my second study, I investigated how *Arsenophonus* titer changed between aphids fed on locust (the host plant where *Arsenophonus* provides an advantage) and fava (the host plant where *Arsenophonus* does not provide an advantage). Within this study I

ran three separate experiments. In my first experiment, I measured titer in three lab colonies fed on either locust or fava. Two of the colonies use were naturally infected with *Arsenophonus* and the third had been transinfected via hemolymph injection. *Arsenophonus* titer was greater on locust compared to fava, and was higher in naturally infected than the transinfected clone. The second experiment used three new aphid colonies, collected from aphids feeding on locust in the field, and monitored the long-term changes in *Arsenophonus* titer for aphids feeding on both locust and fava. Here, I similarly found greater *Arsenophonus* titer on locust compared to fava, but the effect seemed to diminish over time and titer occasionally spiked on fava compared to locust. The titer fluctuations over time imply that *Arsenophonus* function might vary temporally. Each aphid clone had its own distinct fluctuations in titer as well. Finally, my third experiment measured *Arsenophonus* titer in aphids collected directly from different host plants in the field. *Arsenophonus* titer was not greater on locust compared to fava, although there was a trend toward higher values on locust than fava. For one host plant, *Acacia retinodes*, *Arsenophonus* titer was greater compared to the titer of all the other plants. This implies that *Acacia retinodes* may be another host for which *Arsenophonus* infection could be beneficial for *A. craccivora*.

There are two routes future studies should pursue. The first is to more thoroughly investigate the changes in *Arsenophonus* titer over time. The experiment should control for overcrowding and plant quality, and colonies should be sampled more frequently. The second would be to explore how *Arsenophonus* titer from aphids in the field relates to a broader range of host plants from the family Fabaceae. In congruence with measuring the

titer, it would be useful to evaluate nutritional plant secondary phytochemical changes among plants.

References

Agrawal, A. A., and Weber, M. G. (2015). On the study of plant defence and herbivory using comparative approaches: how important are secondary plant compounds. *Ecology Letters, 18*(10), 985-991.

Baumann, P. (2005). Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annual Review of Microbiology, 59*, 155-189.

Becerra, J. X., Noge, K., and Venable, D. L. (2009). Macroevolutionary chemical escalation in an ancient plant–herbivore arms race. *Proceedings of the National Academy of Sciences*, *106*(43), 18062-18066.

Bennett, G. M., and Moran, N. A. (2015). Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. *Proceedings of the National Academy of Sciences, 112*(33), 10169-10176.

Berenbaum, M. (1983). Coumarins and caterpillars: a case for coevolution. *Evolution, 37*(1), 163-179.

Blomquist, C., and Kirkpatrick, B. (2002). Frequency and seasonal distribution of pear psylla infected with the pear decline phytoplasma in California pear orchards. *Phytopathology, 92*(11), 1218-1226.

Brady, C. M., and White, J. A. (2013). Cowpea aphid (*Aphis craccivora*) associated with different host plants has different facultative endosymbionts. *Ecological Entomology*, *38*(4), 433-437.

Brady, C.M., Asplen, M.K., Desneux, N., Heimpel, G.E., Hopper, K.R., Linnen, C.R., Oliver, K.M., Wulff, J.A. and White, J.A., 2014. (2014). Worldwide populations of the aphid *Aphis craccivora* are infected with diverse facultative bacterial symbionts. *Microbial Ecology, 67*(1), 195-204.

Brattsten, L., Wilkinson, C., and Eisner, T. (1977). Herbivore-plant interactions: mixed-function oxidases and secondary plant substances. *Science*, *196*(4296), 1349-1352.

Brisson, J. A., Davis, G. K., and Stern, D. L. (2007). Common genome-wide patterns of transcript accumulation underlying the wing polyphenism and polymorphism in the pea aphid (*Acyrthosiphon pisum*). *Evolution and Development, 9*(4), 338-346.

