DIVERSITY AND FUNCTION OF A PHAGE-INFECTED BACTERIAL SYMBIONT, *HAMILTONELLA DEFENSA,* FROM PEA APHIDS

by

JAYCE WILLIAM BRANDT

(Under the Direction of Michael R. Strand)

ABSTRACT

Insect-microbe symbioses are common in nature. Generally, these symbioses can be categorized into two groups: obligate (primary) symbioses and facultative (secondary) symbioses. The pea aphid, *Acyrthosiphon pisum*, has become a model system for studying bacterial symbionts because it hosts not only a primary nutritional symbiont, *Buchnera aphidicola*, but it can also play host to any number of 8 other facultative symbionts. One such symbiont, *Hamiltonella defensa*, provides its aphid host with protection against the parasitoid wasp *Aphidius ervi*. Field studies have also found that there is variation in the level of protection that *H. defensa* provides. This variation has been attributed to differential infection by a bacteriophage named APSE. Different haplotypes of APSE are responsible for different levels of protection while uninfected *H. defensa* provide no resistance to their aphid hosts. APSE haplotypes vary little from one another in gene content outside of a variable region that contains putative eukaryotic toxin genes. This has led to the hypothesis that protection is due to expression of these phage borne toxin genes that kills the developing wasp larvae in the aphid hemocoel. However, there has been little experimental evidence directly linking toxin genes from APSE to *A. ervi* death. Furthermore, there has been little work describing how APSE persists within *H. defensa* and the impact infection has outside of killing *A. ervi* within the aphid host. This is largely due to the intractability of bacterial symbionts and the complexity of this tripartite symbiosis. Therefore, my work has been to address these gaps in our knowledge through the creation of an in vitro system and functional bioassays in which *H. defensa* and APSE can be grown outside of their aphid hosts.

INDEX WORDS: insects, symbiosis, heritable bacteria, protection, natural enemies, bacteriophage

DIVERSITY AND FUNCTION OF A PHAGE-INFECTED BACTERIAL SYMBIONT, *HAMILTONELLA DEFENSA,* FROM PEA APHIDS

by

JAYCE WILLIAM BRANDT

B.S., Kansas State University, 2012

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

DIVERSITY AND FUNCTION OF A PHAGE-INFECTED BACTERIAL SYMBIONT, *HAMILTONELLA DEFENSA,* FROM PEA APHIDS

by

JAYCE WILLIAM BRANDT

Major Professor: Michael R. Strand

Committee: Kerry M. Oliver

Mark R. Brown

Electronic Version Approved:

Ron Walcott

Interim Dean of the Graduate School

The University of Georgia

December 2019

DEDICATION

To my family, without you this would have been impossible.

ACKNOWLEDGEMENTS

This work was only made possible by the many people who have been with me throughout. First I must thank my advisor Mike Strand for his guidance, mentorship, and most of all his patience. Our many long conversations, while not always pleasant, were instrumental in my growth as a scientist and as a person. I would also like to thank Kerry Oliver who acted as a sounding board and always helped me to keep everything in perspective. I could, likewise, not have completed this research without my final committee member Mark Brown whose enthusiasm and conversation were instrumental. I also thank the department of entomology for providing an encouraging and supportive learning environment and providing me with teaching opportunities to support my research.

I am indebted to my co-authors Germain Chevignon, Jai Eum, and Brett Boyd who contributed greatly to all my work. I would especially like to thank Germain Chevignon whose eye for detail and unwavering standards did so much to enhance the project. I am lucky to count him as both colleague and a friend. I would also like to take this opportunity to thank a few friends. First and foremost, I would like to thank Kerri Coon for her support and unwavering confidence in me. She was always there for me in the darkest times, both literally and figuratively. Whether it was accompanying me during late nights collecting time points or motivating me in the early mornings with one of her many smile-inducing jingles. I would also like to Clesson Higashi for providing me with great conversation when I needed a little distraction or an ear for venting. Finally, I would like to thank the many other students, staff, and faculty that I interacted with and provided me with so many different things that I couldn't possibly list them all. This work was funded by NSF, USDA, and the Department of Entomology.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Insect-microbe interactions are ubiquitous in nature. Interaction can range from beneficial symbioses to detrimental parasites (Herren and Lemaitre 2012). Beneficial symbioses are of particular interest because they can lead to niche expansion of their host insect. Insect microbial symbionts can be broken up into two groups: primary (obligate) symbionts that are required for host development and reproduction or secondary (facultative) symbionts that are not required but provide their host with important ecological traits (Douglas 1998, Moran et al. 2008, Oliver et al. 2010, Oliver & Martinez 2014). The pea aphid (Acyrthosiphon pisum) has become a model for insect-microbe symbioses because they play host to a primary symbiont, *Buchnera aphidicola*, as well was any number of several secondary symbionts (Moran et al. 2005, Brisson & Stern 2006). Secondary symbionts have been shown to provide the aphid with defense against natural predators

and fungal pathogens, heat tolerance, and facilitating host plant usage (Montllor et al. 2002, Oliver et al. 2003, Scarborough et al. 2005, Lukasik et al. 2013, Asplen et al. 2014, Heyworth & Ferrari 2015). One of the most commonly found symbionts in the field is *Hamiltonella defensa* that provides the aphid with protection against the parasitoid wasp *Aphidius ervi* (Russell 2013, Zytenska 2016). Levels of protection can vary by strain of *H. defensa* ranging from complete protection, moderate protection, to no protection at all. This variation is due to differential infection by haplotypes of bacteriophage APSE (Oliver et al. 2009). Currently, there are four haplotypes found in *H. defensa* occurring in North American pea aphids: APSE1, APSE3, APSE2, and APSE8 (Russell et al. 2013, Smith et al. 2015). APSE3 provides complete protection while haplotypes APSE2/APSE8 provides moderate levels of protection, and *H. defensa* that are uninfected with APSE provide no protection. APSE haplotypes are nearly identical only differing in a 3.7-5.8 kb region known as the toxin cassette. Interestingly, levels of protection align with the variable toxins found in the different haplotypes. The most protective haplotype, APSE3, carries a YDrepeat toxin gene (YDr) while the moderately protective APSE2 and APSE8 both contain a cytolethal distending toxin gene (cdtB) (Degnan & Moran 2008, Oliver et al. 2009). There is however, very little known about the underlying mechanisms responsible for killing *A. ervi* within the aphid. Previous work suggests that different APSE haplotypes not only vary in how effective they are at killing *A. ervi* but also the time at which they kill the developing parasitoid. APSE3 infected *H. defensa* kill *A. ervi* early on in development, within 48 hours post parasitism. APSE2/APSE8 infected *H. defensa* kill *A. ervi* over a greater span of their life cycle, anywhere from 48 hours (first instar larvae) to 1-week post parasitism (fourth instar larvae) (Martinez et al. 2014). Given these

differences the leading hypothesis is that the different phage-encoded toxins are directly responsible for *A. ervi* death. However, it has been exceedingly difficult to directly test this hypothesis given the complexity of this symbiosis and the relative intractability of bacterial symbionts and their Hemipteran hosts. There have also been few successes in rearing parasitoid wasps to adulthood *in vitro* (Thompson 1999, Grenier et al. 2012). While there have been few successes in culturing heritable symbionts outside their natural hosts there has been greater success using insect cell lines co-cultured with symbionts (Hypsa & Dale 1997, Darby et al. 2005, Matthew et al. 2005, Pontes & Dale 2006). There have also been previous successes in rearing parasitoid wasps in vitro to life stages preceding adults even if adults were not produced or produced in ways that were economically infeasible for mass rearing as biological control agents (Thompson 1999). Therefore, one of the major goals of this research was to create a system tractable to manipulations that could test the underlying mechanisms of bacteriophage-*H. defensa* mediated protection.

In addition to the lack of experimental evidence for the mechanism for killing *A. ervi* there is also a lack of clear information about the ways in which these bacteriophage and bacteria modulate the symbiosis. One of the many things we do not know about APSE and *H. defensa* is the state in which bacteriophage persist and spread through populations of *H. defensa* and aphids. Currently there has only been one study which showed that APSE creates infectious particles, but the authors were unable to demonstrate its ability to infect naïve *H. defensa* as they lacked uninfected bacteria to work with (van der Wilk et al. 1999). However, they were able to identify and visualize viral particles in aphid hemolymph. Another study by Weldon et al. (2013) showed that bacteria that had lost their APSE infection were detrimental to aphid fitness and grew to higher

densities in aphids. Taken together it is likely that some bacteriophage within the population are undergoing lytic replication and could help keep populations of bacteria from reaching titers which are detrimental to their hosts' health. Horizontally transmitted bacteria such as *H. defensa* go through a severe bottlenecking event during the course of bacterial invasion during embryogenesis within the aphid and this could severely impact the population structures of aphids out in the field. Past studies have also found that aphids can be infected with populations of *H. defensa* that are heterogenous. This coupled with the findings that horizontal infection also occurs at very low rates in the field have an influence on the level and distribution of symbionts. What we currently don't know is how infection with APSE occurs within the aphid or within a given *H. defensa* population and how competition between these different strains occurs and the outcomes. Is it more beneficial to be infected with a highly protective yet seemingly more unstable and detrimental APSE or will strains of APSE that are largely integrated and stable within the *H. defensa* genome outcompete them for aphid resources? Do different strains of APSE act differently in terms of infection status and impact on *H. defensa* fitness? These are only a few questions that have yet to be addressed that can be important when considering how complex symbioses such as this are maintained.

While there are some major holes in our current knowledge about APSE and *H. defensa* we do have some insights from complete genomes of both *H. defensa* and several haplotypes of APSE. *H. defensa* is a gamma-proteobacteria that is closely related to free-living bacteria such as *E. coli* (Moran et al. 2005, Degnan and Moran 2008). Given its facultative symbiont status it has a robust and fairly complete genome when compared to the small and largely degraded genomes of primary obligate bacterial symbionts. *H. defensa* has the machinery to complete most necessary

metabolic processes with few pieces missing that are complemented by aphid and *B. aphidicola* products (Degnan & Moran 2008). However, only one strain of *H. defensa* had been fully sequenced along with its associated APSE haplotype. Only partial genomes were available for other APSE haplotypes so comparative studies were limited in their scope. These studies focused almost entirely on the differences found in predicted genes within a toxin cassette. One of the constraints of these studies was their inability to collect enough material to perform transcriptomics and compare differences in expression between bacteria strains and APSE haplotypes. In fact, there has been no experimental evidence that putative toxins are even expressed let alone responsible for defense against *A. ervi*. Furthermore, studies focus almost entirely on the ecological impact of APSE infection, but we lack basic information about APSE such as how it replicates and persists within *H. defensa.* To date, only one APSE haplotype (APSE1) has even been isolated from *H. defensa* (van der Wilk et al. 1999). This studied, which was the first to identify APSE in *H. defensa,* provided genomic information and micrograph images to place APSE in the family *Podoviridae*. Since this landmark study many more APSE haplotypes have been discovered with major differences in their ecological impact. Surprisingly, there has been no work comparing haplotype differences in expression, replication, and impact on *H. defensa* populations.

Given this information and holes in the literature we set out to fill some of these gaps in the literature. Namely, we wanted to experimentally test the hypothesis that putative toxin genes were responsible for protection and better understand phage biology and differences between haplotypes beyond different levels of protection. In my first chapter I show that *H. defensa* is amenable to culturing outside of its aphid host and possibly open for genetic manipulation

thereafter. We use these newly created bioassays to show that *H. defensa* and APSE are directly responsible for *A. ervi* mortality. In chapter two we present complete genomes of several APSE haplotypes and show that the genomes of APSE were largely identical outside of the toxin cassettes. Thus, APSEs were classified by which toxin cassette they carry along with their associated *H. defensa* host organism. We selected the 3 major clades that have been described in North American pea aphid *H. defensa* to show major differences in how APSE and *H. defensa* strains persist and interact with one another. In my final chapter we use this information and our newly created *in vitro* system for growing *H. defensa* and *A. ervi* outside of aphids to provide some of the first experimental evidence for the mechanism of killing *A. ervi.*

References

Herren, J. & Lemaitre, B. 2012. Insect–microbe interactions: the good, the bad and the others. *Current Opinion in Microbiology*. **15**, 217-219. (doi: 10.1016/j.mib.2012.05.014)

Douglas, A. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera. Annual Review of Entomology*. **43**, 17-37. (doi: 10.1146/annurev.ento.43.1.17).

Moran, N.A., McCutcheon, J.P., Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. *Annual Reviews of Genetics*. **42**, 165-190. (doi: Annurev.Genet.41.110306.130119).

Oliver, K.M., Degnan, P.H., Burke, G.R., Moran, N.A. 2010. Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annual Review of Entomology*. **55**, 247-266. (doi: 10.1146/annurev-ento-112408-085305).

Oliver, K.M. & Martinez, A.J. 2014. How resident microbes modulate ecologically-important traits of insects. *Current Opinion in Insect Science*. **4**, 1-7. (doi:10.1016/j.cois.2014.08.001)

Moran, N.A., Russell, J.A., Koga, R., Fukatsu, T. 2005. Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Applied and Environmental Microbiology*. **71**, 3302-3310. (doi:10.1128/AEM.71.6.3302-3310.2005).

Brisson, J.A. & Stern, D.L. 2006. The pea aphid, *Acyrthosiphon pisum*: an emerging genomic model system for ecological, developmental, and evolutionary studies. *Bioessays*. **28**, 747-755. (doi:10.1002/bies.20436).

Montllor, C.B., Maxmen, A., Purcell, A.H. 2002. Facultative bacterial endosymbionts benefit pea aphids, *Acyrthosiphon pisum*, under heat stress. *Ecological Entomology*. **27**, 189-195. (doi:10.1046/j.1365-2311.2002.00393.x).

Oliver, K.M., Russell, J.A., Moran, N.A., Hunter, M.S. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 1803-1807. (doi:10.1073/pnas.0335320100).

Scarborough, C.L., Ferrari, J., Godfray, H.C.J. 2005. Aphid protected from pathogen by endosymbiont. *Science*. **310**, 1781. (doi:10.1112/science.1120180).

Lukasic, P., van Asch, M., Guo, H., Ferrari, J., Godfray, H.C.J. 2013. Unrelated facultative endosymbionts protect aphids against a fungal pathogen. *Ecology Letters*. **16**, 214-218. (doi:10.111/ele.12031).

Asplen, M.K., Bano, N., Brady, C.M., Desneux, N., Hopper, K.R., Malouines, C., Oliver, K.M., White, J.A., Heimpel, G.E. 2014. Specialisation of bacterial endosymbionts that protect aphids from parasitoids. *Ecological Entomology*. **39**, 736-739. (doi:10.1111/een.12153).

Heyworth, E. & Ferrari, J. 2015. A facultative endosymbiont in aphids can provide diverse ecological benefits. *Journal of Evolutionary Biology*. **28**, 1753-1760. (doi:10.1111/jeb.12705).

Russell, J.A., Weldon, S., Smith, A.H., Kim, K.L., Hu, Y., Lukasik, P., Doll, S., Anastopoulos, I., Novin, M., Oliver, K.M. 2013. Uncovering symbiont-driven genetic diversity across North American pea aphids. *Molecular Ecology*. **22**, 2045-2059. (doi:10.1111/mec.12211)

Zytynska, S.E. & Weisser, W.W. 2016. The natural occurance of secondary bacterial symbionts in aphids. *Ecological Entomology*. **41**, 13-26. (doi:10.1111/een.12281).

Oliver, K.M., Degnan, P.H., Hunter, M.S., Moran, N.A. 2009. Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science*. **325**, 992-994. (doi:10.1126/science.1174463).

Smith, A.H., Lukasik, P., O'Connor, M.P., Lee, A., Mayo, G., Drott, M.T., Doll, S., Tuttle, R., Disciullo, R.A., Messina, A., Russell, J.A. 2015. Patterns, causes, and consequences of defensive microbiome dynamics across multiple scales. *Molecular Ecology*. **24**, 1135-1149. (doi:10.1111/mec.13095).

Degnan, P. & Moran, N.A. 2008. Evolutionary genetics of a defensive facultative symbiont of insects: exchange of toxin-encoding bacteriophage. *Molecular Ecology*. **17**, 916-929. (doi:10.1111/j.1365-294X.2007.03616.x).

Martinez, A.J., Weldon, S.R., Oliver, K.M. 2014. Effects of parasitism on aphid nutritional and protective symbioses. *Molecular Ecology*. **23**, 1594-1607. (doi:10.1111/mec.12550).

Thompson. S.N. 1999. Nutrition and culture of entomophagous insects. *Annual Review of Entomology*. **44**, 561-592. (doi:10.1146/annurev.ento.44.1.561).

Grenier, S. 2012. Artificial rearing of entomophagous insects, with emphasis on nutrition and parasitoids – general outlines from personal experience. *Karaelmas Science and Engineering Journal*. **2**, 1-12.

Hypsa, V. & Dale, C. 1997. In vitro culture and phylogenetic analysis of "*Candidatus Arsenophonus triatominarum"* an intracellar bacterium from the triatomine bug, *Triatoma infestans*. *International Journal of Systemic Bacteriology*. **47**, 1140-1144. (doi:10.1099/00207713-47-4-1140).

Darby, A.C., Chandler, S.M., Welburn, S.C., Douglas A.E. 2005. Aphid-symbiotic bacteria cultured in insect cell lines. *Applied and Environmental Microbiology*. **71**, 4833-4839. (doi:10.1128/aem.71.8.4833-4839.2005).

Matthew, C.Z., Darby, A.C., Young, S.A., Hume, L.H., Welburn, S.C. 2005. The rapid isolation and growth dynamics of the tsetse symbiont *Sodalis glossinidius*. *Fems Microbiology Letters*. **248**, 69-74. (doi:10.1016/j.femsle.2005.05.024).

Pontes, M.H. & Dale, C. 2006. Culture and manipulation of insect facultative symbionts. *Trends in Microbiology*. **14**, 406-412. (doi:10.1016/j.tim.2006.07.004).

van der Wilk, F., Dullemans, A.M., Verbeek, M., van den Heuvel, J. 1999. Isolation and characterization of APSE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrthosiphon pisum*. *Virology*. **262**, 104-113. (doi:10.1006/viro.1999.9902).

Weldon, S.R., Strand, M.R., Oliver, K.M. 2013. Phage loss and the breakdown of a defensive symbiosis in aphids. *Proceedings of the Royal Society of London B: Biological Sciences*. **280**, 10122103. (doi:10.1098/rspb.2012.2103).

CHAPTER 2

CULTURE OF AN APHID SYMBIONT DEMONSTRATES ITS DIRECT ROLE IN DEFENSE AGAINST PARASITOIDS[1]

[1] Brandt, J.W., Chevignon G., Oliver, K.M., Strand, M.R. 2017. *Proceedings of the Royal Society B*. 284(1866):20171925.
Reprinted here with permission of the publisher

2.1 Abstract

Heritable symbionts are common in insects with many contributing to host defense. *Hamiltonella defensa* is a facultative, bacterial symbiont of the pea aphid, *Acyrthosiphon pisum* that provides protection against the endoparasitoid wasp *Aphidius ervi*. Protection levels vary among strains of *H. defensa* that are differentially infected by bacteriophages named APSEs. In contrast, little is known about mechanism(s) of resistance due to the intractability of host-restricted microbes for functional study. Here, we developed methods for culturing strains of *H. defensa* that varied in the presence and type of APSE. Most *H. defensa* strains proliferated at 27°C in co-cultures with the TN5 cell line or as pure cultures with no insect cells. The strain infected by APSE3, which provides high levels of protection in vivo, produced a soluble factor(s) that disabled development of *A. ervi* embryos independent of any aphid factors. Experimental transfer of APSE3 also conferred the ability to disable *A. ervi* development to a phage-free strain of *H. defensa*. Altogether, these results provide a critical foundation for characterizing symbiont derived factor(s) involved in host protection and other functions. Our results also demonstrate that phage-mediated transfer of traits provides a mechanism for innovation in host restricted symbionts.

Keywords: bacteria, endosymbiont, mutualism, mobile element, horizontal transfer

2.2. Introduction

Insects, including species of medical and agricultural importance, frequently host maternally transmitted bacterial symbionts [1]. By providing novel genes that expand the functional repertoire

of the host insect, heritable symbionts often play key roles in mediating host ecology [2, 3]. Phloem-feeding aphids (Hemiptera: Aphidoidea), for example, depend on the obligate (=primary) symbiont *Buchnera aphidicola,* which lives in specialized cells called bacteriocytes and supplies nutrients deficient in the aphid diet [4]. Aphids also commonly host one or more species of facultative (=secondary) symbionts that can live either extracellularly in the hemocoel or intracellularly in bacteriocytes or neighbouring sheath cells [5, 6]. Facultative symbionts are known to provide several conditional benefits to hosts including resistance to parasitoid wasps and pathogens, protection against thermal stress, and facilitation of host plant usage [7-16]. Some facultative symbionts are also infected by bacteriophages [17-19] that can enhance host fitness by provisioning novel traits and/or altering within-host bacterial abundance [18, 20, 21].

The γ-proteobacterium *Hamiltonella defensa* is among the most common aphid facultative symbionts, occurring in 34% of sampled species [22, 23]. *H. defensa* confers protection against parasitoids in at least three aphid species, including the pea aphid, *Acyrthosiphon pisum* [8, 14, 24]. Field and laboratory surveys identify multiple strains of *H. defensa* in *A. pisum* with most persistently infected by a podovirus-like bacteriophage named APSE [25-27]. Four APSE variants (APSE1, 2, 3 & 8) are currently known. The viral genome of each shares strong conservation in regulatory and structural genes but varies in a 3.7-5.8 kb domain containing predicted eukaryotic-like toxin genes [18, 19, 28, 29]. A major mortality agent of *A. pisum* is the endoparasitoid wasp *Aphidius ervi* [30-33]. In examined strains, *H. defensa* confers protection against this wasp only when infected with APSE [34]. In APSE3-associated strains, the spontaneous loss of phage eliminates a protective phenotype and increases titres of *H. defensa*, which adversely affect aphid

fitness [20, 21]. Protection levels against *A. ervi* and the timing of wasp mortality in aphids differ among *H. defensa* strains and their associated APSE variants [20, 24, 29, 35-37]. Recent results indicate some genotypes of *A. pisum* exhibit resistance toward *A. ervi* in the absence of *H. defensa* or other facultative symbionts [38]. Endogenous and symbiont defenses also target only specific parasitoid species [29, 39]. These findings collectively suggest aphid defense against parasitoids is multifactorial with APSE-encoded toxins or other gene products interacting additively or synergistically with factors produced by aphids.

Asexual reproduction and the ability to manipulate symbiont infections make aphids excellent models for studying the phenotypic effects of facultative symbionts [6, 7]. However, in vivo studies have provided few insights into the mechanisms underlying symbiont-mediated phenotypes. The ability to cultivate normally host-restricted symbionts and their bacteriophages outside of aphids and other insects would open new avenues for studying function. While most heritable symbionts of insects remain unculturable, a few species have been propagated in co-cultures with insect cells [40-44] or as pure cultures [43, 45]. The ~2.1-Mb genome of *H. defensa* is significantly smaller than its closest free-living relatives and lacks a number of genes with metabolic functions [46]. Nonetheless, Darby et al. [40] reported that one strain of *H. defensa* (called T type in this study) from black bean aphids (*Aphis fabae*) and a strain of the closely related species *Regiella insecticola* (U-type) from *A. pisum* persisted in vitro when co-cultured with the C6/36 cell line from the mosquito *Aedes albopictus* or the S2 cell line from *Drosophila melanogaster*. The growth and abundance of *H. defensa* and *R. insecticola* in these cultures were not quantified but descriptions of their maintenance suggest both primarily persisted intracellularly

at low abundance [40]. In contrast, *H. defensa* grows to high titres as *A. pisum* develop from first instar nymphs to adults [21, 36].

Here, we report the in vitro culture of four *H. defensa* strains that exhibit variable protection against parasitoids in vivo owing to differences in APSE infections (Table 2.1). We then demonstrate that one strain of APSE-infected *H. defensa* kills *A. ervi* in the absence of any aphid factor, and that horizontal transfer of this APSE to a previously phage-free strain of *H. defensa* confers the same killing activity.

TABLE 2.1 Characteristics of *H. defensa* strains cultured.

H. defensa strain	APSE variant	Primary phage toxin	Protection in aphids
A2C	no phage	no phage	none
NY26	APSE2	CdtB	moderate
AS3	APSE3	YDR	high
ZA17	APSE8	CdtB	moderate

2.3 Methods

(a) Insects and cell lines. Clonal *A. pisum* lines carrying the A2C, AS3, NY26 and ZA17 strains of *H. defensa* were maintained on fava bean plants at 20° C and a 16 h light: 8 h dark photoperiod as previously described [20, 35-37]. *A. ervi* used in the study were reared in the laboratory as a large population founded from commercially purchased (Arbico) and field-collected wasps on an

A. pisum line with no facultative symbionts; adult wasps were provided continuous access to honey and water at 20° C under a 16 light: 8 h dark photoperiod [36]. Culture media and insect cell lines used in the study are summarized in Table S1.

(b) Establishment of *H. defensa* cultures. *A. pisum* adults containing the A2C, AS3, NY26 or ZA17 strain of *H. defensa* were surface sterilized by submerging in 5% bleach, 1% Tween20, and 0.1% ROCCAL-D (Pfizer) (v/v in water) for 30 sec followed by rinsing in sterile water for 30 sec. Each aphid was then transferred to a 25 µl drop of medium and pierced using sterile forceps to release haemolymph from the body cavity into the medium. Each drop was then transferred to an individual culture well in a 24 well culture plate (Corning) containing 1 ml of medium with or without fetal bovine serum (FBS) and particular insect cell lines (Table S2.1). Primary cultures were maintained under ambient atmosphere at 20° or 27° C. Established cultures were passaged weekly or biweekly by adding 1 x 10^6 bacteria to 25 cm culture flasks (Falcon) containing 4.0 ml of TC100 medium plus 10% FBS with or without 1 x 10^6 TN5 cells (Table S2.1).

(c) Estimating *H. defensa* titres. Cohorts (n=5) of two-day old adult aphids infected with the A2C, AS3, NY26 or ZA17 strains of *H. defensa* were placed in 500 µl of phosphate buffered saline (PBS). The cuticle of each aphid was then pierced using forceps as described above. After removing the aphids, each drop was transferred to 1.5 ml microfuge tube and gently centrifuged at 50 x *g* to pellet aphid cells and debris. The supernatant was then transferred to a new tube and spun at 8000 x *g* for 10 min at 4° C to pellet *H. defensa*. After decanting the supernatant, DNA was extracted as described [21] and stored at -20°C. DNA was similarly isolated from *H. defensa*

in established cultures by collecting 500 µl of medium 3, 6, 9 or 15 days after new cultures were passaged. *H. defensa* titres were estimated by qPCR using specific primers for the single-copy *H. defensa dnaK* gene (Table S2.2) and established methods [21, 35, 47]. Ten µl reactions contained 0.5 µl of 5 µM of each primer, 5 µl of the Rotor-Gene SYBR Green Master Mix and 1 µl of a 1:100 dilution of template and 3 µl of ultrapure water. Four technical replicates were run per sample using a Rotor-Gene Q cycler (Qiagen). Data were acquired during the extension step, and analyzed with the Rotor-Gene application software.

(d) Diagnostic PCR. Diagnostic PCR reactions were run using specific primers that amplified unique domains in: 1) the *dnaK* gene present in all strains of *H. defensa*, 2) the head protein gene (*p24*) present in all APSE variants, 3) the *ydr* gene in APSE3 that persistently infects AS3, 4) the *cdtB* gene in APSE2 and APSE8 that persistently infect NY26 and ZA17 respectively, 5) a hypothetical protein unique to ZA17, 6) the RTX-toxin gene of APSE8 that is unique to NY26, and 7) the *dnaK* gene of *B. aphidicola* (Table S2.2). PCR reactions (25 µl) contained 2.5 µl of 25 mM Mg^{2+}, 0.5 µl of mixed dNTPs (2.5 µM each), 0.5 µl of each primer (5 µM), 1 µl of a 1/100 dilution of template DNA, 0.2 µl of HotMaster Taq DNA Polymerase, and 19.6 µl of ultrapure water. Samples were run using a T100 thermal cycler (Biorad). Eight µl of each PCR reaction were then run on a 1% agar gel stained with ethidium bromide and visualized using a BioDoc-iT imaging system.

(e) Microscopy. *H. defensa* were slide mounted and examined with a Leica DMXRXE upright epifluorescent microscope with differential interference optics and a Leica digital camera. Established cultures in flasks containing TN5 cells were also examined after staining with 20 µg/ml acridine orange (Sigma) or 1 µg/ml Hoechst 33342 (Sigma) in water using a Leica DM IRE2 inverted epifluorescent microscope with phase-contrast optics fitted with a Hamamatsu digital camera. *A. ervi* eggs were stained with acridine orange and 1 µg/ml propidium iodide (Sigma) followed by examination as described above. Images were captured using Leica or SimplePCI software (Compix). For transmission electron microscopy (TEM), *H. defensa* and TN5 cells from established cultures were pelleted 4 days post-passage by centrifugation at 2000 x *g* at 4°C. After decanting the supernatant the pellet consisting of bacteria and insect cells was fixed for 2 h in in 0.1 M cacodylate-HCl buffer (pH 7.0) containing 2% glutaraldehyde (v/v), 2% paraformaldehyde (v/v) and 0.2% picric acid (v/v). After rinsing in 0.1 M cacodylate-HCl buffer and repelleting, cells were resuspended in 2% agar at 50° C followed by centrifugation for 10 min and storage at 4°C until the agar hardened. Agar-embedded samples were then post-fixed in 1% OsO_4 in 0.1 M cacodylate-HCl buffer for 1 h, rinsed in deionized water, and dehydrated in graded ethanol (30-100%) with 15 min wash steps. Samples were then rinsed 1 x in acetone and 2x in propylene oxide for 15 min before infiltrating in propylene oxide: Epon-Araldite (1:1, 1:3) for 2 h. Samples were embedded in 100% Epon-Araldite, which polymerized over 24 h in a 75° C oven. Sections were cut with a Reichert Ultracut S ultramicrotome and stained with uranyl acetate and lead citrate. Sections were then examined and photographed using a JEOL JEM1011 (JEOL, Inc., Peabody, MA) transmission electron microscope at 100kV.

(f) *A. ervi* assays. *A. pisum* without facultative symbionts were individually parasitized by *A. ervi.* Aphids were surface sterilized as previously described 2 or 36 h post-parasitism and then dissected to isolate either newly laid *A. ervi* eggs or 1st instar larvae. One egg or larva was placed per well of a 96-well culture plate that contained 100 µl TC100 + 30% FBS. A2C, AS3 or ZA17 from 3-day-old established cultures were then added to wells at a density of $1x10^4$ to $1x10^6$ bacteria per ml. *A. ervi* eggs were allowed to develop for 72 h at 27°C while first instars were maintained for 48 h. Embryonic stages of *A. ervi* were recorded every 24 h.

(g) APSE infectivity experiments. One ml of medium from 3-day old established cultures of the AS3 strain was collected and centrifuged at 8000 x *g.* Cell pellet template DNA was extracted as described above. The supernatant was collected and filtered through an 0.2 µm syringe filter to eliminate all bacterial cells. Two µl of TURBO DNase (Ambion) and 5 µl of buffer was then added to 50 µl of this supernatant for 40 min at 37° C. After adding EDTA (10 mM) to inactivate the DNase, supernatant template DNA was extracted [21]. APSE specific primers (Table S2.2) were used to amplify viral DNA using qPCR as described above. APSE3 copy number per ml of medium was calculated by multiplying the qPCR estimate by the dilution factor and elution volume. Two independently acquired biological replicates were performed for each treatment with samples internally replicated 4 times. A2C titre was determined by qPCR as described above. Six new A2C cultures were initiated by adding 1 x 10^7 bacteria to 4 ml of fresh TC100 + 10% FBS and 1 x 10^4 TN5 cells in a 6 well culture plate (Corning). One x 10^6 APSE3 in medium collected

from the AS3 strain was then added to A2C cultures (N=3) at a multiplicity of infection (MOI) of 0.1 while 3 other A2C cultures were grown without addition of APSE3. After nine days at 27°C APSE3 and *H. defensa* copy number per ml in the cell pellets, and APSE3 copy number in DNase pre-treated supernatants were determined by qPCR. The effects of APSE3-infected A2C on development of *A. ervi* eggs were then compared to non-infected (phage-free) A2C by adding 1 x 10^6 bacteria per well and assessing the proportion of eggs that developed to the 128-cell stage as described above.

(h) Figure assembly and statistical analyses. Light and electron microscopy images were assembled using Adobe Illustrator and Photoshop. All statistical analyses were performed using R (www.r-project.org) or JMP Pro version 11 (SAS Institute Inc., Cary, NC). Figures were generated using R or GraphPad Prism 5.0.

2.4 Results

(a) *H. defensa* grows extracellularly in vivo and in vitro. Previous estimates using qPCR indicated that titres of phage-free and APSE-infected *H. defensa* increase in *A. pisum* from ~1 x 10^7 per first instar nymph to more than ~1 x 10^9 per 2 day old adult [21]. In the current study, we determined that the titres of A2C, AS3, NY26 and ZA17 specifically in the haemolymph of 2-day old adults exceeded 1 x 10^9 bacteria (Fig. S2.3). Moreover, most if not all *H. defensa* resided in plasma as non-motile pleomorphic rods rather than haemocytes [48], which supported that the vast

majority of bacteria in aphids persist extracellularly (Fig. S2.3). We therefore used adult haemolymph as the source of *H. defensa* for establishing primary cultures. We initially focused on the phage-free A2C strain and tested several different types of insect cell culture media with our without FBS and different insect cell lines (Table S2.1). At 27° C and ambient atmosphere we observed bacteria with identical morphology to *H. defensa* in all primary cultures including those containing C6/36 and S2 cells used by Darby et al. [40]. However, after 3 days bacteria only noticeably increased in abundance in culture wells containing TC100 medium plus 10% FBS and the TN5 cell line (Table S2.1). We therefore used these conditions to initiate primary cultures for the APSE-infected AS3, NY26 and ZA17 strains (Table S2.1). Similar to A2C, AS3 and ZA17 increased in abundance under ambient atmosphere at 27° C, whereas NY26 showed no increase in abundance at 27° C but did so when maintained at 18° C. Approximately $1x10^6$ bacteria from primary cultures were transferred after 7 (A2C, AS3, ZA17) or 14 days (NY26) to 25 cm^2 cell culture flasks containing 4 ml of TC100 plus 10% FBS and 1 x 10^4 TN5 cells. Each strain was thereafter passaged weekly (A2C, AS3, ZA17) or biweekly (NY26) under the same conditions and henceforth called established cultures. Diagnostic PCR assays confirmed that each established culture contained only the strain of *H. defensa* that was present in the donor aphid (Fig. S2.4).

(b) Established cultures grow to high titre as extracellular bacteria. Titre assays indicated that each *H. defensa* strain increased to a maximum of ~1 x 10^{10} bacteria per ml in established cultures (Fig. 2.2A). TN5 cells did not increase in abundance when co-cultured with each strain *H. defensa* (Fig. 2.2B). However, almost no TN5 cells were stained by the vital dye propidium iodide, which

indicated most remained viable in the presence of *H. defensa.* TN5 cells co-cultured with each strain of *H. defensa* also remained attached to culture plates and morphologically looked identical to TN5 cells cultured in TC100 plus 10% FBS without bacteria. *H. defensa* in established cultures remained pleomorphic but commonly formed pairs or short chains of cells not observed in *A. pisum* (Fig. S2.5A-D). Staining with the DNA specific dye Hoechst 33342 indicated bacteria were predominantly extracellular and did not enter TN5 cells (Fig. S2.5E, F). We also examined established cultures of the A2C, AS3 and ZA17 strains by TEM. Ultrastructure at low magnification showed an abundance of variably shaped *H. defensa* that were extracellular (Fig. S2.6A, B). Higher magnification images also showed that the cell envelope of *H. defensa* consisted of an inner and outer membrane but lacked a symbiosome membrane [49] present in some intracellular symbionts including *Buchnera* (Fig. S2.6C). Inspection of hundreds of TN5 cells identified only one containing an *H. defensa* cell in the cytoplasm, which indicated bacteria rarely entered this cell type (Fig. S2.6D). Established cultures of *H. defensa* were stored at -80°C by adding 15% glycerol (vol/vol) to culture medium. After thawing on ice, aliquots stored for more than 1 month grew similarly to continuously propagated cultures when re-inoculated into flasks containing TC100 plus 10% FBS and TN5 cells.

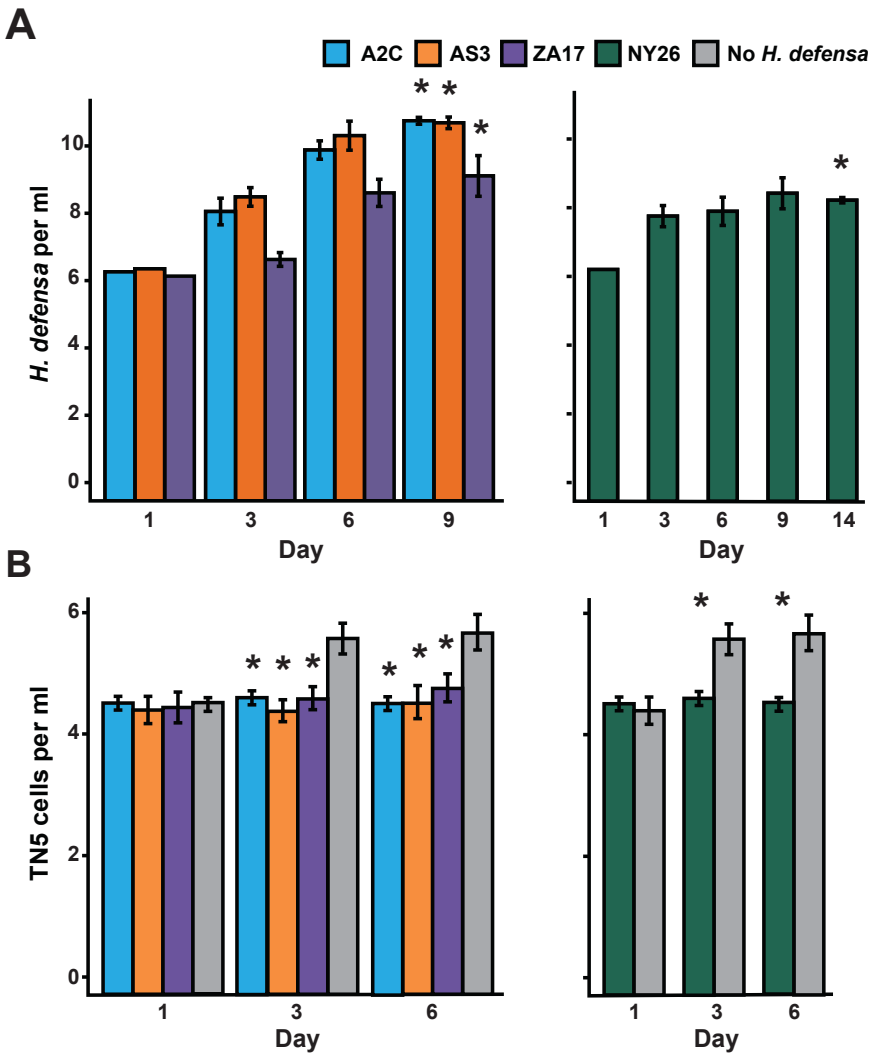


Figure 2.2. Growth of *H. defensa* and TN5 cells in established cultures of A2C, AS3, NY26 and ZA17. (A) Genome copy number (± 95% confidence intervals) per ml of medium of A2C, AS3, and ZA17 (left graph) or NY26 (right graph) estimated by qPCR over 9 or 15 days and analyzed by repeated measures ANOVA. Asterisks indicate that titers are significantly more abundant compared to day 1 (pairwise t-tests, α=0.01). (B) Growth of TN5 cells in the presence and absence of A2C, AS3, and ZA17 (left graph) or NY26 (right graph). * indicate that flasks with *H. defensa* contained significantly fewer TN5 cells than flasks with no *H. defensa* at day 3 and day 6 (ANOVA followed by a post-hoc Tukey-Kramer Honest Significant Difference test; α=0.05).