Brisson, J. A., Ishikawa, A., and Miura, T. (2010). Wing development genes of the pea aphid and differential gene expression between winged and unwinged morphs. *Insect Molecular Biology, 19*(s2), 63-73.

Chang, N. W., and Wade, M. J. (1994). The transfer of *Wolbachia pipientis* and reproductive incompatibility between infected and uninfected strains of the flour beetle, *Tribolium confusum*, by microinjection. *Canadian Journal of Microbiology, 40*(11), 978-981.

Conner, W. E., Boada, R., Schroeder, F. C., González, A., Meinwald, J., and Eisner, T. (2000). Chemical defense: bestowal of a nuptial alkaloidal garment by a male moth on its mate. *Proceedings of the National Academy of Sciences, 97*(26), 14406-14411.

Dale, C., Beeton, M., Harbison, C., Jones, T., and Pontes, M. (2006). Isolation, pure culture, and characterization of "*Candidatus Arsenophonus arthropodicus*," an intracellular secondary endosymbiont from the hippoboscid louse fly *Pseudolynchia canariensis*. *Applied and Environmental Microbiology, 72*(4), 2997-3004.

Darby, A. C., Choi, J. H., Wilkes, T., Hughes, M. A., Werren, J. H., Hurst, G. D. D., and Colbourne, J. K. (2010). Characteristics of the genome of *Arsenophonus nasoniae*, son-killer bacterium of the wasp *Nasonia*. *Insect Molecular Biology, 19*, 75-89.

Degnan, P. H., and Moran, N. A. (2008). Evolutionary genetics of a defensive facultative symbiont of insects: exchange of toxin-encoding bacteriophage. *Molecular Ecology, 17*(3), 916-929.

Dixon, A. (1977). Aphid ecology: life cycles, polymorphism, and population regulation. *Annual Review of Ecology and Systematics, 8*(1), 329-353.

Dixon, A. (1985). Structure of aphid populations. *Annual Review of Entomology, 30*(1), 155-174.

Douglas, A. E. (1993). The nutritional quality of phloem sap utilized by natural aphid populations. *Ecological Entomology, 18*(1), 31-38.

Douglas, A. E. (2015). Multiorganismal insects: diversity and function of resident microorganisms. *Annual Review of Entomology, 60*, 17-34.

Downie, D., and Gullan, P. (2005). Phylogenetic congruence of mealybugs and their primary endosymbionts. *Journal of Evolutionary Biology, 18*(2), 315-324.

Doyle, J. J., and Luckow, M. A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiology, 131*(3), 900-910.

Duron, O., Bouchon, D., Boutin, S., Bellamy, L., Zhou, L. Q., Engelstadter, J., and Hurst, G. D. (2008). The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biology*, *6*(1), 27.

Dussourd, D. E., and Eisner, T. (1987). Vein-cutting behavior: insect counterploy to the latex defense of plants. *Science, 237*(4817), 898-902.

Ehrlich, P. R., and Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, *18*(4), 586-608.

Eisner, T., Johnessee, J. S., Carrel, J., Hendry, L. B., and Meinwald, J. (1974). Defensive use by an insect of a plant resin. *Science, 184*(4140), 996-999.

Enders, L. S., and Miller, N. J. (2016). Stress-induced changes in abundance differ among obligate and facultative endosymbionts of the soybean aphid. *Ecology and Evolution*, *6*(3), 818-829.

Ferrari, J., and Vavre, F. (2011). Bacterial symbionts in insects or the story of communities affecting communities. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1569), 1389-1400.

Ferrari, J., West, J. A., Via, S., and Godfray, H. C. J. (2012). Population genetic structure and secondary symbionts in host-associated populations of the pea aphid complex. *Evolution, 66*(2), 375-390.