(c) Some strains of *H. defensa* grow as pure cultures in liquid medium. After one year and more than 50 passages, we tested each strain for the ability to grow in TC100 plus 10% FBS but no TN5 cells. Results showed that the A2C, AS3 and ZA17 strains grew at similar rates and to similar densities (~1 x 10^{10} per ml) at 27° C as bacteria co-cultured with TN5 cells. Each strain could also be routinely passaged and maintained as a pure culture in TC100 plus 10% FBS but would not grow in TC100 in the absence of FBS. The NY26 strain in contrast could not be cultured in the absence of TN5 cells. Given these results, we reassessed whether primary cultures of A2C, AS3 and ZA17 could be established from adult aphids by adding hemolymph to TC100 plus 10% FBS and no TN5 cells. Repeated efforts failed to generate any primary cultures that grew. Attempts to propagate *H. defensa* from established cultures on TC100 plus 10% FBS agar plates under ambient or a microaerobic atmosphere also failed to produce colonies for any strain.

(d) AS3 disables development of *Aphidius ervi*. Similar to symbionts, most endoparasitoid wasps require specialized conditions for development that make in vitro culture difficult [50-52]. While no wasps in the genus *Aphidius* have been successfully reared to adulthood in vitro, prior studies established that *A. ervi* eggs develop to the first instar in TC100 plus 10% FBS [53]. *A. ervi* eggs have a thin chorion (egg shell) and develop in the haemolymph of aphid hosts. Eggs undergo complete cleavage to initially form a 128-cell embryo that is enveloped by an extraembryonic membrane of polar body origin within the chorion [53]. The embryo then ruptures from the chorion as a spherical morula that completes development into a first instar wasp larva [53].. Previous in vivo assays further showed that *A. ervi* eggs: 1) usually die during embryogenesis in

A. pisum containing APSE3-infected *H. defensa*, 2) exhibit lower levels of mortality that can occur before or after completion of embryogenesis in *A. pisum* containing APSE2- or APSE8-infected *H. defensa,* and 3) exhibit no mortality in *A. pisum* containing phage-free *H. defensa* [36]. We thus assessed whether established cultures of *H. defensa* affected development of *A. ervi* eggs by focusing on the APSE3-infected AS3, APSE8-infected ZA17, and phage-free A2C strains. We did not bioassay APSE2-infected NY26 because of its low temperature requirements and APSE2 contains a very similar primary toxin gene (*cdtB*) as APSE8 [36].

At a starting density of 1 x 10^6 bacteria per ml, AS3 significantly reduced the proportion of newly laid eggs that developed into 128-cell stage embryos, whereas A2C and ZA17 did not (Fig. 2A). AS3 at 1 x 10^5 per ml also significantly reduced the proportion of *A. ervi* eggs that developed to the 128-cell stage but 1 x 10^4 bacteria per ml did not (Fig. 2.3A). Epifluorescent microscopy showed that most eggs co-cultured with 1 x 10^6 AS3 per ml failed to develop beyond the 32-cell stage. Staining with propidium iodide further showed that eggs co-cultured with AS3 initially exhibited a loss of viability of cells forming the extraembryonic membrane at 24 h post-culture (Fig. 2.3B). This was followed by mortality of most cells in the extraembryonic membrane plus many embryonic cells by 48 h (Fig. 2.3B). In contrast, most embryos showed no loss of viability at 24 or 48 h when co-cultured with ZA17, A2C or no *H. defensa* (Fig. 2.3B).

We next cultured AS3 or A2C (negative control) at a starting density of 1 x 10^6 per ml for 6 h in TC100 plus 10% FBS. We then removed the bacteria and added newly laid *A. ervi* eggs to the conditioned medium. Most *A. ervi* eggs in AS3-conditioned medium failed to develop to the 128-cell stage, while most eggs in A2C-conditioned or non-conditioned (control) medium

developed to the 128-cell stage (Fig. 2.3C). To assess whether ZA17 affected *A. ervi* larvae, bacteria were co-cultured with first instars followed by assessment of larval viability after 24 and 48 h. Results revealed no reduction in viability when compared to larvae co-cultured with A2C or no bacteria (Fig. 2.3D).

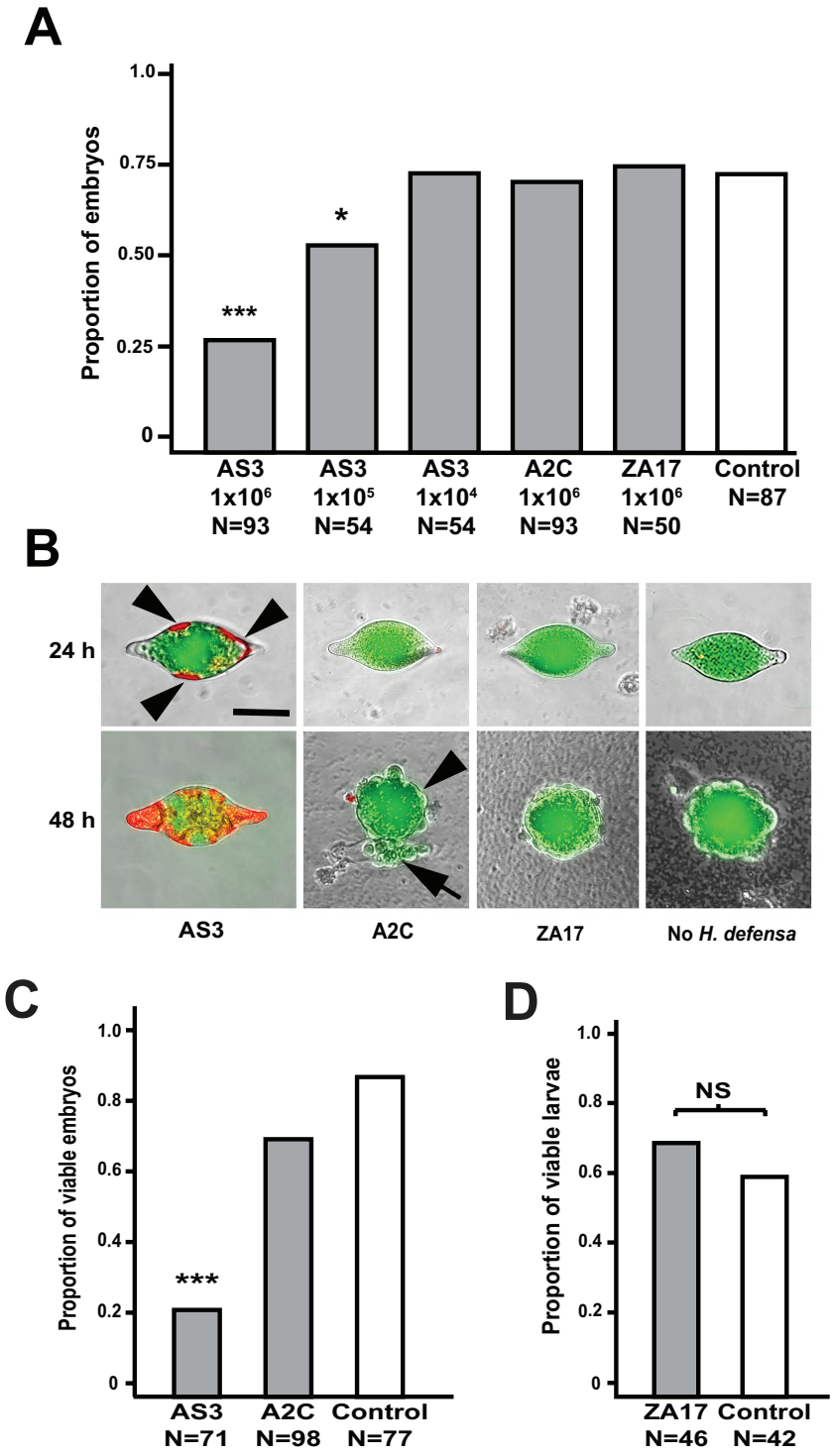

A
Proportion of embryos
1.0
0.75
0.50
0.25
0
***
*
AS3 1x10^6 N=93
AS3 1x10^5 N=54
AS3 1x10^4 N=54
A2C 1x10^6 N=93
ZA17 1x10^6 N=50
Control N=87
B
24 h
48 h
AS3
A2C
ZA17
No H. defensa
C
Proportion of viable embryos
1.0
0.8
0.6
0.4
0.2
0
***
AS3 N=71
A2C N=98
Control N=77
D
Proportion of viable larvae
1.0
0.8
0.6
0.4
0.2
0
NS
ZA17 N=46
Control N=42

Figure 2.3. The AS3 strain of *H. defensa* disrupts development of *A. ervi* during early embryogenesis. (A) Proportion of *A. ervi* eggs that developed to the 128-cell stage after 36 h when co-cultured with the AS3, A2C, or ZA17 strains with no *H. defensa* (Control). Starting densities per ml indicated on X-axis. Asterisks indicate the treatment significantly differed from the no bacteria control (X^2=62.7; df=5; p<0.0001 followed by post-hoc pairwise Fisher Exact Tests). (B) Epiflourescent micrographs of *A. ervi* eggs co-cultured with AS3, A2C, ZA17 or no bacteria and stained with propidium iodide (red) and acridine orange (green) after 24 or 48 h. Several cells in the extraembryonic membrane of eggs co-cultured with AS3 are stained by propidium iodide at 24 h (arrowheads) while many embryonic cells are stained at 48 h. Eggs cultured with A2C, ZA17 or no bacteria exhibit little or no staining by propidium iodide. The morula (arrowhead) that ruptured out of the chorion (arrow) is visible at 48 h for the image shown under A2C, while only a morula is shown at 48 h under ZA17 and No *H. defensa*. Scale bar equals 200 μm. (C) Proportion of *A. ervi* eggs that developed to the 128-cell stage at 36 h was significantly lower in medium conditioned by AS3 than medium conditioned by A2C or medium with no *H. defensa* (Control) (X^2 = 61.2, df=2, P≤0.0001 followed by post-hoc pairwise Fisher Exact Tests). (D) Proportion of viable first instar *A. ervi* after 48 h did not differ when cultured with ZA17 or no *H. defensa* (Control) (X^2 = 0, df=1, P=1).

(e) APSE3 confers anti-*A. ervi* defense to the A2C strain. Comparative genomic data indicate that all APSEs persist as integrated proviruses in *H. defensa* genomes, with data also supporting that some APSE variants replicate [19, 21, 28]. Phylogenetic evidence also suggests that APSEs have horizontally moved among *H. defensa* strains [25]. However, no experimental evidence supports that any APSE is infectious or that phage acquisition can alter the phenotype of strains in defense against parasitoids. Since APSE3-infected AS3 disabled development of *A. ervi,* we assessed whether established cultures of this strain produced APSE3 that could infect phage-free A2C. After isolating DNA from the AS3 strain in established cultures 5 days post-inoculation, qPCR analysis showed that bacterial cells contained similar copy numbers of the APSE3 and AS3 genomes, which supported that approximately one copy of the APSE3 genome was present per

bacterial cell (Fig. 2.4A). However, qPCR analysis of DNase pretreated supernatants also detected 2.7 x 10^8 copies per ml of the APSE3 genome in the medium, which supported that AS3 in culture produced virus particles (Fig. 2.4A). We therefore used this supernatant to inoculate cultures containing 1 x 10^7 A2C with APSE3 at an MOI of 0.1. Analysis 9 days post-infection detected a 10-fold higher copy number of the APSE3 versus A2C genome in DNA templates from bacterial cells, and 4 x 10^{10} copies of the APSE3 genome per ml in DNase pre-treated supernatants (Fig. 2.4B). This supported that APSE3 from the AS3 strain infected and replicated in A2C. We then compared the effects of adding 1 x 10^5 APSE3-infected A2C or non-infected A2C to 1 ml cultures containing newly laid *A. ervi* eggs. Results showed that APSE3-infected A2C reduced the proportion of *A. ervi* eggs that developed to the 128-cell stage relative to uninfected A2C (Fig. 2.4C). This reduction was also near identical to the effects of APSE3-infected AS3 (see Fig. 2.4A).

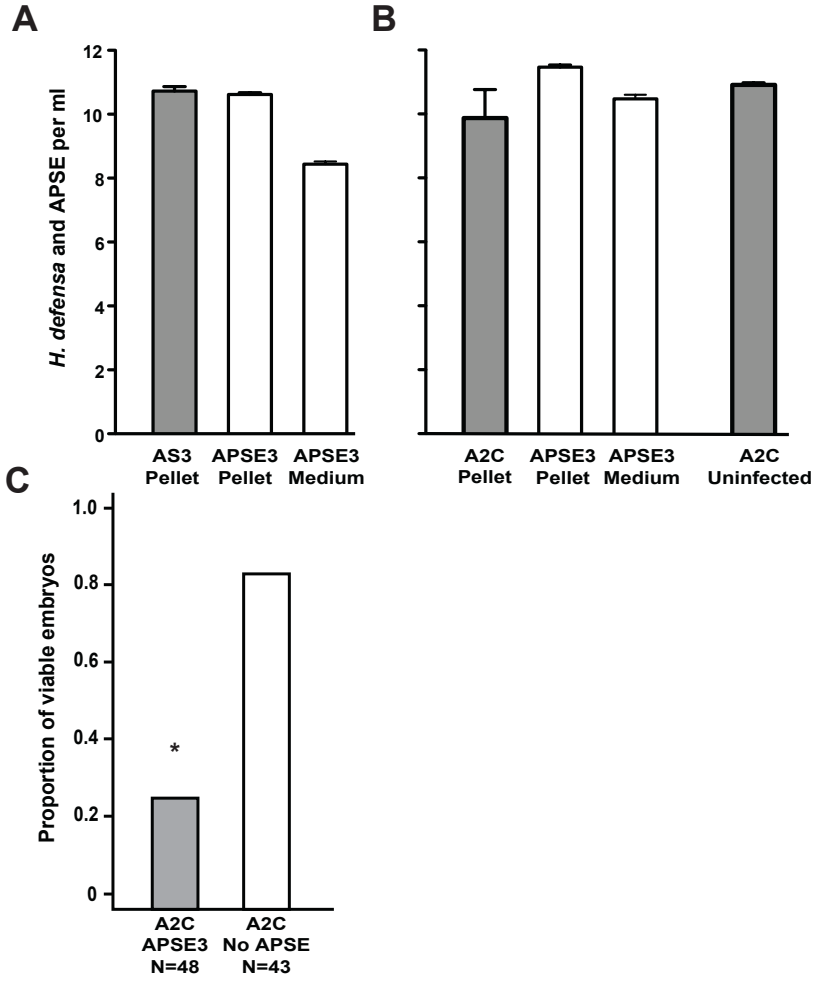


Figure 2.4. Established cultures of AS3 produce APSE3 particles that infect and alter the ability of A2C to kill *A. ervi* embryos. (A) Genome copy number (± 95% confidence intervals) of AS3 and APSE3 in bacterial pellets, and APSE3 in medium pretreated with DNase. Data generated from analysis of 1 ml of medium from two independent cultures. (B) Genome copy number (± 95% confidence intervals) of A2C and APSE3 in bacterial pellets, and APSE3 in medium pretreated with DNase. Data generated from analysis of 1 ml of medium from 3 independent cultures 9 days post-infection. (C) Proportion of *A. ervi* eggs that developed to the 128-cell stage after 36 h when co-cultured with the APSE3-infected A2C or non-infected A2C. Both were bioassayed at a starting density of 1 x 10^5 per ml. The asterisk above the bar for APSE3-infected A2C indicates this treatment significantly differed from A2C control (X^2=66.5; df=1; p<0.0001)

2.5 Discussion

No obligate and only a few facultative heritable symbionts of insects have been cultured outside of hosts. One difficulty in culturing these organisms is that heritable symbionts are generally restricted to host tissues and genome reduction results in many species having specialized requirements for growth that are difficult to mimic in vitro [43]. Another challenge is that the slow growth of many heritable symbionts elevates the risk of contamination when trying to establish primary cultures [43]. Most successes in culturing heritable insect symbionts have used insect cells that could serve as hosts for intracellular growth along with media developed for culture of insect cells rather than bacteria [40-42, 44].

Given these precedents, we took a similar approach to culturing *H. defensa.* However, our priorities focused on identifying conditions in which strains of *H. defensa* that vary in levels of in vivo protection against *A. ervi* grow to comparable titres as in the aphid *A. pisum.* This allowed us to conduct experiments in vitro that would indicate whether symbiont products from different strains disable development of *A. ervi*. Consistent with in vivo results [20], in vitro assays indicate that APSE3-infected AS3 disable embryonic development of *A. ervi*. However, our in vitro data differ from prior in vivo studies by showing for the first time that: 1) AS3 kills *A. ervi* eggs in the absence of any aphid-produced product, and 2) mortality is due to a soluble factor released from *H. defensa*/APSE that can enter wasp eggs prior to rupturing from the chorion as a morula. We also found that while cytotoxic to *A. ervi* this factor does not reduce the viability of TN5 cells, although TN5 cell growth was lower when co-cultured with *H. defensa* when compared to growth

without bacteria. This finding is consistent with observations that wasps are killed in vivo, but aphids suffer relatively minor costs to infection with *H. defensa* [31].

Also consistent with prior in vivo studies [20], we found that the phage-free strain A2C did not affect *A. ervi* development. However, our results show that APSE3 from the AS3 strain infect A2C, which then exhibits the ability to disable development of *A. ervi* embryos. These results indicate that some APSE variants retain the capacity to infect *H. defensa* and horizontally transfer the factor(s) that kill *A. ervi* embryos. While host restriction inevitably leads to heritable symbiont decay over time [54], phage-mediated horizontal transfer of ecologically important traits between heritable symbionts provides a mechanism to maintain symbiont innovation. In contrast, our finding that APSE8-infected ZA17 had no adverse effects on development of *A. ervi* eggs or larvae in vitro suggests that interactions with aphid factors are required for disabling wasp development in vivo. Consistent with this, APSE-8 *H. defensa* confers very little protection against *A. ervi* in some aphid genotypes, but moderate protection in others [29].

Our results identified TC100 plus 10% FBS as the preferred medium for *H. defensa* growth. However, we could not establish primary cultures from aphids in this medium without TN5 cells although the A2C, AS3 and ZA17 strains could be propagated as pure liquid cultures after multiple passages in vitro. We observed lower survival and/or growth of *H. defensa* in the other media and cell lines we tested including C6/36 and S2 cells. Thus, culture conditions likely account for the slow in vitro growth and intracellular persistence reported previously for *H. defensa* [40]. No cell lines have been established from aphids but our results are consistent with other studies showing that cells from non-host insect species benefit the culture of some heritable symbionts [40-44].

The near absence of intracellular *H. defensa* in TN5 cells strongly argues that infection of this cell line is not essential for persistence or growth of *H. defensa.* What TN5 cells provide that is required for initiating primary cultures and maintenance of the NY26 strain remains unclear. Unknown adaptations by A2C, AS3 and ZA17 after multiple passages in vitro could play a role in why these strains could be shifted to growth in pure liquid cultures. The factors that underlie the requirement of NY26 for culture at lower temperatures with TN5 cells will also require further study.

In summary, our results provide an essential foundation for identifying the *H. defensa*-associated molecules that provide protection to aphids against parasitoids, and understanding the function(s) of different APSE variants. The ability to propagate particular strains of *H. defensa* in pure cultures also provides opportunities for characterizing genomic and transcriptome variation and the potential for genetically manipulating this symbiont.

2.6 Acknowledgements

We thank Mary Ard in the University of Georgia Microscopy facility for assistance with transmission electron microscopy. This project was supported by NSF IOS 1256794 to K.M.O and M.R.S.

2.7 References

[1] Moran, N.A., McCutcheon, J.P. & Nakabachi, A. 2008 Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics* **42**, 165-190. (doi:doi:10.1146/Annurev.Genet.41.110306.130119).

[2] Moran, N.A. 2007 Symbiosis as an adaptive process and source of phenotypic complexity. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8627-8633. (doi:10.1073/pnas.0611659104).

[3] Oliver, K.M. & Martinez, A.J. 2014 How resident microbes modulate ecologically-important traits of insects. *Current Opinion in Insect Science* **4**, 1-7. (doi:10.1016/j.cois.2014.08.001).

[4] Douglas, A. 1998 Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annual Review of Entomology* **43**, 17-37. (doi:doi:10.1146/annurev.ento.43.1.17).

[5] Moran, N.A., Russell, J.A., Koga, R. & Fukatsu, T. 2005 Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Applied and Environmental Microbiology* **71**, 3302-3310. (doi:doi:10.1128/AEM.71.6.3302-3310.2005).

[6] Oliver, K.M., Degnan, P.H., Burke, G.R. & Moran, N.A. 2010 Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. In *Annual Review of Entomology* (pp. 247-266. (doi:10.1146/annurev-ento-112408-085305)

[7] Brisson, J.A. & Stern, D.L. 2006 The pea aphid, *Acyrthosiphon pisum*: an emerging genomic model system for ecological, developmental and evolutionary studies. *Bioessays* **28**, 747-755. (doi:10.1002/bies.20436).

[8] Asplen, M.K., Bano, N., Brady, C.M., Desneux, N., Hopper, K.R., Malouines, C., Oliver, K.M., White, J.A. & Heimpel, G.E. 2014 Specialisation of bacterial endosymbionts that protect aphids from parasitoids. *Ecological Entomology* **39**, 736-739. (doi:10.1111/een.12153).

[9] Heyworth, E. & Ferrari, J. 2015 A facultative endosymbiont in aphids can provide diverse ecological benefits. *Journal of Evolutionary Biology* **28**, 1753-1760. (doi:doi:10.1111/jeb.12705).

[10] Łukasik, P., van Asch, M., Guo, H., Ferrari, J. & Godfray, C.J. 2013 Unrelated facultative endosymbionts protect aphids against a fungal pathogen. *Ecology Letters* **16**, 214-218. (doi:doi:10.111/ele.12031).

[11] Montllor, C.B., Maxmen, A. & Purcell, A.H. 2002 Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. *Ecological Entomology* **27**, 189-195. (doi:10.1046/j.1365-2311.2002.00393.x).

[12] Oliver, K.M., Russell, J.A., Moran, N.A. & Hunter, M.S. 2003 Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1803-1807. (doi:10.1073/pnas.0335320100).

[13] Scarborough, C.L., Ferrari, J. & Godfray, H.C.J. 2005 Aphid protected from pathogen by endosymbiont. *Science* **310**, 1781-1781. (doi:10.1126/science.1120180).

[14] Schmid, M., Sieber, R., Zimmermann, Y.S. & Vorburger, C. 2012 Development, specificity and sublethal effects of symbiont-conferred resistance to parasitoids in aphids. *Functional Ecology* **26**, 207-215. (doi:10.1111/j.1365-2435.2011.01904.x).

[15] Vorburger, C., Gehrer, L. & Rodriguez, P. 2010 A strain of the bacterial symbiont *Regiella insecticola* protects aphids against parasitoids. *Biology Letters* **6**, 109-111. (doi:10.1098/rsbl.2009.0642).

[16] Wagner, S.M., Martinez, A.J., Ruan, Y.M., Kim, K.L., Lenhart, P.A., Dehnel, A.C., Oliver, K.M. & White, J.A. 2015 Facultative endosymbionts mediate dietary breadth in a polyphagous herbivore. *Functional Ecology* **29**, 1402-1410. (doi:10.1111/1365-2435.12459).

[17] Metcalf, J.A. & Bordenstein, S.R. 2012 The complexity of virus systems: the case of endosymbionts. *Current Opinion in Microbiology* **15**, 546-552. (doi:http://dx.doi.org/10.1016/j.mib.2012.04.010).

[18] Moran, N.A., Degnan, P.H., Santos, S.R., Dunbar, H.E. & Ochman, H. 2005 The players in a mutualistic symbiosis: Insects, bacteria, viruses, and virulence genes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 16919-16926. (doi:10.1073/pnas.0507029102).

[19] van der Wilk, F., Dullemans, A.M., Verbeek, M. & van den Heuvel, J. 1999 Isolation and characterization of APSE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrthosiphon pisum*. *Virology* **262**, 104-113. (doi:10.1006/viro.1999.9902).

[20] Oliver, K.M., Degnan, P.H., Hunter, M.S. & Moran, N.A. 2009 Bacteriophages encode factors required for orotection in a symbiotic mutualism. *Science* **325**, 992-994. (doi:10.1126/science.1174463).

[21] Weldon, S., Strand, M. & Oliver, K. 2013 Phage loss and the breakdown of a defensive symbiosis in aphids. *Proceedings of the Royal Society of London B: Biological Sciences* **280**, 20122103. (doi:doi:10.1098/rspb.2012.2103).

[22] Henry, L.M., Maiden, M.C.J., Ferrari, J. & Godfray, H.C.J. 2015 Insect life history and the evolution of bacterial mutualism. *Ecology Letters* **18**, 516-525. (doi:10.1111/ele.12425).

[23] Zytynska, S.E. & Weisser, W.W. 2016 The natural occurrence of secondary bacterial symbionts in aphids. *Ecological Entomology* **41**, 13-26. (doi:10.1111/een.12281).

[24] Oliver, K.M., Moran, N.A. & Hunter, M.S. 2005 Variation in resistance to parasitism in aphids is due to symbionts not host genotype. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 12795-12800. (doi:10.1073/pnas.0506131102).

[25] Degnan, P.H. & Moran, N.A. 2008 Evolutionary genetics of a defensive facultative symbiont of insects: exchange of toxin-encoding bacteriophage. *Molecular Ecology* **17**, 916-929. (doi:doi:10.1111/j.1365-294X.2007.03616.x).

[26] Henry, L.M., Peccoud, J., Simon, J.-C., Hadfield, J.D., Maiden, M.J., Ferrari, J. & Godfray, H.C.J. 2013 Horizontally transmitted symbionts and host colonization of ecological niches. *Current Biology* **23**, 1713-1717. (doi:doi:10.1016/j.cub.2013.07.029).

[27] Russell, J.A., Weldon, S., Smith, A.H., Kim, K.L., Hu, Y., Lukasik, P., Doll, S., Anastopoulos, I., Novin, M. & Oliver, K.M. 2013 Uncovering symbiont-driven genetic diversity across North American pea aphids. *Molecular Ecology* **22**, 2045-2059. (doi:10.1111/mec.12211).

[28] Degnan, P.H. & Moran, N.A. 2008 Diverse phage-encoded toxins in a protective insect endosymbiont. *Applied and Environmental Microbiology* **74**, 6782-6791. (doi:10.1128/aem.01285-08).

[29] Martinez, A.J., Doremus, M.R., Kraft, L.J., Kim, K.L. & Oliver, K.M. 2017 Multi-modal defenses in aphids offer redundant protection and increased costs likely impeding a protective mutualism. *Journal of Animal Ecology*. (doi:10.1111/1365-2656.12675).

[30] Hufbauer, R.A. 2002 Aphid population dynamics: does resistance to parasitism influence population size? *Ecological Entomology* **27**, 25-32. (doi:10.1046/j.1365-2311.2002.0379a.x).

[31] Oliver, K.M., Campos, J., Moran, N.A. & Hunter, M.S. 2008 Population dynamics of defensive symbionts in aphids. *Proceedings of the Royal Society B-Biological Sciences* **275**, 293-299. (doi:10.1098/rspb.2007.1192).

[32] Schellhorn, N.A., Kuhman, T.R., Olson, A.C. & Ives, A.R. 2002 Competition between native and introduced parasitoids of aphids: nontarget effects and biological control. *Ecology* **83**, 2745-2757. (doi:DOI: 10.2307/3072012).

[33] Smith, A.H., Lukasik, P., O'Connor, M.P., Lee, A., Mayo, G., Drott, M.T., Doll, S., Tuttle, R., Disciullo, R.A., Messina, A., et al. 2015 Patterns, causes and consequences of defensive

microbiome dynamics across multiple scales. *Molecular Ecology* **24**, 1135-1149. (doi:10.1111/mec.13095).

[34] Oliver, K.M., Smith, A.H. & Russell, J.A. 2014 Defensive symbiosis in the real world -advancing ecological studies of heritable, protective bacteria in aphids and beyond. *Functional Ecology* **28**, 341-355. (doi:doi:10.1111/1365-2435.12133).

[35] Doremus, M.R. & Oliver, K.M. 2017 Aphid heritable symbiont exploits defensive mutualism. *Applied and Environmental Microbiology* **83**. (doi:10.1128/aem.03276-16).

[36] Martinez, A.J., Weldon, S.R. & Oliver, K.M. 2014 Effects of parasitism on aphid nutritional and protective symbioses. *Molecular Ecology* **23**, 1594-1607. (doi:10.1111/mec.12550).

[37] Oliver, K.M., Noge, K., Huang, E.M., Campos, J.M., Becerra, J.X. & Hunter, M.S. 2012 Parasitic wasp responses to symbiont-based defense in aphids. *BMC Biology* **10**. (doi:10.1186/1741-7007-10-11).

[38] Martinez, A.J., Ritter, S.G., Doremus, M.R., Russell, J.A. & Oliver, K.M. 2014 Aphid-encoded variability in susceptibility to a parasitoid. *BMC Evolutionary Biology* **14**. (doi:10.1186/1471-2148-14-127).

[39] McLean, A.H.C. & Godfray, H.C.J. 2015 Evidence for specificity in symbiont-conferred protection against parasitoids. *Proceedings of the Royal Society B-Biological Sciences* **282**. (doi:10.1098/rspb.2015.0977).

[40] Darby, A.C., Chandler, S.M., Welburn, S.C. & Douglas, A.E. 2005 Aphid-symbiotic bacteria cultured in insect cell lines. *Applied and Environmental Microbiology* **71**, 4833-4839. (doi:10.1128/aem.71.8.4833-4839.2005).

[41] Hypsa, V. & Dale, C. 1997 In vitro culture and phylogenetic analysis of "*Candidatus Arsenophonus triatominarum*," an intracellular bacterium from the triatomine bug, *Triatoma infestans*. *International Journal of Systematic Bacteriology* **47**, 1140-1144. (doi:doi:10.1099/00207713-47-4-1140).

[42] Matthew, C.Z., Darby, A.C., Young, S.A., Hume, L.H. & Welburn, S.C. 2005 The rapid isolation and growth dynamics of the tsetse symbiont *Sodalis glossinidius*. *Fems Microbiology Letters* **248**, 69-74. (doi:10.1016/j.femsle.2005.05.024).

[43] Pontes, M.H. & Dale, C. 2006 Culture and manipulation of insect facultative symbionts. *Trends in Microbiology* **14**, 406-412. (doi:10.1016/j.tim.2006.07.004).

[44] Welburn, S.C., Maudlin, I. & Ellis, D.S. 1987 In vitro cultivation of *Rickettsia*-like-organisms from *Glossina* spp. *Annals of Tropical Medicine and Parasitology* **81**, 331-335. (doi:doi:10.1080/00034983.1987.11812127).

[45] Sabri, A., Leroy, P., Haubruge, E., Hance, T., Frere, I., Destain, J. & Thonart, P. 2011 Isolation, pure culture and characterization of *Serratia symbiotica* sp nov., the R-type of secondary endosymbiont of the black bean aphid *Aphis fabae*. *International Journal of Systematic and Evolutionary Microbiology* **61**, 2081-2088. (doi:10.1099/ijs.0.024133-0).

[46] Degnan, P.H., Yu, Y., Sisneros, N., Wing, R.A. & Moran, N.A. 2009 *Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors.

Proceedings of the National Academy of Sciences of the United States of America **106**, 9063-9068. (doi:10.1073/pnas.0900194106).

[47] Burke, G.R., Thomas, S.A., Eum, J.H. & Strand, M.R. 2013 Mutualistic Polydnaviruses Share Essential Replication Gene Functions with Pathogenic Ancestors. *Plos Pathogens* **9**. (doi:10.1371/journal.ppat.1003348).

[48] Laughton, A.M., Garcia, J.R., Altincicek, B., Strand, M.R. & Gerardo, N.M. 2011 Characterisation of immune responses in the pea aphid, *Acyrthosiphon pisum*. *Journal of Insect Physiology* **57**, 830-839. (doi:10.1016/j.jinsphys.2011.03.015).

[49] Baumann, P. 2005 Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annual Review of Microbiology* **59**, 155-189. (doi:10.1146/annurev.micro.59.030804.121041).

[50] Grenier, S. 2012 Artificial rearing of entomophagous insects, with emphasis on nutrition and parasitoids -general outlines from personal experience. *Karaelmas Science and Engineering Journal* **2**, 1-12.

[51] Pennacchio, F. & Strand, M.R. 2006 Evolution of developmental strategies in parasitic hymenoptera. In *Annual Review of Entomology* (pp. 233-258.

[52] Thompson, S.N. 1999 Nutrition and culture of entomophagous insects. *Annual Review of Entomology* **44**, 561-592. (doi:10.1146/annurev.ento.44.1.561).

[53] Grbic, M. & Strand, M.R. 1998 Shifts in the life history of parasitic wasps correlate with pronounced alterations in early development. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1097-1101. (doi:10.1073/pnas.95.3.1097).

[54] Bennett, G.M. & Moran, N.A. 2015 Heritable symbiosis: The advantages and perils of an evolutionary rabbit hole. *Proceedings of the National Academy of Sciences* **112**, 10169-10176. (doi:10.1073/pnas.1421388112).

CHAPTER 3

APSE HAPLOTYPES VARY IN LIFE CYCLE DESPITE BEING NEARLY IDENTICAL IN GENE CONTENT AND EXPRESSION

[1] Brandt, J.W., Chevignon G., Oliver, K.M., Strand, M.R. To be submitted to: .

3.1 Abstract

There is a great diversity of bacteriophage found infecting all manner of prokaryotic organisms on Earth. Phage have been shown to have important impacts on bacterial population, ecology, and physiology. This is especially true of temperate phage which can also act as mechanisms of horizontal transfer for ecologically important traits. One such phage, known as APSE, infects the secondary endosymbiont *Hamiltonella defensa* in pea aphids. APSE has been shown to impart a protective phenotype to *H. defensa* so that aphids carrying certain phage-bacteria combinations are variably resistant to parasitism by a parasitic wasp. Currently there are four haplotypes of APSE known to occur in North American pea aphids that provide either moderate (APSE2 and 8) or complete (APSE1 and 3) protection against parasitoids. While ecological details and impacts of these different haplotypes are fairly well established, we know little about how these different haplotypes behave or persist within *H. defensa*. Here, we use a comparative genomic and transcriptomic approach to better understand differences and similarities between haplotypes. We also use qPCR and in vitro culturing techniques to show that APSE haplotypes not only differ in how they persist and replicate within *H. defensa* population, but also in their ability to infect new *H. defensa* strains despite largely conserved genomes and expression profiles.

3.2 Introduction

Bacteriophages constitute the most abundant and diverse form of life known (Rohwer 2003, Deresinski 2009). They can infect all forms of prokaryotic organisms on Earth with up to 10

different phage types in a single bacterial species (Clokie et al. 2011). Phages have been shown to alter bacterial population, ecology, and physiology depending on the life cycle of the bacteriophage and bacterium. Phage replication falls on a continuum from virulent (lytic) to temperate (lysogenic) (Brussow et al. 2004, Miller and Day 2008, Metcalf and Bordenstein 2012, Nanda et al. 2015). Lytic phages begin replication immediately following infection of a bacterium and produce and assemble new virus particles that are released through bacterial cell lysis resulting in the bacteria's death. In contrast, lysogenic phages infect bacteria and integrate into the host genome as a prophage (Barksdale and Arden 1974, Kutter 2004, Calendar 2005). Temperate lysogenic phages retain the machinery needed for lytic replication and lysogeny (Feiner et al. 2015). Through integration in to their bacterial host genome, temperate phage are able to introduce new genes into bacterial populations while also modifying population dynamics through lytic replication (Feiner et al. 2015, Davies et al. 2016).

One such phage, named *Acyrthosiphon pisum* secondary endosymbiont (APSE), infects the pea aphid facultative endosymbiont, *Hamiltonella defensa*. APSEs provide protection to *H. defensa* host, *A. pisum*, to the parasitoid wasp *Aphidius ervi* (Oliver et al. 2009, Weldon et al. 2013). APSEs are phylogenetically related to the *Podoviridae* (van der Wilk et al. 1999). APSEs have isometric heads with exceptionally short tails and are around 55 nm in size (van der Wilk et al. 1999, Clark et al 2001, Lima-Mendez et al 2008, Lavigne et al 2008, Fouts et al 2013, Lopes et al 2014). Sequencing of several *H. defensa* strains containing proviral APSE haplotypes indicates each viral genome shares core genes with features in replication but differ in a region of the genome containing eukaryotic toxin genes that have been hypothesized to protect aphids against

parasitoids (Degnan et al. 2008, Oliver et al. 2009, Chevignon et al. 2018). However, there has been no biological evidence linking toxin genes directly to killing wasp parasitoids. While we have made strides in understanding the basis of aphid protection through the creation of an *in vitro* system which allows us to culture *H. defensa* and its associated bacteriophage outside of aphid hosts we still know relatively little about basic APSE biology (Brandt et al. 2017).

Here we present a comparative genomic study of 5 APSE haplotypes. New genomic data coupled with our ability to culture *H. defensa* and APSE *in vitro* allowed us to also perform a comparative characterization of APSEs using electron microscopy imaging, quantitative and diagnostic PCR. Our results offer interesting insights into APSE and other bacteriophage biology by showing that despite largely conserved genomes and transcriptomes APSE haplotypes can vary greatly in how they replicate and their impacts on bacteria and larger organisms.

3.3 Methods

(a) Insect rearing and *H. defensa* cultures. *A. pisum* lines of identical genetic background were previously established from single parthenogenetic females that hosted the phage-free A2C, APSE3-infected AS3, APSE2-infected NY26 and APSE8-infected ZA17 strains of *H. defensa* (Oliver et al. 2009; 2012; Martinez et al. 2014; Doremus & Oliver 2017). In vitro cultures were established from these aphid lines and maintained in 25 cm^2 culture flasks (Falcon) under ambient atmosphere at 27° C in 5 ml of TC100 medium (Sigma) plus 10% of fetal bovine serum (FBS) (Hyclone) (Brandt et al. 2017). The NY26 strain was maintained at 20° C in the same culture medium plus 1 x 10^4 TN5 cells, which is an adherent cell line from the moth *Trichoplusia ni* (Brandt et al. 2017). A2C, AS3, and ZA17 were passaged weekly after growing to an average maximum density of 1 x 10^{10} bacteria per ml while NY26 was passaged biweekly after growing to the same density.