Fisher, R. M., Henry, L. M., Cornwallis, C. K., Kiers, E. T., and West, S. A. (2017). The evolution of host-symbiont dependence. *Nature Communications, 8*, 15973.

Frago, E., Dicke, M., and Godfray, H. C. J. (2012). Insect symbionts as hidden players in insect–plant interactions. *Trends in Ecology and Evolution, 27*(12), 705-711.

Frost, K., Willis, D., and Groves, R. (2011). Detection and variability of aster yellows phytoplasma titer in its insect vector, *Macrosteles quadrilineatus* (Hemiptera: *Cicadellidae*). *Journal of Economic Entomology, 104*(6), 1800-1815.

Fujii, Y., Kageyama, D., Hoshizaki, S., Ishikawa, H., and Sasaki, T. (2001). Transfection of *Wolbachia* in Lepidoptera: the feminizer of the adzuki bean borer *Ostrinia*

scapulalis causes male killing in the Mediterranean flour moth *Ephestia kuehniella*. *Proceedings of the Royal Society of London B: Biological Sciences, 268*(1469), 855-859.

Futuyma, D. J., and Agrawal, A. A. (2009). Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences, 106*(43), 18054-18061.

Ghanim, M., and Kontsedalov, S. (2009). Susceptibility to insecticides in the Q biotype of *Bemisia tabaci* is correlated with bacterial symbiont densities. *Pest Management Science, 65*(9), 939-942.

Glaser, R. L., and Meola, M. A. (2010). The native *Wolbachia* endosymbionts of *Drosophila melanogaster* and *Culex quinquefasciatus* increase host resistance to West Nile virus infection. *PLOS ONE, 5*(8), e11977.

Hammer, T. J., and Bowers, M. D. (2015). Gut microbes may facilitate insect herbivory of chemically defended plants. *Oecologia, 179*(1), 1-14.

Hansen, A. K., and Moran, N. A. (2014). The impact of microbial symbionts on host plant utilization by herbivorous insects. *Molecular Ecology, 23*(6), 1473-1496.

Henry, L. M., Peccoud, J., Simon, J.-C., Hadfield, J. D., Maiden, M. J., Ferrari, J., and Godfray, H. C. J. (2013). Horizontally transmitted symbionts and host colonization of ecological niches. *Current Biology, 23*(17), 1713-1717.

Hosokawa, T., Kikuchi, Y., and Fukatsu, T. (2007). How many symbionts are provided by mothers, acquired by offspring, and needed for successful vertical transmission in an obligate insect-bacterium mutualism? *Molecular Ecology, 16*(24), 5316-5325.

Iturbe-Ormaetxe, I., Walker, T., and LO'Neill, S. (2011). *Wolbachia* and the biological control of mosquito-borne disease. *EMBO reports, 12*(6), 508-518.

Janson, M., Stireman, O., Singer, S., and Abbot, P. (2008). Phytophagous insect–microbe mutualisms and adaptive evolutionary diversification. *Evolution*, *62*(5), 997-1012.

Jousselin, E., Coeur d'acier, A., Vanlerberghe-Masutti, F., and Duron, O. (2013). Evolution and diversity of *Arsenophonus* endosymbionts in aphids. *Molecular Ecology, 22*(1), 260-270.

Koga, R., Bennett, G. M., Cryan, J. R., and Moran, N. A. (2013). Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. *Environmental Microbiology, 15*(7), 2073-2081.

Kredich, N. M., and Tomkins, G. M. (1966). The enzymic synthesis of L-cysteine in *Escherichia coli* and *Salmonella typhimurium*. *Journal of Biological Chemistry, 241*(21), 4955-4965.

Krieger, R. I., Feeny, P. P., and Wilkinson, C. F. (1971). Detoxication enzymes in the guts of caterpillars: an evolutionary answer to plant defenses? *Science, 172*(3983), 579-581.

Lamelas, A., Gosalbes, M. J., Manzano-Marín, A., Peretó, J., Moya, A., and Latorre, A. (2011). *Serratia symbiotica* from the aphid *Cinara cedri*: a missing link from facultative to obligate insect endosymbiont. *PLOS Genetics, 7*(11), e1002357.

Leclair, M., Polin, S., Jousseaume, T., Simon, J.C., Sugio, A., Morlière, S., Fukatsu, T., Tsuchida, T. and Outreman, Y. (2016). Consequences of coinfection with protective symbionts on the host phenotype and symbiont titres in the pea aphid system. *Insect Science* , *24*(5), 798-808.

Lee, S. H., Baek, J. H., and Yoon, K. A. (2016). Differential properties of venom peptides and proteins in solitary vs. social hunting wasps. *Toxins, 8*(2), 32.

Leonardo, T. E., and Muiru, G. T. (2003). Facultative symbionts are associated with host plant specialization in pea aphid populations. *Proceedings of the Royal Society of London B: Biological Sciences, 270*(Suppl 2), S209-S212.

Li, X., Schuler, M. A., and Berenbaum, M. R. (2002). Jasmonate and salicylate induce expression of herbivore cytochrome P450 genes. *Nature, 419*(6908), 712-715.

Liu, X. P., Gong, C. M., Fan, Y. Y., Eiblmeier, M., Zhao, Z., Han, G., and Rennenberg, H. (2013). Response pattern of amino compounds in phloem and xylem of trees to soil drought depends on drought intensity and root symbiosis. *Plant Biology, 15*, 101-108.

Login, F. H., and Heddi, A. (2013). Insect immune system maintains long-term resident bacteria through a local response. *Journal of Insect Physiology, 59*(2), 232-239.

Manning, P. A., Pugsley, A. P., and Reeves, P. (1977). Defective growth functions in mutants of *Escherichia coli* K12 lacking a major outer membrane protein. *Journal of Molecular Biology, 116*(2), 285-300.

McCutcheon, J. P., McDonald, B. R., and Moran, N. A. (2009). Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proceedings of the National Academy of Sciences, 106*(36), 15394-15399.

McCutcheon, J. P., and von Dohlen, C. D. (2011). An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Current Biology, 21*(16), 1366-1372.

McLean, A. H., Parker, B. J., Hrček, J., Henry, L. M., and Godfray, H. C. J. (2016). Insect symbionts in food webs. *Philosophical Transactions of the Royal Society B, 371*(1702), 20150325.

Meng, J., Zhang, C., Chen, X., Cao, Y., and Shang, S. (2014). Differential protein expression in the susceptible and resistant *Myzus persicae* (Sulzer) to imidacloprid. *Pesticide Biochemistry and Physiology, 115*, 1-8.

Moran, N. A. (2007). Symbiosis as an adaptive process and source of phenotypic complexity. *Proceedings of the National Academy of Sciences, 104*(suppl 1), 8627-8633.

Moran, N. A., Tran, P., and Gerardo, N. M. (2005). Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum Bacteroidetes. *Applied and Environmental Microbiology, 71*(12), 8802-8810.

Munson, M., Baumann, P., Clark, M., Baumann, L., Moran, N., Voegtlin, D., and Campbell, B. (1991). Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. *Journal of Bacteriology, 173*(20), 6321-6324.

Musser, R. O., Hum-Musser, S. M., Eichenseer, H., Peiffer, M., Ervin, G., Murphy, J. B., and Felton, G. W. (2002). Herbivory: caterpillar saliva beats plant defences. *Nature, 416*(6881), 599-600.

Nevo, E., and Coll, M. (2001). Effect of nitrogen fertilization on *Aphis gossypii* (Homoptera: Aphididae): variation in size, color, and reproduction. *Journal of Economic Entomology, 94*(1), 27-32.