(b) Estimating *Hamiltonella defensa* and APSE titers. *H. defensa* and APSE isolation have been previously described (Brandt et al 2017). Briefly, ten 2-day-old adult aphids were opened using forceps and hemolymph extracted into 500 µl of phosphate buffered saline (PBS). After removing the aphid body, PBS and hemolymph was transferred to a 1.5 ml microfuge tube and gently centrifuged at 50g to pellet aphid cells and debris. The supernatant (containing APSE and *H. defensa*) was then transferred to a new sterile tube and spun at 8000g for 10 min at 4°C to pellet *H. defensa.* After decanting and reserving the supernatant, DNA was extracted from *H. defensa* as

described in Weldon et al 2013 and stored at -20°C. The encapsidated APSE was then filtered through a 0.2 mm syringe filter to eliminate all bacterial cells. Two microliters of TURBO DNase (Ambion) and 5 µl of buffer was then added to 50 µl of this supernatant for 40 min at 37°C. This treatment degraded any naked DNA in the medium but cannot degrade DNA that is packaged inside of a virus particle. After DNase inactivation, supernatant template DNA was then extracted (Weldon et al 2013) and stored at -20°C. DNA was similarly isolated from *H. defensa* and APSE in established cultures by collecting 500 ml of medium 9 days after new cultures were passaged. *H. defensa* and APSE titers were estimated by qPCR using specific primers for the single-copy *H. defensa* dnaK gene and APSE major head protein gene (Table 3) (Brandt et al 2017). Ten microliters of qPCR reactions contained 0.5 µl of 5 mM of each primer, 5 µl of the Rotor-Gene SYBR Green Master Mix and 1 µl of template DNA and 3 µl of ultrapure water. Four technical replicates were run per sample using a Rotor-Gene Q cycler (Qiagen) with data acquired during the extension step. *H. defensa* and APSE copy number ml^{-1} of medium was calculated by multiplying the qPCR estimate by the dilution factor and elution volume. Titer data were analyzed by analysis of variance (two-way ANOVA) followed by the post hoc Tukey HDS test.

(c) APSEs infectivity experiments. Previous studies showed that APSE3 is able to infect *H. defensa* line A2C (Brandt et al 2017). Similar methods for APSE infection were used here. A2C titer was determined by qPCR as described above. Twelve new A2C cultures were initiated by adding $1x10^7$ bacteria to 4 ml of fresh TC100 plus 10% FBS and $1x10^4$ TN5 cells in a six-well culture plate (Corning). APSE3, APSE2 and APSE8 in medium from the AS3, NY26 and ZA17

strains respectively were then added to three A2C cultures (n = 3) at a multiplicity of infection (MOI) of 1, 0.1 and 0.01 while three other A2C cultures were grown without APSEs. After 9 days at 27°C, 0.5 ml of medium was centrifuged followed by qPCR (described above) to determine *H. defensa* copy number in the cell pellet and APSEs copy in the DNAse pretreated supernatant. After 7 days 10 µl of each A2C infected culture previously described were passed to 4 ml of fresh TC100 plus 10% FBS and $1x10^4$ TN5 cells in order to establish a second generation (g2) of infected A2C. After 9 days DNA was extracted from medium and pelleted as described above from the g2 culture APSE along with *H. defensa* titers were then assessed by qPCR as described above. The same procedure was performed to assess the infectivity of APSEs isolated from in vivo aphid adults.

(d) APSEs genome analysis. APSE genomes were retrieved from different publicly available sources reported in previous studies. APSE1 (accession: NC_000935) was isolated, sequenced, and characterized in van der Wilk et al 1999. APSE2 5AT (accession: NC_012751) was sequenced along with the *H. defensa* 5AT strain genome in Degnan et al 2009. APSE2 NY26 (accession: CP017605), APSE3 AS3 (accession: CP017610), and APSE8 ZA17 (accession: CP017613) were also sequenced along with *H. defensa* in Chevignon et al 2018. The five APSE genomes were aligned with MAFFT (Katoh and Standley 2013) implemented in Geneious (Kearse et al 2012) and their respective annotations were manually corrected and normalized based on the annotation pipeline described in Chevignon et al 2018. Protein similarities were assessed by extracting coding sequences (CDS) from each genome haplotype followed by sequence translation and MAFFT alignment (Katoh and Standley 2013) implemented in Geneious (Kearse et al 2012).

(e) APSE transcriptome characterization. In vitro APSE3 and APSE8 transcriptomes were isolated from the previously described AS3 and ZA17 *H. defensa* strain transcriptome (Chevignon et al 2018). For in vivo APSE3 and APSE8 transcriptomes the same protocol was used with some modification in order to enrich *H. defensa* from aphids. Briefly, a pool of approximately 400 fourth instar and adult aphids were dissected to release hemolymph into 1 ml drops of PBS. After removing the aphid carcasses each drop was transferred to 1.5 ml microfuge tube and gently centrifuged at 50g to pellet aphid cells and debris. The supernatant (containing APSE and *H. defensa*) was then transferred to a new tube and spun at 8000g for 10 min at 4°C to pellet *H. defensa.* Total RNA was isolated from AS3 (containing APSE3) and ZA17 (containing APSE8) strains using the mirVana miRNA Isolation Kit (Ambion, Thermo Fisher Scientific) followed by DNase treatment using the TURBO DNA-free Kit (Ambion, Thermo Fisher Scientific) and ethanol precipitation in the presence of glycogen. A total of three 400 aphid pools per strain were extracted generating 3 independent biological replicates. We sequentially depleted rRNA using the Ribo-Zero (Bacteria) Magnetic Kit (Epicentre, Illumina) according to manufacturer's protocol followed by Ribo-Zero-Treated RNA purification using the Illumina modified RNeasy MinElut Cleanup Kit (Qiagen). RNA templates were quality checked by the Georgia Genomics Facility using an Agilent 2100 Bioanalyzer (Agilent Technologies) and Fragment Analyzer (Advanced Analytical). Standard paired-end sequencing libraries (75 bp) were generated using the Kapa stranded RNA-seq Library Preparation Kit (Kapa Biosystems) and Illumina sequenced on a NextSeq (150 Cycles) Mid Output Flow Cell. RNAseq reads for each strain were aligned to *A. pisum* (accession:

ABLF00000000), *B. aphidicola* (accession: ACFK01000001), ApV (accession: NC_003780) and *H. defensa* / APSE (accession: CP017610 and CP017613) genomes using bowtie2 (Langmead and Salzberg, 2012) with the stringent parameters --very-sensitive, --no-mixed, and --no-discordant. Read counts for each CDS were generated using htseq-count (Anders et al. 2015). Differential expression analyses were assessed by DESeq2 (Love et al 2014) implemented in R (R Core Team 2013). Gene expression was considered significantly different when fold change was > 4 and p-value < 0.01. Because of the difference of aligned reads to *H. defensa* genome between in vitro (20 to 46 million reads) and in vivo (1 to 8 million reads) we performed subsampling of the in vitro reads libraries in order to obtain the same number of reads for in vivo and in vitro samples (Wu and Wang 2016). Subsampling of the in vitro libraries was done with the seqtk tool (https://github.com/lh3/seqtk) (Table 2).

(f) Electron microscopy. *H. defensa* was isolated from adult aphids or from in vitro cultures as described above. After decanting the supernatant, the pellet consisting of bacteria and insect cells was fixed for 2 h in 0.1 M cacodylate–HCl buffer (pH 7.0) containing 2% glutaraldehyde (v/v), 2% paraformaldehyde (v/v) and 0.2% picric acid (v/v). After rinsing in 0.1 M cacodylate–HCl buffer and repelleting cells were resuspended in 2% agar at 50°C followed by centrifugation for 10 min and storage at 4°C until the agar hardened. Agar-embedded samples were then post-fixed in 1% OsO4 in 0.1 M cacodylate–HCl buffer for 1 h, rinsed in deionized water, and dehydrated in graded ethanols (30–100%) with 15min wash steps. Samples were then rinsed in acetone followed by a rinse in propylene oxide for 15 min before infiltrating in proplyene oxide: Epon-Araldite (1:1,

1:3) for 2 h. Samples were embedded in 100% Epon-Araldite, which polymerized over 24 h in a 75°C oven. Sections were cut with a Reichert Ultracut S ultramicrotome and stained with uranyl acetate and lead citrate. Sections were then examined and photographed using a JEOL JEM1011 (JEOL, Inc., Peabody, MA) Transmission electron microscope (TEM) at 100 kV.

(g) PCR-based integration sites characterization. Complete genome assemblies for the A2C, AS3, NY26, and ZA17 strains previously identified the precise sites of integration for APSE2, APSE3, and APSE8 (Chevignon et al 2018). Using this information, we designed specific primers that amplified 5' (attL) and 3' (attR) APSE integration boundaries (ST. 3). To identify transduced particles we designed primers downstream of the integration site in the *H. defensa* chromosome allowing us to amply fragments from 245bp to 1150bp for attL site and from 96bp to 1716bp for attR site. PCR reactions (25 µl) contained 2.5 µl of 25 mM Mg^{2+}, 0.5 µl of mixed dNTPs (2.5 mM each), 0.5 µl of each primer (5 mM), 2 µl of template DNA, 0.2 µl of HotMaster Taq DNA polymerase and 18.6 µl of ultrapure water. Samples were run using a T100 thermal cycler (Bio-Rad) for 35 cycles. Eight microliters of each PCR reaction were then run on a 1% agar gel stained with ethidium bromide and visualized using a BioDoc-iT imaging system.

3.4 Results

(a) Genomic organization of APSEs. To date, 5 full APSE genomes are available: APSE1 (van der wilk et al. 1999), APSE2 from *Hamiltonella defensa* strain 5AT (Degnan et al. 2008A), APSE2

from NY26, APSE8 from ZA17, and APSE3 from AS3 (Chevignon et al. 2018). Whole genome alignment and annotation of those 5 APSE genomes highlighted differences among haplotypes. The least similar haplotypes (APSE2-NY26 and APSE3-AS3) shared 79.8% identity while the most similar haplotypes (APSE2-5AT and APSE2-NY26) shared 99.8% identity (Table 3.1). Genome annotation indicated that APSE genomes consist of 3 domains: (i) a region harboring genes coding for proteins involved in phage replication (ii) the toxin cassette containing genes coding for protein homologues of putative eukaryotic toxins and (iii) a region consisting of genes coding for structural proteins in capsids (Fig. 3.1).

In accordance with Degnan et al. 2008 we found that the most divergent region between APSE haplotypes was the toxin cassette. The replication machinery was largely conserved among haplotypes, but at least 3 possibly important deletions could be identified. The first deletion of 1.2 kb is present in APSE1 and APSE3; this sequence contains 2 genes, G and H, which encodes for a DNA-binding protein and a hypothetical protein respectively. The second deletion of 918 kb is present in the p45 gene of APSE3 and APSE8 which codes for the DNA polymerase I. The third and final deletion in the replication machinery region is in the P47 gene which codes for a "Phage protein." This deletion is variable in size amongst APSE haplotypes ranging from 78 to 207 bp depending on the haplotype. The particle assembly region is the most similar region between APSE haplotypes with a single 174 bp deletion in the P30 gene of APSE1, APSE3, and APSE8. This gene codes for a phage packaged DNA stabilization protein. Besides these deletions occurring in certain APSE haplotypes we also found some divergent sequences at the protein level. These

differences are summarized in (Fig. 3.2) with gene name and predicted function found in Supplemental Table 2.

Table 3.1: Percentage identity matrix between five APSE haplotypes. APSE haplotype relative identity similarities. Parent *H. defensa* line in parentheses when known. The higher the similarity between the two strains the darker the cell color.

	APSE1	**APSE2 (5AT)**	**APSE2 (NY26)**	**APSE8 (ZA17)**	**APSE3 (AS3)**
APSE1		83.37	83.58	82.37	81.49
APSE2 (5AT)	83.37		99.76	93.03	79.55
APSE2 (NY26)	83.58	99.76		93.23	79.75
APSE8 (ZA17)	82.37	93.03	93.23		83.45
APSE3 (AS3)	81.49	79.55	79.75	83.45	

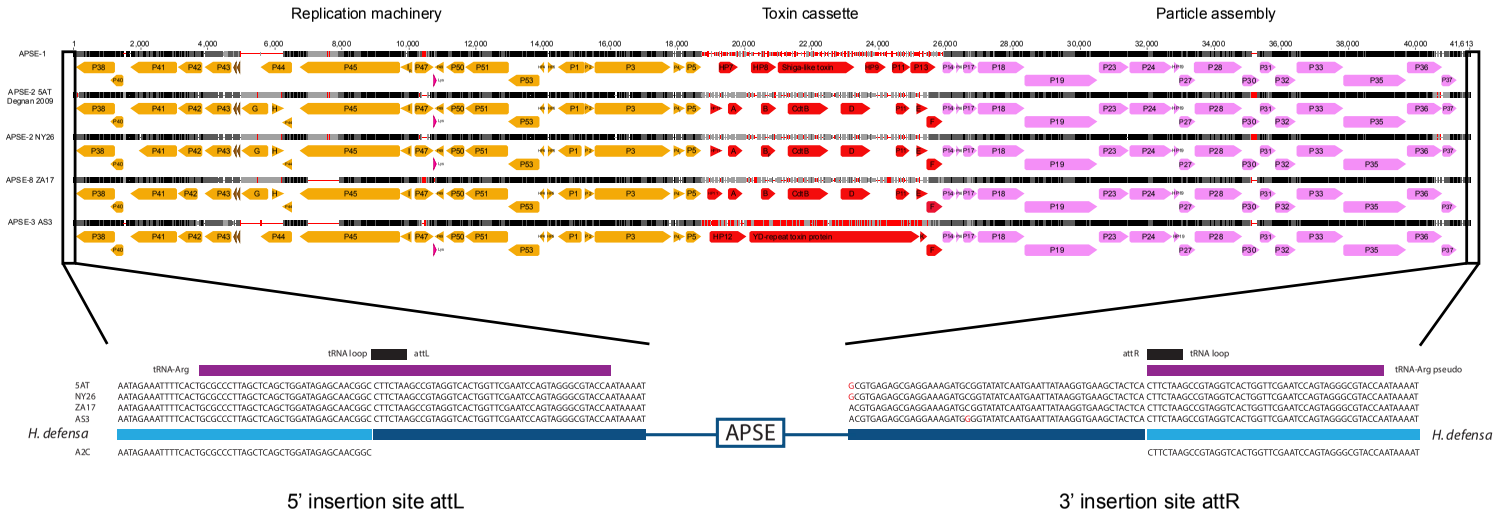


Figure 3.1: APSE Haplotypes whole genome alignment. Linear representations of different APSE haplotypes aligned to previously reported APSE1. CDS are color coded by functional regions where Orange=replication machinery, red=toxin cassette, and pink=particle assembly. Each CDS is labeled with a name that corresponds to a predicted function which can be found in Supplemental table 2. The lower part of the figure highlights the sequences of the attL and attR site when APSE is integrated into *H. defensa* genome. The purple bar highlights *H. defensa* tRNA-Arg while the black bar corresponds to the position of the anticodon in the tRNA-loop. Light blue bars correspond to *H. defensa*

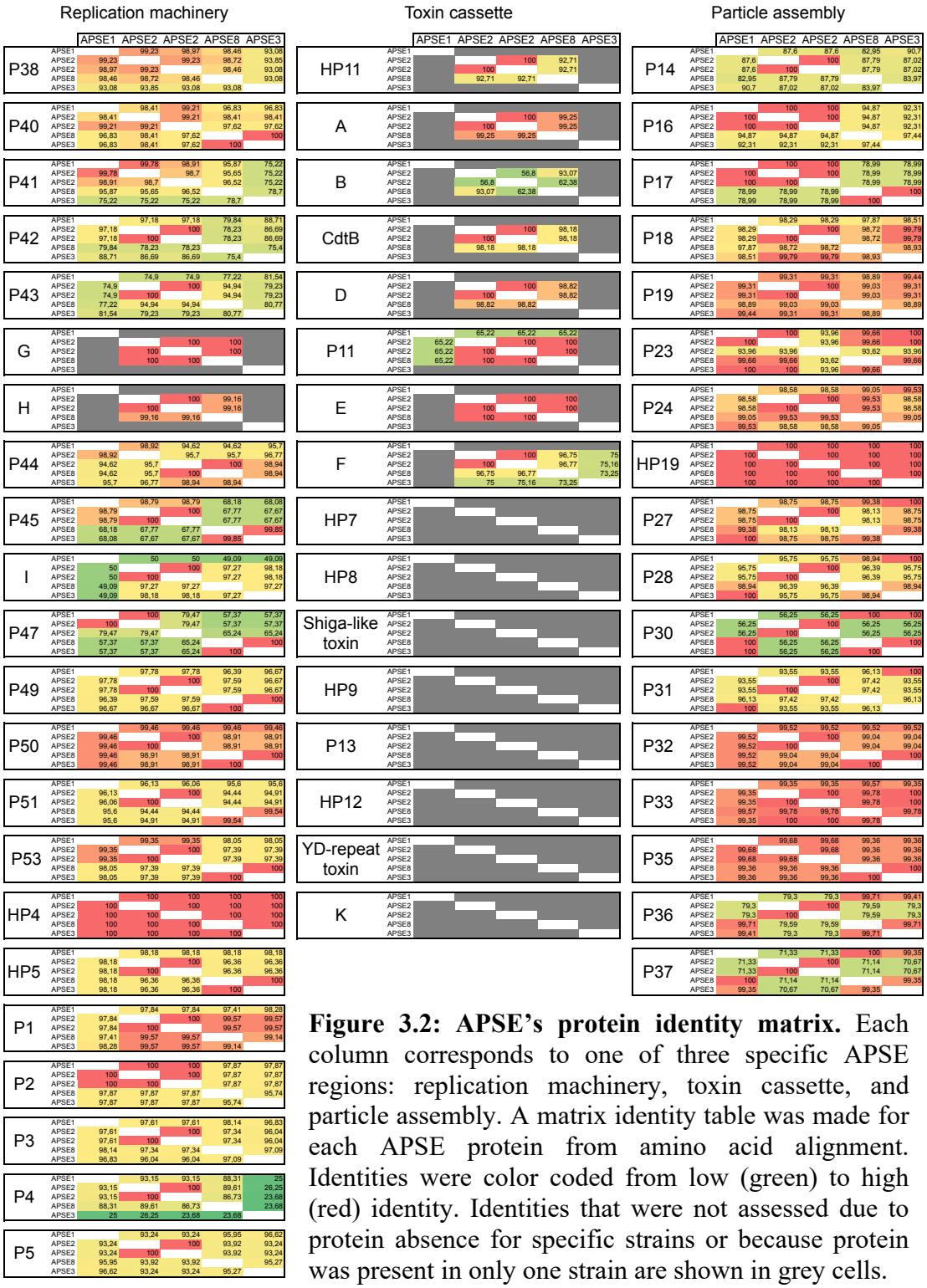


Figure 3.2: APSE's protein identity matrix. Each column corresponds to one of three specific APSE regions: replication machinery, toxin cassette, and particle assembly. A matrix identity table was made for each APSE protein from amino acid alignment. Identities were color coded from low (green) to high (red) identity. Identities that were not assessed due to protein absence for specific strains or because protein was present in only one strain are shown in grey cells.