Nováková, E., Husník, F., Šochová, E., and Hypša, V. (2015). *Arsenophonus* and *Sodalis* symbionts in louse flies: an analogy to the *Wigglesworthia* and *Sodalis* system in tsetse flies. *Applied and Environmental Microbiology, 81*(18), 6189-6199.

Nováková, E., Hypša, V., Klein, J., Foottit, R. G., von Dohlen, C. D., and Moran, N. A. (2013). Reconstructing the phylogeny of aphids (Hemiptera: Aphididae) using DNA of the obligate symbiont *Buchnera aphidicola*. *Molecular Phylogenetics and Evolution, 68*(1), 42-54.

Oliver, K. M., Degnan, P. H., Burke, G. R., and Moran, N. A. (2010). Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annual Review of Entomology*, *55*, 247-266.

Oliver, K. M., and Martinez, A. J. (2014). How resident microbes modulate ecologically-important traits of insects. *Current Opinion in Insect Science, 4*, 1-7.

Oliver, K. M., Smith, A. H., and Russell, J. A. (2014). Defensive symbiosis in the real world -advancing ecological studies of heritable, protective bacteria in aphids and beyond. *Functional Ecology, 28*(2), 341-355.

Pan, H.P., Chu, D., Liu, B.M., Xie, W., Wang, S.L., Wu, Q.J., Xu, B.Y. and Zhang, Y.J., (2013). Relative amount of symbionts in insect hosts changes with host-plant adaptation and insecticide resistance. *Environmental Entomology, 42*(1), 74-78.

Parkinson, J. F., Gobin, B., and Hughes, W. O. (2017). The more, the merrier? Obligate symbiont density changes over time under controlled environmental conditions, yet holds no clear fitness consequences. *Physiological Entomology* , *42*(2), 163-172.

Portman, S. L. (2013). Variability in host plant chemistry effects flight muscle development, gene expression, and flight capacity of adult lepidopteran herbivores (PhD thesis). Pennsylvania State University. Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3585615)

Richards, L. A., Dyer, L. A., Forister, M. L., Smilanich, A. M., Dodson, C. D., Leonard, M. D., and Jeffrey, C. S. (2015). Phytochemical diversity drives plant–insect community diversity. *Proceedings of the National Academy of Sciences*, *112*(35), 10973-10978.

Russell, J. A., and Moran, N. A. (2005). Horizontal transfer of bacterial symbionts: heritability and fitness effects in a novel aphid host. *Applied and Environmental Microbiology, 71*(12), 7987-7994.

Sabater-Muñoz, B., Toft, C., Alvarez-Ponce, D., and Fares, M. A. (2017). Chance and necessity in the genome evolution of endosymbiotic bacteria of insects. *The ISME Journal, 11*(6), 1291-1304.

Sagot, F. J. S., Henri, H., Zchori-Fein, E., Latorre, A., Moya, A., and Barbe, V. (2015). Two host clades, two bacterial arsenals: evolution through gene losses in facultative endosymbionts. *Genome Biology and Evolution*, *7*(3), 839-855.

Sandstrom, J., and Pettersson, J. (1994). Amino acid composition of phloem sap and the relation to intraspecific variation in pea aphid (*Acyrthosiphon pisum*) performance. *Journal of Insect Physiology, 40*(11), 947-955.

Schoonhoven, L. M., Van Loon, J. J., and Dicke, M. (2005). *Insect-plant biology*: Oxford University Press on Demand.