(b) APSE differs in relative genome copy number and virion abundance in vitro and between haplotypes. Given the relative conservation of replication and packaging genes in all APSE haplotypes we would expect to see very little differences between haplotypes of APSE in the way in which they replicate (Degnan & Moran 2008). However, in vivo data from previous studies reveal that different haplotypes of APSE differ in their relative abundance to *H. defensa* (Weldon et al. 2013). Experimental aphid lines were created by injecting different *H. defensa* strains carrying APSE2 into 2 different previously uninfected aphid lines. These lines revealed that *H. defensa* and APSE2 abundance had an inverse relationship. That is, one line had relatively more APSE genomes per *H. defensa* and thus fewer *H. defensa* per aphid. This study also showed that some strains of *H. defensa* are capable of spontaneously losing APSE3 and that loss of APSE3 leads to a significant increase in the number of *H. defensa* per aphid. Anecdotally, we also found that populations of *H. defensa* carrying APSE3 rarely lost the phage and thus their protective phenotype in vitro (Brandt et al. 2017). Despite many years of rearing aphid lines in captivity and frequent screening we have never observed or seen reported a loss of any other APSE haplotype. These observations concur with the conclusions of Weldon et al. that APSE2 has a higher vertical transmission rate than APSE3. These observations are in stark contrast to largely conserved replication and packaging machinery found amongst APSE haplotypes and led us investigate how APSE haplotypes replicate and persist within *H. defensa*. We were also sensitive to the fact that there may be differences between in vitro and in vivo systems given that relative abundance of APSE changed with differing aphid lines (Weldon et al. 2013). This possibility was again highlighted by the fact that after many attempts we consistently failed to capture any images of

APSE from established in vitro cultures using transmission electron microscopy (TEM) (results not shown). In contrast, a previous study had captured APSE1 particles from aphids using TEM (van der Wilk et al. 1999). Our hypothesis that there were differences in phage production despite a large conservation of core genes was confirmed when we successfully produced TEM micrographs of APSE haplotypes 3 and 8 from adult aphids (Fig 3.3 BC).

These results raised further questions as to how APSE replicates within *H. defensa* and how changing conditions affect APSE lysogeny. We decided to answer these questions using qPCR-based methods to record the number of *H. defensa* and APSE. We also compared not only in vivo and established in vitro cultures (> 200 generations in cultures), but also newly created in vitro cultures (generation 1). We did this to answer 3 very important questions: 1) How many APSE genome copies per *H. defensa* are present? (ie. Is APSE accumulating in *H. defensa* and how much is accumulating?) 2) How much of this accumulated phage is DNase protected and present outside of *H. defensa*? (ie. How much of this accumulated phage is being processed and packaged into virions?) and 3) Do newly established in vitro cultures differ in both of these parameters from in vivo and established (>200 generations) cultures? (ie. Does a change in environmental conditions and nutrient availability effect lysogeny in APSE?)

Our results showed that in vivo there are more than 50 copies of the APSE8 genome for every *H. defensa* genome copy whereas there were 16 and 5 genome copies of APSE2 and APSE3 per *H. defensa* were found respectively (fig. 3.4A). Interestingly, the number of APSE copies per bacteria significantly drops in vitro for APSE2 and APSE8 whereas APSE3 has no significant decrease. This decrease in copy number happens after just a single passage of the bacteria in vitro

and maintains all APSE haplotypes at less than 10 copies per *H. defensa* for over 200 passages in vitro (Fig. 3.4A).

Not only do haplotypes vary in the number of APSE copy number per *H. defensa*, but they also vary in the amount of free and DNase protected APSE in either aphid hemolymph (in vivo) or culture medium (in vitro). Interestingly, the number of free particles does not align with increased APSE copy numbers. We found that APSE3 maintained a consistent level of free particles while APSE2 and 8 have rapid changes in the number of free particles (Fig. 3.4B). In new cultures APSE2 and APSE8 produced a significantly higher percentage of free particles with 11% and 16% free particles respectively. However, after only a few passages in vitro these numbers decline drastically and appear to stabilize at very low percentages (less than 5%). APSE3 on the other hand always produces a low percentage of free particles (2-5%) even during the switch from in vivo to in vitro conditions. Altogether, our data shows an accumulation of APSE genome copy number inside *H. defensa* in vivo, but these particles are possibly not processed, packaged, and released until the establishment of in vitro cultures (first passage or generation). Established cultures (over 200 generations) then show reduced APSE replication and particle production at least in the 2 moderately protective haplotypes APSE2 and APSE8. This possibly points to a shift to a mostly lysogenic lifestyle in the more stable and nutritious in vitro conditions. Similar shifts to lysogenic life cycles with expression of stress genes and shifts in nutrient availability have been observed in the temperate phages of *E. coli* (Edlin & Bitner 1977, Little and Michalowski 2010).

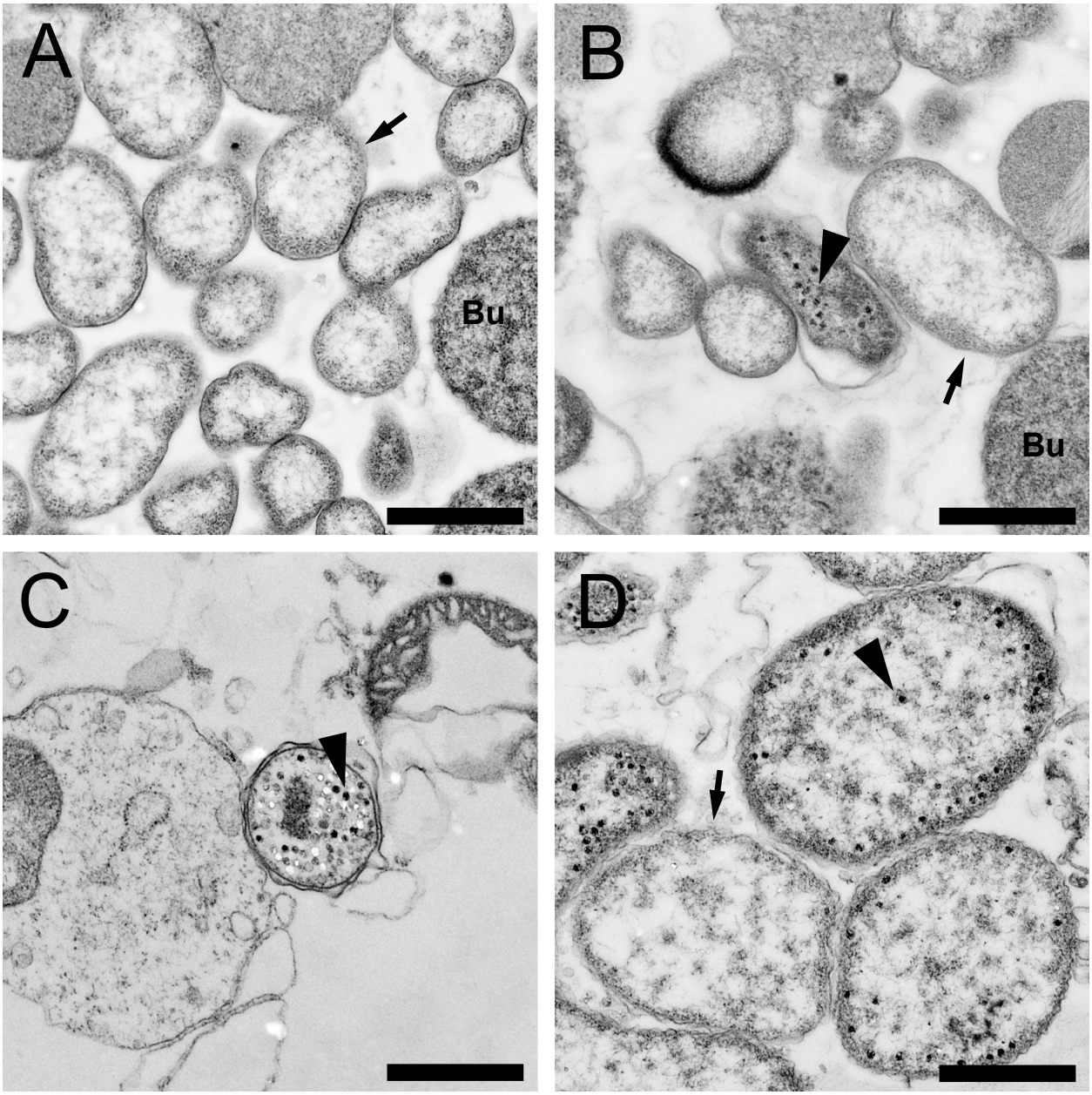


Figure 3.3: Transmission electron microscopy images of *H. defensa*. TEM micrographs of *H. defensa* isolated from adult pea aphids for the A2C (A), AS3 (B), and ZA17 (C) strains as well as *H. defensa* isolated from established culture of A2C strain 3 days after infection with in vitro APSE3 (D). Arrow: *H. defensa* free of APSE. Arrowhead: APSE particle inside *H. defensa*. Bu: *Buchnera aphidicola*. Scale bar equals 800 nm.

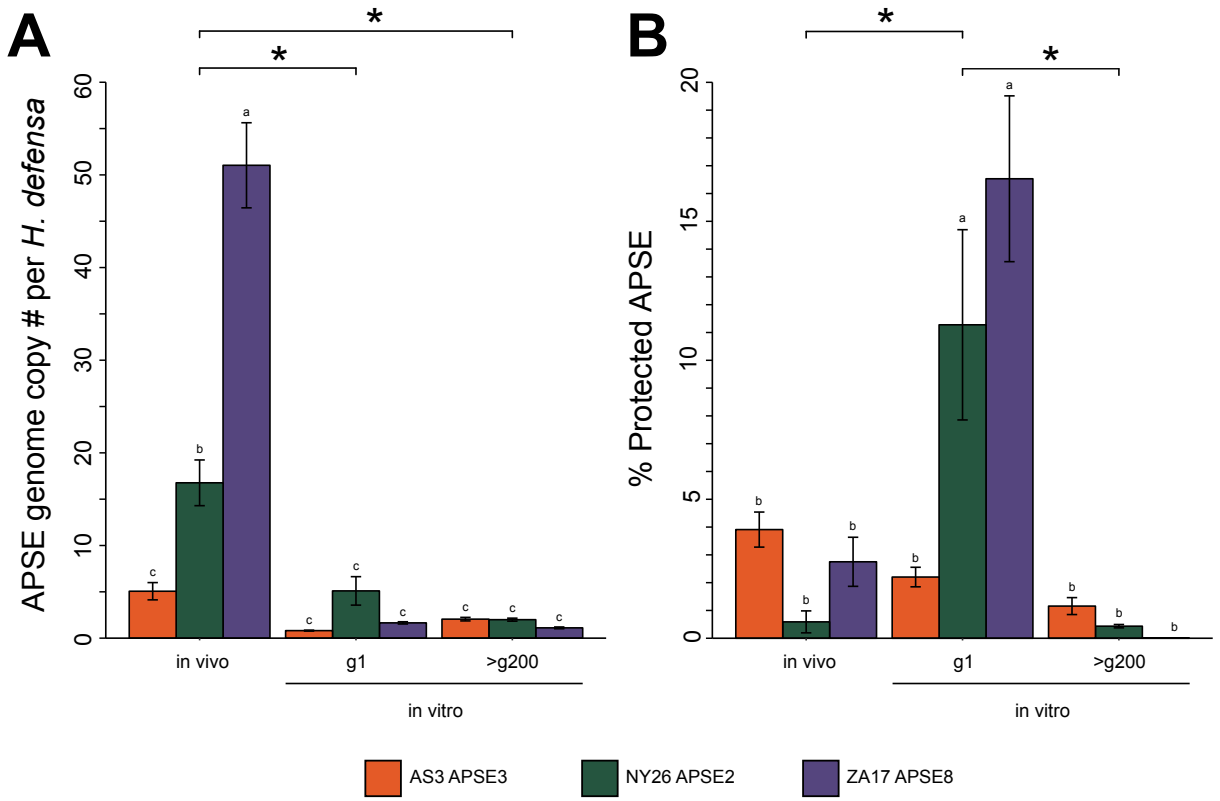


Figure 3.4: APSE particle production in vivo and in vitro. Number of APSE genome copy per *H. defensa* (A) and percentage of free APSE particles in the medium (B) for three different conditions: APSE from adult aphids (in vivo), newly established cultures of *H. defensa* (generation 1 = g1), or from continuously maintained cultures of *H. defensa* (> generation 200 = >g200) and for three different *H. defensa* strains (AS3 APSE3, NY26 APSE2 and ZA17 APSE8). Lines represent standard error and asterisks (*) indicate that the conditions were significantly different ($p < 0.05$). Letters above the bars show significant differences between treatments ($p < 0.05$) (two-way ANOVA followed by post hoc Tukey HSD).

(c) APSE3 infects and replicates in A2C line. We previously determined that APSE3 infected the A2C strain of *H. defensa* imparting a protective phenotype but did not test the infectivity of other haplotypes (Brandt et al. 2017). Here, we tested whether different APSE haplotypes infect A2C and also whether particles produced in *A. pisum* versus in vitro differ in infectivity. Therefore, we inoculated the A2C strain of *H. defensa* with APSE2, APSE3, and APSE8 from aphids and in vitro at MOI 1, 0.1, and 0.01. qPCR assays showed a significant decrease in the number of *H. defensa* 9 days after infection with APSE3 (MOI 1 and 0.1) when compared to the control uninfected A2C. However, APSE3 at a low MOI (0.01) had no effect on the number of *H. defensa* in culture 9 days after infection. Infection with APSE2 or APSE8 had no effect on the number of *H. defensa* no matter the MOI (Fig 3.5A). After one successive generation we continued to see a decreased number of *H. defensa* in cultures infected with APSE3, but now even the lowest MOI showed a decreased number of bacteria after 9 days of growth post culture passage (Fig 3.5B). Bacteriophage collected from aphids never reduced the number of *H. defensa* in culture after 9 days of growth no matter the APSE haplotype or MOI (Fig. 3.5C). In concert with a decrease in the number of *H. defensa* in APSE3 infected cultures there was a drastic increase in the number of APSE detected (up to $1x10^{12}$ APSE genome copies) at all 3 MOIs tested sourced from culture. In cultures infected with APSE2 and APSE8 there was no increase in the number of APSE from what was added to culture flasks (Fig 3.5D). This trend continued into the second generation of the test cultures, but there was a decrease in the number of APSE2 and APSE8 genomes found in cultures (Fig. 3.5E). When APSE haplotypes sourced from aphid hemolymph were added to cultures there was no increase in APSE genome copy number no matter the haplotype or MOI (Fig. 3.5F). Taken

together, these results indicate that the only APSE haplotype capable of infecting the A2C strain *H. defensa* was APSE3. Interestingly, this is only true for bacteriophage that are collected from established cultures as APSE collected from aphid hemolymph were never able to replicate in cultured A2C. Further supporting this observation, we find that only APSE3 sourced from cultures had multiple copies of APSE per *H. defensa* and also detectable amounts of free particles in newly infected cultures and these numbers were comparable to the established naturally APSE3 infected AS3 *H. defensa* strain (Fig. 3.5G-L). Additionally, we were able to capture APSE3 particles within *H. defensa* using TEM further supporting APSE3 was able to infect and replicate within A2C *H. defensa* (fig. 3.3D).

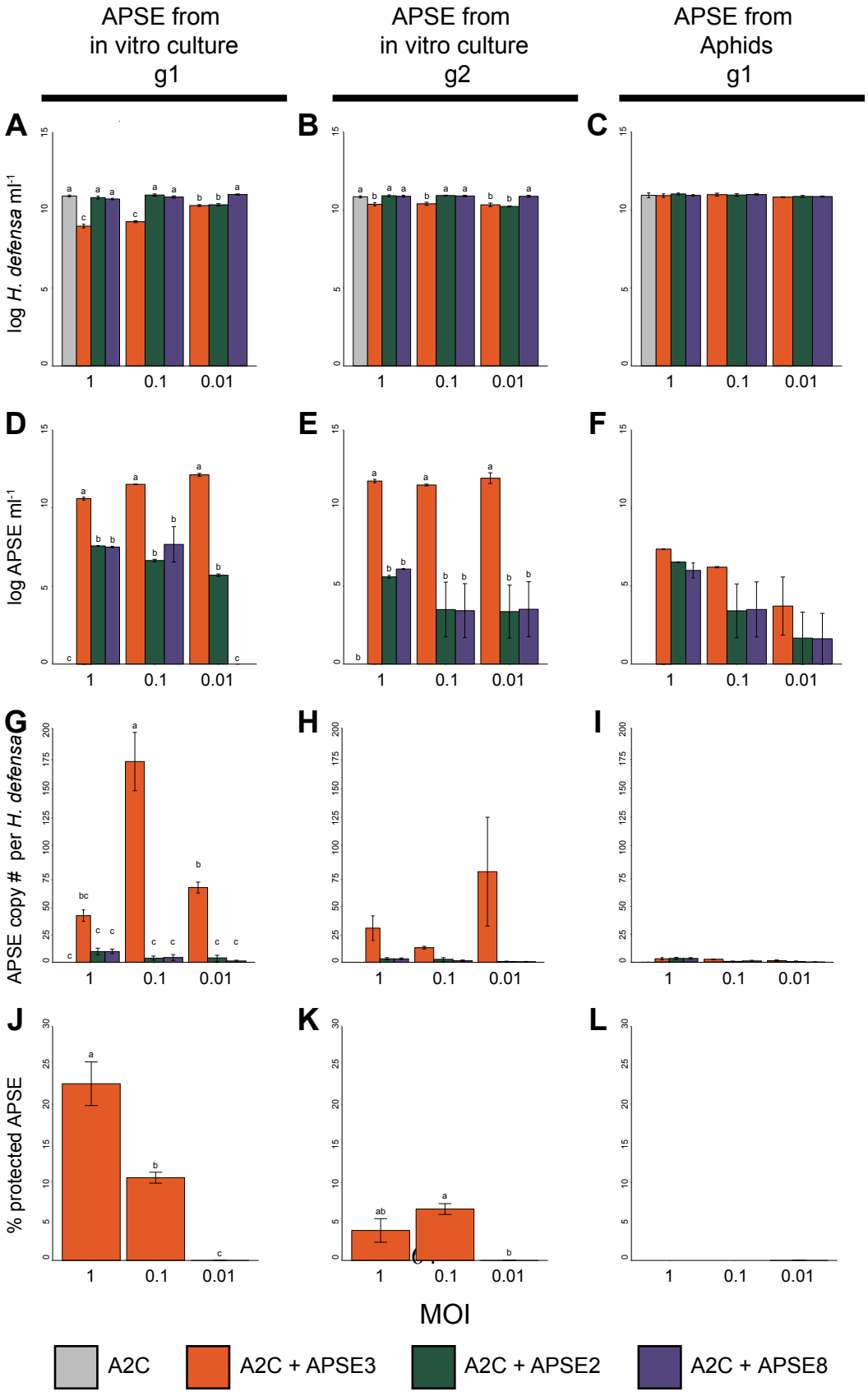

APSE from in vitro culture g1
APSE from in vitro culture g2
APSE from Aphids g1
A
B
C
D
E
F
G
H
I
J
K
L
log H. defensa ml-1
log APSE ml-1
APSE copy # per H. defensa
% protected APSE
1
0.1
0.01
MOI
A2C
A2C + APSE3
A2C + APSE2
A2C + APSE8

Figure 3.5: APSE3 infected the A2C strain of *H. defensa* but APSE2 and 8 do not. Inoculated starting cultures of $1x10^7$A2C *H. defensa* with APSE haplotypes from different sources. qPCR measurements were taken after 9 days incubation at 27°C. (A) Only APSE3 decreased the number of *H. defensa* in inoculated cultures at relatively high MOI (1 and 0.1) after 9 days when APSE was sourced from in vitro cultures. (B) After one successive generation (18 days post inoculation) there is no difference in the number of *H. defensa* in culture. (C) APSE sourced from adult aphids has no effect on *H. defensa* numbers in culture no matter the haplotype. (D) Only haplotype APSE3 sourced from in vitro cultures increased in number of genome copies from the starting inoculum ($1x10^7$, $1x10^6$, or $1x10^5$ for MOI 1, 0.1, 0.01 respectively) in the first generation (9 days post inoculation). (E) Haplotype APSE3 is present in high numbers (~$1x10^{10}$ genome copies) in the second generation (18 days post inoculation) while haplotypes APSE2 and APSE8 numbers have decreased from starting inoculum. (F) When APSE is sourced from adult aphids there is no difference in the number of APSE genomes after 9 days and no increase from the starting inoculation number. (G) Only APSE3 has a significantly higher number of APSE genome copies per *H. defensa* after 9 days. The ratio of APSE3 to *H. defensa* is highest in cultures started at MOI 0.1 (~ 200x) while MOI 1 and 0.01 still have significantly more copies of APSE per *H. defensa* (~50x and 75x) than other haplotypes at any MOI (≤1x). (H) Only in vitro sourced APSE3 inoculated cultures have ≥1x APSE genome per *H. defensa* in the second generation (18 days post inoculation) but are not significantly different from one another as in the first generation. (I) No in vivo sourced APSE had ≥1x more APSE than *H. defensa.* (J) After 9 days there is significantly more protected free APSE3 in medium from cultures started at MOI 1 (~20% of total APSE3) than MOI 0.1 (~10%) and MOI 0.01 (<1%). (K) After one successive generation both MOI 1 and 0.1 continue to produce free APSE3 at about 5% total APSE3 while cultures inoculated at MOI 0.01 do not seem to produce any free APSE3. APSE2 and APSE8 were never successfully recovered from medium only DNase treated samples. (L) No free APSE3 is found in cultures inoculated with aphid sourced APSE3 which is consistent with it being unable to infect A2C *H. defensa*. For all panels different letter correspond to significant differences ($p<0.05$, two-way ANOVA followed by post hoc Tukey HSD).