Serbus, L. R., Ferreccio, A., Zhukova, M., McMorris, C. L., Kiseleva, E., and Sullivan, W. (2011). A feedback loop between *Wolbachia* and the *Drosophila gurken*

mRNP complex influences *Wolbachia* titer. *Journal of Cell Science, 124*(24), 4299-4308.
Serbus, L. R., White, P. M., Silva, J. P., Rabe, A., Teixeira, L., Albertson, R., and Sullivan, W. (2015). The impact of host diet on *Wolbachia* titer in *Drosophila*. *PLOS Pathogens, 11*(3), e1004777.
Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. (2000). Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp APS. *Nature, 407*(6800), 81-86.
Stafford-Banks, C. A., Rotenberg, D., Johnson, B. R., Whitfield, A. E., and Ullman, D. E. (2014). Analysis of the salivary gland transcriptome of *Frankliniella occidentalis*. *PLOS ONE, 9*(4), e94447.
Strauss, E., Kinsland, C., Ge, Y., McLafferty, F. W., and Begley, T. P. (2001). Phosphopantothenoylcysteine synthetase from *Escherichia coli* identification and characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria. *Journal of Biological Chemistry, 276*(17), 13513-13516.
Su, Q., Zhou, X., and Zhang, Y. (2013). Symbiont-mediated functions in insect hosts. *Communicative and Integrative Biology, 6*(3), e23804.
Sugawara, E., and Nikaido, H. (1992). Pore-forming activity of OmpA protein of *Escherichia coli*. *Journal of Biological Chemistry, 267*(4), 2507-2511.
Tamas, I., Klasson, L. M., Sandström, J. P., and Andersson, S. G. (2001). Mutualists and parasites: how to paint yourself into a (metabolic) corner. *FEBS Letters, 498*(2-3), 135-139.
Thao, M. L., Clark, M. A., Baumann, L., Brennan, E. B., Moran, N. A., and Baumann, P. (2000). Secondary endosymbionts of psyllids have been acquired multiple times. *Current Microbiology, 41*(4), 300-304.
Thao, M. L., Moran, N. A., Abbot, P., Brennan, E. B., Burckhardt, D. H., and Baumann, P. (2000). Cospeciation of psyllids and their primary prokaryotic endosymbionts. *Applied and Environmental Microbiology, 66*(7), 2898-2905.
Toju, H., Tanabe, A. S., Notsu, Y., Sota, T., and Fukatsu, T. (2013). Diversification of endosymbiosis: replacements, co-speciation and promiscuity of bacteriocyte symbionts in weevils. *The ISME Journal, 7*(7), 1378.
Toomey, R. E., and Wakil, S. J. (1966). Studies on the mechanism of fatty acid synthesis XVI. Preparation and general properties of acyl-malonyl acyl carrier protein-condensing enzyme from *Escherichia coli*. *Journal of Biological Chemistry, 241*(5), 1159-1165.
Tsuchida, T., Koga, R., Matsumoto, S., and Fukatsu, T. (2011). Interspecific symbiont transfection confers a novel ecological trait to the recipient insect. *Biology Letters, 7*(2), 245-248.
Urban, J. M., and Cryan, J. R. (2012). Two ancient bacterial endosymbionts have coevolved with the planthoppers (Insecta: Hemiptera: Fulgoroidea). *BMC Evolutionary Biology, 12*(1), 87.
Van Valen, L. (1977). The red queen. *The American Naturalist, 111*(980), 809-810.
Vogel, K. J., and Moran, N. A. (2011). Effect of host genotype on symbiont titer in the aphid—*Buchnera* symbiosis. *Insects, 2*(3), 423-434.

Wagner, S.M., Martinez, A.J., Ruan, Y.M., Kim, K.L., Lenhart, P.A., Dehnel, A.C., Oliver, K.M. and White, J.A. (2015). Facultative endosymbionts mediate dietary breadth in a polyphagous herbivore. *Functional Ecology*, *29*(11), 1402-1410.

Wang, Y. (2002). The function of OmpA in *Escherichia coli*. *Biochemical and Biophysical Research Communications, 292*(2), 396-401.

Wilkes, T. E., Darby, A. C., Choi, J. H., Colbourne, J. K., Werren, J. H., and Hurst, G. D. D. (2010). The draft genome sequence of *Arsenophonus nasoniae*, son-killer bacterium of *Nasonia vitripennis*, reveals genes associated with virulence and symbiosis. *Insect Molecular Biology, 19*, 59-73.

Wink, M. (2013). Evolution of secondary metabolites in legumes (Fabaceae). *South African Journal of Botany, 89*, 164-175.

Xue, J., Zhou, X., Zhang, C.X., Yu, L.L., Fan, H.W., Wang, Z., Xu, H.J., Xi, Y., Zhu, Z.R., Zhou, W.W. and Pan, P.L. (2014). Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biology, 15*(12), 521.

Zhang, Y.C., Cao, W.J., Zhong, L.R., Godfray, H. C. J., and Liu, X.D. (2016). Host plant determines the population size of an obligate symbiont (*Buchnera aphidicola*) in aphids. *Applied and Environmental Microbiology, 82*(8), 2336-2346.

Thorsten Hansen

Vita

Place of Birth:

Gurnee, Illinois USA

Education:

B.S., **Biology.** University of Illinois at Champaign-Urbana, Champaign-Urbana, IL. Completed May 2015

Relevant Experience:

2015-Present **Graduate Research Assistant, Endosymbiont Ecology Lab, University of Kentucky** [Dr. Jennifer White]

2013-2015 **Lab Technician/Undergraduate Researcher, Endosymbiont Genomics Lab, University of Illinois Urbana-Champaign** [Dr. Allison Hansen]

2011-2013 **Lab Technician, Medical Entomology Lab, Illinois Natural History Survey Prairie Research Institute, Illinois** [Dr. Ephantus Muturi]

June 2012-August 2012 **Summer Intern, Field Museum, Chicago, Illinois** [Rebekah Baquiran]

Publications

Thairu, M. W., I. H. Skidmore, R. Bansal, E. Nováková, **T. E. Hansen**, H. Li-Byarlay, S. A. Wickline, and A. K. Hansen. "Efficacy of RNA interference knockdown using aerosolized short interfering RNAs bound to nanoparticles in three diverse aphid species." *Insect Molecular Biology* 26.3 (2017): 356-368.

Presentations:

Hansen, T.E. and J.A. White. 2017**.** Identifying Mechanisms of Host Plant Specialization in *Aphis craccivora* and its Bacterial Symbionts. ESA, Denver, CO. MS student competition.

Hansen, T.E. and J.A. White. 2017**.** Identifying Mechanisms of Host Plant Specialization in *Aphis craccivora* and its Bacterial Symbionts. NCB, Indianapolis, IN. Third place, MS student competition.

Thairu, M. W., I. H. Skidmore, R. Bansal, E. Nováková, **T. E. Hansen**, H. Li-Byarlay, S. A. Wickline, and A. K. Hansen. 2016. Aerosolized siRNAs and Nanoparticles successfully knocks down Gene Expression of BCAT in the Soybean Aphid. IAGC, Orlando, FL.

Hansen, T.E. and J.A. White. 2016**.** Transcriptomic Approaches to Elucidate Symbiont-mediated Mechanisms of Host Plant Specialization in an Aphid Pest. NCB, Cleveland, OH.

Hansen, T.E. and J.A. White. 2016**.** Transcriptomic Approaches to Elucidate Symbiont-mediated Mechanisms of Host Plant Specialization in an Aphid Pest. The CEEB, Lexington, KY.

Hansen, T.E. and A.K. Hansen (No relation). 2015**.** Who's the boss?: Clarification of which partner controls amino acid production in the aphid-Buchnera aphidicola symbiosis. OVEA Lexington, KY.

Awards and Achievements:

Third Place- Student Oral Paper Competition; Entomological Society of America-North Central Branch, Indianapolis, IN (June 2017)

North Central Branch Student Travel Award; Indianapolis, IN (June 2017)

North Central Branch Student Travel Award; Indianapolis, IN (June 2016)

Robert H. Davis Memorial Fund-Integrative Biology Department Urban-Champaign, IL (May 2014)