(d) Gene expression varies little between haplotypes but vary between in vitro and in vivo systems. In order to confirm the observed phenotype regarding the dichotomy of APSE particle production we decided to pursue an RNAseq analysis of APSE expression, both in vitro and in vivo, for two different *H. defensa* strains: ZA17 which carries APSE8 and AS3 which carries APSE3. In order to obtain high quality samples, we took advantage of our ability to culture *H. defensa* free of any aphid read contamination and without any library amplification. In order to get

clean reads from in vivo samples we developed a centrifugation-based protocol to enrich for bacteria from whole aphids. The sequencing results had a good amount of *H. defensa* reads from 22 to 35% of total reads for AS3, unfortunately only 5% of total reads mapped to *H. defensa* from the ZA17 strain (Supplemental table 2). This low level of *H. defensa* reads is mainly due to contamination of *A. pisum* reads, *Buchnera aphidicola*, and reads from an unexpected RNA virus named ApV (van der Wilk et al. 1997). A great majority of the *A. pisum* reads corresponded to eukaryotic rRNA that have not been removed by the bacteria-specific Ribo-Zero rRNA removal kit. Due to the important difference of aligned reads to the *H. defensa* genome between the samples we performed subsampling of the in vitro reads libraries in order to obtain the same read number in vivo and in vitro. Finally, because the present study focuses on APSE characterization, we isolated the expression of only APSE genes.

We found most genes were upregulated in vivo for the APSE8 haplotype but few differences in expression for APSE3 (Fig. 3.6A-B). This upregulation in APSE8 could account for the larger number APSE genomes per *H. defensa* and particle production found in vivo when compared to in vitro cultures. However, there was no obvious patterns in the genes upregulated making interpretation difficult. The mostly consistent expression of genes in APSE3 aligned well with its consistency in genome per *H. defensa* number and percentage free particles between systems. However, it does raise questions as to why in vivo sourced APSE3 was unable to infect A2C *H. defensa.* Surprisingly, we also found that there very few differences in the level of expression of genes between haplotypes in vitro despite some interesting differences in genome copy per *H. defensa* and infectivity and similar gene content (Fig. 3.6C). Only genes present in the

given haplotype (e.g. cdtB in APSE8 and YDr in APSE3) are significantly different. However, in vivo APSE8 tended to be more highly expressed in most genes when compared to APSE3, but there was, once again, no apparent pattern to these more highly expressed genes (Fig. 3.6D).

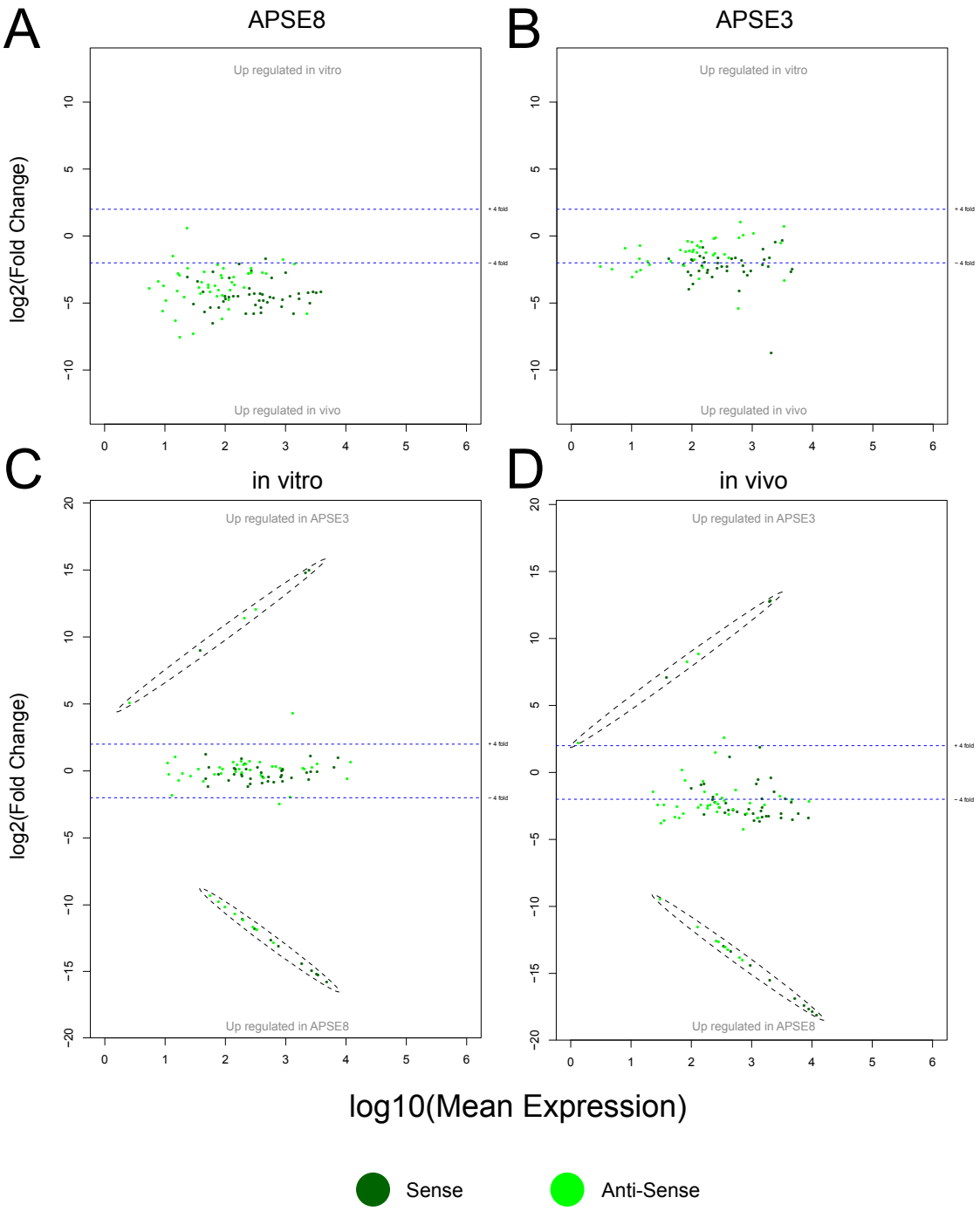

A
APSE8
Up regulated in vitro
Up regulated in vivo
+ 4 fold
− 4 fold
log2(Fold Change)
B
APSE3
Up regulated in vitro
Up regulated in vivo
+ 4 fold
− 4 fold
C
in vitro
Up regulated in APSE3
Up regulated in APSE8
+ 4 fold
− 4 fold
log2(Fold Change)
D
in vivo
Up regulated in APSE3
Up regulated in APSE8
+ 4 fold
− 4 fold
log10(Mean Expression)
Sense
Anti-Sense

Figure 3.6: In vivo and in vitro transcriptome of APSE3 and APSE8. Expression comparison of APSE8 between in vivo and in vitro (A). Expression comparison of APSE3 between in vivo and in vitro (B). Expression comparison between APSE3 and APSE8 in vitro (C). Expression comparison between APSE3 and APSE8 in vivo (D). Blue dashed lines bound the + 4-fold and – 4-fold change between comparisons. Dark green points indicate sense genes whereas light green points indicate antisense genes. Dashed circles encompass genes that are strain specific.

(e) ASPE performs specialized transduction. Previous studies of the APSE genome and its morphology have placed it in the family *Podoviridae* according to sequence and morphological similarities (van der Wilk et al. 1999, Degnan et al. 2008). The family has been classically described morphologically as having an icosahedral head with a short non-contractile tail (Lavigne et al. 2008, King et al. 2011). They are also described as having a linear double-stranded DNA (dsDNA) chromosome which is packaged into virions via rolling circle replication (King et al. 2011). In this model of replication, a precursor of a long DNA molecule consisting of several viral genomes laid end to end is cut non-specifically into 'headfuls,' head capsules that are slightly larger than the amount of DNA that is required to encode the genome. In this way, the resulting phage mature phage DNA particles are terminally repetitious and circularly permuted (Streisinger et al. 1967, Botstein et al. 1973). This method of replication is efficient and targeted, but it is not perfect and host DNA can be mistakenly packaged into head capsules at low frequencies (Ebel-Tsipsis et al. 1972). This phenomenon is known as generalized transduction as it packages host DNA non-specifically and is common among *Podoviridae* (Weisburg 1996, Cheng et al. 2012, Byrne & Kropinski 2005). Alternatively, some bacteriophage perform specialized transduction by which host DNA sequences flanking the prophage are packaged along with viral DNA during maturation (Cavanagh & Miller 1986). While specialized transduction is found in other

bacteriophage families it has only been reported in a handful of *Podoviridae* species (Pourcel et al. 2017).

However, virus taxonomy is extremely difficult and based on only a few exemplar species and horizontal transfer among viruses is rampant (Lavigne et al. 2008). As such, we wanted to confirm APSE's placement in *Podoviridae* and test assumptions about its method of replication. This was especially important because APSE had been previously reported as having a circularized genome by Degnan et al. (2008) and this would be unique among *Podoviridae*. We also had the goal of understanding APSE's ability to horizontally transfer genes to *H. defensa* through either general or specialized transduction.

In an attempt to characterize the APSE replication method we designed specific primers (Supplemental table 3) in order to amplify the APSE integration boundaries in the *H. defensa* chromosome (attF and attL) as well as the unintegrated chromosome (attB) and the circular or concatenated APSE (attP) (Fig 7). Using these primers, we found that we still recovered signals from both ends of the integration boundary in samples that should only contained APSE particles. This result suggested that APSE packaged portions of the *H. defensa* genome close to the integration site of APSE. In order to characterize the content of the free APSE particles isolated from medium we used primers to amplify multiple fragment sizes from 245-1150 bp for the attL site, 96-1716 bp for the attR site, 604-1943 bp for the attB site, and 204-251 bp for the attP site. Our results showed that no matter the size of amplified fragments we were always able to amplify both integration sites (attL and attR) in bacterial pellets. Interestingly, we received a signal from DNA isolated from free APSE particles with a size dependent efficacy. As we increased the size

of the target amplifiable fragment there was a reduced level of amplification both in vivo and in vitro. This pattern was present regardless of which boundary was expanded, but longer stretches of host DNA flanking the attR site were recovered in all strains and conditions (Fig 7A-D). These results suggested that APSE were capable of packaging portions of the *H. defensa* genome from both the 5' and 3' end simultaneously (Fig. 7E). Interestingly, amplification of the unintegrated attB site was recovered for all strains of *H. defensa* in vivo suggesting that all communities of *H. defensa* contain some population of uninfected bacteria. In vitro strains of *H. defensa* were similar in this regard except ZA17 (APSE8) which showed almost complete integration into the *H. defensa* population. This aligns well with our previous observations that an uninfected ZA17 in vitro line has never been spontaneously generated and it produces a very small percentage of virions in vitro. Interestingly, we also recovered an attB signal in APSE particles for *H. defensa* strains NY26(APSE2) and AS3(APSE3) in vivo, but not in vitro. We were always able to amplify the attP site for all strains across all conditions suggesting that circular or concatenated forms of APSE are present in both *H. defensa* and free in the hemolymph or medium (Fig7G-H). Taken together, these results reveal that APSE packaged stretches of the *H. defensa* genome closest to the integration boundaries and were capable of specialized transduction around the integration site. Results also suggested that APSE is capable of packaging a larger portion of the *H. defensa* genome in vivo when compared to in vitro. This access packaging was also not symmetrical and was biased to the attR site both in vitro and in vivo. We also provide further evidence that APSE is unique among *Podoviridae* in possibly packaging a circular dsDNA genome.

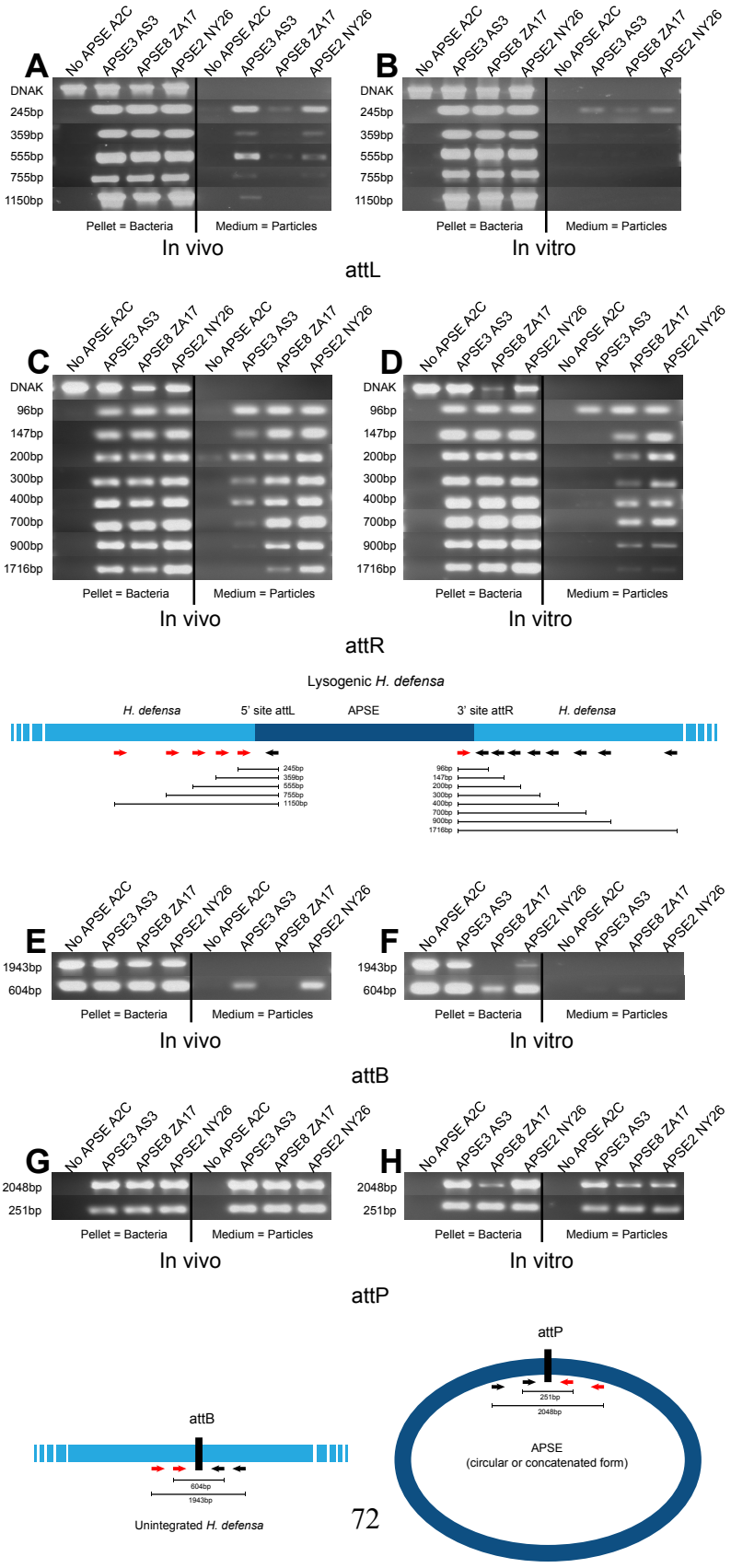

A
B
No APSE A2C
APSE3 AS3
APSE8 ZA17
APSE2 NY26
DNAK
245bp
359bp
555bp
755bp
1150bp
Pellet = Bacteria
Medium = Particles
In vivo
In vitro
attL
C
D
96bp
147bp
200bp
300bp
400bp
700bp
900bp
1716bp
attR
Lysogenic H. defensa
H. defensa
5' site attL
APSE
3' site attR
E
F
1943bp
604bp
attB
G
H
2048bp
251bp
attP
attB
Unintegrated H. defensa
APSE
(circular or concatenated form)

Figure 3.7: PCR-based characterization of attR, attL, attB and attP sites. Amplifications were done on DNA extracted from pellet or medium of A2C, AS3, ZA17 and NY26 *H. defensa* strains isolated from adult aphids or from in vitro established culture. (A, B) Amplification of the attL site. (C, D) Amplification of the attR site. (E, F) Amplification of the attB site. (G, H) Amplification of the attP site. Unscaled schematic representations of the different primer positions are given to help identify amplified fragment localization, red and black arrows correspond to forward and reverse primers respectively. It should be noted that the amplification of the attP site could correspond to the circular or to the concatenated form of APSE.

3.5 Discussion

Bacteriophage make up a large and diverse albeit relatively forgotten portion of lifeforms (Rohwer et al. 2003, Derenski et al. 2009). Phage are capable of infecting all forms of prokaryotic organisms with many organisms being infected with multiple phage simultaneously. They are also capable of drastically altering bacterial population, ecology, and physiology depending bactiophage and bacteria life cycle (Kutter 2004, Calendar 2005, Clokie et al. 2011, Oliver and Martinez 2014, Nanda et al. 2015, Davies et al. 2016). APSE is a temperate bacteriophage most closely related to the family *Podoviridae* and has been shown to be an important part of a complex tri-partite symbiosis (van Der Wilk et al. 1999, Degnan et al. 2008, Oliver 2009). In this symbiosis a bacterium, *Hamiltonella defensa,* infected with APSE provides their pea aphid hosts with protection against a parasitoid wasp (Oliver et al. 2005). Comparative studies of APSE haplotypes have largely concerned differences in protective phenotypes, ecological distribution, and genomic comparisons. However, very few studies have looked at the general biology of APSE haplotypes. These types of studies have been untenable in the past due to the complexity of working with tripartite symbioses. One major hurdle was obtaining enough material to adequately perform

genomics and transcriptomics for secondary symbionts and their associated phages when they are in relatively low abundance when compared to host and primary symbiotic bacteria DNA and RNA. However, using our newly developed culturing methods we were uniquely poised to construct and compare very complete and high-quality genomes and transcriptomes along with characterizing important aspects of APSE biology such as how they replicate, infect, and impact *H. defensa* for both in vivo and in vitro conditions (Brandt et al. 2017, Chevignon et al. 2018).

Comparisons of different APSE haplotype genomes agreed with previous studies and showed that there are very few differences with the most divergent region being the toxin cassette(Degnan et al. 2008) We also reported that the other 2 major regions of the APSE genome, the replication and particle assembly regions, were largely conserved with only 3 major deletions across all 5 haplotypes and 1 major deletion in 3 haplotypes (APSE1, 3, and 8) respectively. This conservation even extended to the predicted protein level as well with only 79.8% in the most divergent haplotypes and 99.8% conservation in the most closely related haplotypes. These results corroborated previous findings and built upon them by providing 3 new complete APSE haplotype genomes from new lines of *H. defensa* (APSE2 from NY26, APSE3 from AS3, and APSE8 from ZA17) with improved sequencing depths that allowed for better annotations and new comparisons to be made. The sequencing of APSE2 from the *H. defensa* line NY26 also allowed for the first comparison between the same APSE haplotype from divergent strains of *H. defensa*. Despite the very close relationship we still reported only 99.76% similarity between the two strains showing that despite a relatively short amount of time these haplotypes are no longer completely the same. With the sequencing of more *H. defensa* strains and their associated APSE haplotypes we could

be able to construct a much better phylogenetic tree and better understand how phage spread through symbiotic bacteria populations and influence their evolution.

Our results also revealed differences between APSE haplotypes in the relative number of phage genomes to *H. defensa* in both aphids and in vitro cultures. The largest variation was found in vivo where APSE8 genome copies were 50x more plentiful than *H. defensa* genomes in the bacterial pellet with APSE2 having 16x and APSE3 5x more APSE than *H. defensa*. Weldon et al. (2013) had previously reported differences in the number of APSE genomes per *H. defensa* but in that study they were unable to determine whether this was due to differing *H. defensa* mortality caused by infection, differing levels of integration by APSE haplotypes, or concatenated APSE genomes held within the bacteria. Our results suggest that *H. defensa* mortality due to infection alone is unlikely the cause of these differences as only APSE3 caused any reduction in the number of *H. defensa* upon infection and we could detect no difference in the absolute number of *H. defensa* in established cultures or from adult aphids. Furthermore, we were able to capture TEM micrographs of all APSE haplotypes infecting *H. defensa* from aphids, but never found APSE in the process of lysing an *H. defensa*. Given these results it is more likely the increased abundance of APSE genomes is due to the accumulation of either encapsidated or concatenated virus within *H. defensa*.

While in vitro cultures of *H. defensa* have made it possible to uncover new insights into APSE and *H. defensa* relationships we did find some differences between in vitro and in vivo conditions. Specifically, we found that APSE2 and APSE8 have significantly fewer APSE genome copies per *H. defensa* in vitro and an even smaller percentage of those APSE genomes are found

as protected virus particles free in the medium. In fact, we recorded a dramatic shift in the number of these free particles upon transition from in vivo to in vitro systems. The first generation in vitro NY26 (APSE2) and ZA17 (APSE8) strains of *H. defensa* show a pronounced spike of free APSE upon entering culture, but subsequent generations have a very low percentage of APSE present as free particles. It is possible that this spike in free particles is due to massive *H. defensa* lysis from the stress of moving to culture conditions and thus releasing accumulated APSE. Another possibility is that lysogenic forms of APSE2 and 8 are more fit under culturing conditions and thus outcompete lytic forms much like P1, P2, and Mu lysogens of *E. coli* reproduce more rapidly than non-lysogens in glucose limited conditions (Edlin et al. 1977). Additionally, APSE8 showed higher transcriptional activity across the genome in vivo. In contrast, APSE3 was not significantly different at any point in APSE genome copy number, percent free particles, or expression level for almost all genes.

Given these results perhaps it is unsurprising that APSE3 is once again different from the other two haplotypes in its ability to infect the A2C strain of *H. defensa*. APSE3 was the only haplotype capable of infecting A2C, but surprisingly only when APSE3 was sourced from established in vitro cultures. APSE isolated from aphid hemolymph was never able to infect A2C in vitro cultures. It is unknown why we were never able to successfully infect A2C with aphid derived APSE3 considering it was found to be both genetically and transcriptionally identical to in vitro derived APSE3. These results only further highlight how little we know about APSE and other bacteriophage reproduction and infectivity.

Considering these somewhat conflicting results we developed primers for the integration sites of APSE and a centrifuge and filtering regime to isolate APSE from *H. defensa* to better understand how APSE haplotypes differed in their level of integration. While we found that most APSE is integrated into the *H. defensa* genome no matter the haplotype at least a small portion of the population also existed in a nonintegrated state. More importantly though, we found that APSE is capable of specialized transduction. Specialized transduction differs from general transduction in that the pieces of bacterial genomes directly flanking the integration site are packaged along with the virus genome while generalized transduction is mistakenly packaged bacterial DNA into the viral capsid (Cavanagh & Miller 1986, Snyder et al. 2013). While there have been several studies showing bacteriophage from the family *Podoviridae* can perform generalized transduction, such as P22 and F116, there have been few examples of specialized transduction in the family (Ebel-Tsipis et al. 1972, Byrne & Kropinski 2005, Pourcel et al. 2017). Specialized transduction could be important for many possible reasons. It could be a pathway for the introduction of genes or morons (bits of bacterial DNA that do not make a functional product) or it could be used as a defense against other invading bacteriophages (Weisberg 1996, Ravin et al. 2000, Cheng-Hung et al. 2012). Understanding the role and frequency of this transduction could give us valuable insights to APSE and *H. defensa* evolution and how this complex tripartite association was initiated and how it is maintained. We also present the first evidence that APSE might form a circularized genome which would be unique amongst *Podoviridae* (Lavigne et al. 2008). Overall, our results illuminate previously unknown details about APSE haplotype diversity, reproduction, and

maintenance within *H. defensa*. We also uncover differences between in vitro and in vivo systems which could inform future studies using in vitro systems to understand complex relationships.

3.6 Acknowledgements

We thank Mary Ard in the University of Georgia Microscopy facility for assistance with transmission electron microscopy. We would also like to thank Kyungsun Kim for aid in rearing and collecting aphids for genome and transcriptome samples. This work was supported by a grant from the US National Science Foundation (IOS 1256794) to KMO and MRS.

3.7 References

Rohwer, F. 2003. Global Phage Diversity. *Cell*. **113**, 141. (doi: 10.1016/S0092-8674(03)00276-9)

Deresinski, S. 2009. Bacteriophage Therapy: exploiting smaller fleas. *Clinical Infectious Diseases.* **48**, 1096-1101. (doi: 10.1086/597405)

Clokie, M.R., Millard, A.D., Letarov, A.V., Heaphy, S. 2011. Phages in Nature. *Bacteriophage*. **1**, 31-45. (doi: 10.4161/bact.1.1.14942)

Brussow, H., Canchaya, C., Hardt, W.D. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiology and Molecular Biology Reviews*. **68**, 560-602. (doi: 10.1128/MMBR.68.3.560-602.2004)

Miller, R.V. & Day, M.J. 2008. Contribution of lysogeny, pseudolysogeny, and starvation to phage ecology. *Bacteriophage Ecology: Population Growth, Evolution, and Impact of Bacterial Viruses*. 1st edition, 114-144. Cambridge University Press. Cambridge, United Kingdom. (doi: 10.1017/CBO9780511541483.008)

Metcalf, J.A. & Bordenstein, S.R. 2012 The complexity of virus systems: the case of endosymbionts. *Current Opinion in Microbiology.* **15**, 546-552. (doi: 10.1016/j.mib.2012.04.010).

Nanda, A.M., Thormann, K., Frunzke, J. 2015. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *Journal of Bacteriology*. **197**, 410-419. (doi: 10.1128/JB.02230-14)

Barksdale, L. & Arden, S.B. 1974. *Persisting bacteriophage infections, lysogeny, and phage conversions. Annual Reviews of Microbiology*. **28**, 265-299. (doi: 10.1146/annurev.mi.28.100174.001405)

Kutter, E. & Sulakvelidze, A. 2004. *Bacteriophages: Biology and Applications*. 1st Edition. CRC Press. Boca Raton, Florida. (ISBN: 9780849313363)

Calendar R. 2005. *The Bacteriophages.* 2nd Edition. Oxford University Press. New York, New York. (ISBN: 9780195148503)

Feiner, R., Argov, T., Rabinovich, L., Sigal, N., Borovok, I., Herskovits, A.A. 2015. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nature Reviews Microbiology*. **13**, 641-650. (doi: 10.1038/nrmicro3527)

Davies, E.V., James, C.E., Williams, D., O'Brien, S., Fothergill, J.L., Haldenby, S., Paterson, S., Winstanley, C., Brockhurst, M.A. 2016. Temperate phages both mediate and drive

adaptive evolution in pathogen biofilms. *Proceedings of the National Academy of Science of the United States of America*. **113**, 8266-8271. (doi: 10.1073/pnas.1520056113)

Oliver, K.M., Degnan, P.H., Hunter, M.S. & Moran, N.A. 2009 Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* **325**, 992-994. (doi:10.1126/science.1174463).

Brandt, J.W., Chevignon, G., Oliver, K.M., Strand, M.R. 2017. Culture of an Aphid Symbiont Demonstrates its Direct Role in Defense Against Parasitoids. *Proceedings of the Royal Society of London B: Biological Sciences*. **284**, 20171925. (doi: 10.1098/rspb.2017.1925).

Weldon, S., Strand, M. & Oliver, K. 2013 Phage loss and the breakdown of a defensive symbiosis in aphids. *Proceedings of the Royal Society of London B: Biological Sciences* **280**, 20122103. (doi: 10.1098/rspb.2012.2103)

van der Wilk, F., Dullemans, A.M., Verbeek, M. & van den Heuvel, J. 1999 Isolation and characterization of APSE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrthosiphon pisum*. *Virology* **262**, 104-113. (doi:10.1006/viro.1999.9902)

Clark, A.J., Inwood, W., Cloutier, T., Dhillon, T.S. 2001. Nucleotide sequence of coliphage HK620 and the evolution of lambdoid phages. *Journal of Molecular Biology*. **311**, 657-679. (doi: 10.1006/jmbi.2001.4868)

Lima-Mendez, G., Van Helden, J., Toussaint, A., Laplae, R. 2008. Reticulate representation of evolutionary and functional relationships between phage genomes. *Molecular Biology and Evolution*. **25**, 762-777. (doi: 10.1093/molbev/msn023)

Lavigne, R., Seto, D., Mahadevan, P., Ackermann, H.W., Kropinski, A.M. 2008. Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools. *Research in Microbiology*. **159**, 406-414. (doi: 10.1016/j.resmic.2008.03.005)

Fouts, D.E., Lumpp, J., Bishop-Lilly, K.A., Rajavel, M., Willner, K.M., Butani, A., Henry, M., Biswas, B., Li, M., Albert, M.J., Loessner, M.J., Calendar, R., Sozhamannan, S. 2013. Whole genome sequencing and comparative genomic analyses of two Vibrio cholerae O139 Bengal-specific Podoviruses to other N4-like phages reveal extensive genetic diversity. *Virology*. **10**, 165. (doi: 10.1186/1743-422X-10-165)

Lopes, A., Tavares, P., Petit, M.A., Guerois, R., Zinn-Justin, S. 2014. Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics*. **15**, 1027. (doi: 10.1186/1471-2164-15-1027)

Degnan, P.H. & Moran, N.A. 2008 Diverse phage-encoded toxins in a protective insect endosymbiont. *Applied and Environmental Microbiology* **74**, 6782-6791. (doi:10.1128/aem.01285-08).

Chevignon, G., Boyd, B.M., Brandt, J.W., Oliver, K.M., Strand, M.R. 2018. Culture-Facilitated Comparative Genomics of the Facultative Symbiont Hamiltonella defensa. *Genome Biology and Evolution*. **10**, 786-802. (doi: 10.1093/gbe/evy036)

Oliver, K.M., Noge, K., Huang, E.M., Campos, J.M., Becerra, J.X., Hunter, M.S. 2012. Parasitic wasp responses to symbiont-based defense in aphids. *BMC Biology*. **10**, 11. (doi: 10.1186/1741-7007-10-11)

Martinez, A.J., Weldon, S.R., Oliver, K.M. 2014. Effects of parasitism on aphid nutritional and protective symbioses. *Molecular Ecology*. **23**, 1594-1607. (doi: 10.1111/mec.12550)

Doremus, M.R. & Oliver, K.M. 2017. Aphid Heritable Symbiont Exploits Defensive Mutualism. *Applied Environmental Microbiology*. **83**, e03276-16. (doi: 10.1128/AEM.03276-16)

Katoh, K. & Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution.* **30**, 772-780. (doi: 10.1093/molbev/mst010)

Kearse M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. **28**, 1647-1649. (doi: 10.1093/bioinformatics/bts199)

Langmead, B. & Salzberg, S.L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. **9**, 357-359. (doi: 10.1038/nmeth.1923.)

Anders, S., Pyl, P.T., Huber, W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. **31**, 166-169. (doi: 10.1093/bioinformatics/btu638)

Love, M.I., Huber, W., Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. **15**, 550. (doi: 10.1186/s13059-014-0550-8)

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Wu, P.Y. & Wang, M.D. 2016 The Selection of Quantification Pipelines for Illumina RNA-seq Data Using a Subsampling Approach. *IEEE EMBS Int Conf Biomed Health Inform* **2016**, 78-81. (doi: 10.1109/BHI.2016.7455839)

Moran, N.A., Degnan, P.H., Santos, S.R., Dunbar, H.E. & Ochman, H. 2005 The players in a mutualistic symbiosis: Insects, bacteria, viruses, and virulence genes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 16919-16926. (doi:10.1073/pnas.0507029102).

Degnan, P.H. & Moran, N.A. 2008A Evolutionary genetics of a defensive facultative symbiont of insects: exchange of toxin-encoding bacteriophage. *Molecular Ecology* **17**, 916-929. (doi:10.1111/j.1365-294X.2007.03616.x).

van den Heuvel, J.F., Hummelen, H., Verbeek, M., Dullemans A.M., van der Wilk F. 1997 Characteristics of acyrthosiphon pisum virus, a newly identified virus infecting the pea aphid. *Journal of Invertebrate Pathology.* **70**, 169-176. (doi: 10.1006/jipa.1997.4691)

Oliver, K.M., Moran, N.A. & Hunter, M.S. 2005 Variation in resistance to parasitism in aphids is due to symbionts not host genotype. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 12795-12800. (doi:10.1073/pnas.0506131102).

Little, J.W. & Michalowski, C.B. 2010 Stability and Instability in the Lysogenic State of Phage Lambda. *Journal of Bacteriology* **192**, 6064-6076. (doi: 10.1128/JB.00726-10)

King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. 2011. Podoviridae. in *Virus Taxonomy.* 63-85. Elsevier. (ISBN: 9780123846846).

Botstein, D., Waddell, C.H., King, J. Mechanism of Head Assembly and DNA Encapsulation in *Salmonella* Phage P22. *Journal of Molecular Biology*. **80**, 669-695.

Streisinger, G., Emrich, J., Stahl, M. 1967 Chromosome structure in phage T4.III. Terminal redundancy and length determination. *Proceedings of the National Academy of Sciences of the United States of America.* **57**, 292.

Edlin, G, Lin, L., Bitner, R. 1977 Reproductive fitness of P1, P2, and Mu lysogens of Escherichia coli. *Journal of Virology*. **21**, 506-564.

Cavanagh, M.M. & Miller, R.V. 1986 Specialized Transduction of Pseudomonas aeruginosa PAO by Bacteriophage D3. *Journal of Bacteriology*. **165**, 448-452.

Snyder, L., Peters, J.E., Henkin, T.M., Champness, W. 2013 Lysogeny: the λ Paradigm and the Role of Lysogenic Conversion in Bacterial Pathogenesis. *Molecular Genetics of Bacteria (4th ed.)*. 340–343. Washington, DC: ASM Press. (ISBN 9781555816278).

Ebel-Tsipis, J., Fox, M.S., Botstein, D. 1972 Generalized transduction by bacteriophage P22 in Salmonella typhimurium. II. Mechanism of integration of transducing DNA. *Journal of Molecular Biology*. **71**, 449-469.

Byrne, M. & Kropinski, A.M. 2005 The genome of the Pseudomonas aeruginosa generalized transducing bacteriophage F116. *Gene*. **346**, 187-194. (doi: 10.1016/j.gene.2004.11.001).

Pourcel, C., Midoux, C., Hauck, Y., Vergnaud, G., Libera, L. 2017 Large Preferred Region for Packaging of Bacterial DNA by phiC725A, a Novel Pseudomonas aeruginosa F116-Like Bacteriophage. *PLOS One*. **12**, eo169684. (doi: 10.1371/journal.pone.0169684).

Weisberg, R.A., 1996 Specialized transduction. *Escherichia coli and Salmonella: cellular and molecular biology*, **2nd edn, vol 2**, 2442-2448. ASM Press, Washington DC.

Cheng-Hung, L., Pei-Yu, C., Chiou-Ying, Y., Nien-Tsung, L. 2012 Genome, Integration, and Transduction of a Novel Temperate Phage of Helicobacter pylori. *Journal of Virology.* **86**, 8381-8392. (doi:10.1128/JVI.00446-12).

CHAPTER 4

IDENTIFICATION OF SMALL FACTORS IN A TRI-PARTITE SYMBIOSIS RESPONSIBLE FOR KILLING PARASITOID EMBRYOS

[1] Brandt, J.W., Eum, J., Chevignon G., Oliver, K.M., Brown, M.R., Strand, M.R. To be submitted to:.

4.1 Abstract

The pea aphid, *Acyrthosiphon pisum*, has become a model system for ecology and evolution of symbiosis. This is especially true of the tri-partite symbiosis between the pea aphid and the protective symbiont *Hamiltonella defensa* and its bacteriophage APSE. Together *H. defensa* and APSE impart a protective phenotype which kills the parasitoid wasp *Aphidius ervi* early on in its development inside the pea aphid host. Previously, we had developed an in vitro system for growing *A. ervi* embryos in the presence or absence of *H. defensa* outside of their aphid host. Here we use this simplified system and our ability to control APSE infection status of *H. defensa* to better identify the mechanism by which the APSE-*H. defensa* symbiosis kills wasp larvae early on in embryogenesis.

4.2 Introduction

Recently, there has been an explosion in our understanding of how important and diverse insect-microbe interaction can be to the survival and fitness of insects throughout nature. These interactions can range from temporary gut residence to maternally inherited bacterial symbionts and have been shown to impact all facets of insect biology as well as population dynamics (Douglas 1998, Moran et al. 2008, Oliver et al. 2010, Oliver & Martinez 2014). Among these interactions the pea aphid, *Acyrthosiphon pisum,* and its many possible bacterial partners has emerged as a model for these interactions (Moran et al. 2005, Brisson & Stern 2006). The pea aphid carries an obligate nutritional symbiont as well as several facultative symbionts that provide

the pea aphid with important ecological traits while not being necessary for development and reproduction in all conditions (Montllor et al. 2002, Oliver et al. 2003, Scarborough et al. 2005, Lukasik et al. 2013, Asplen et al. 2014, Heyworth & Ferrari 2015). One such species is the protective symbiont *Hamiltonella defensa* which is one of the most commonly found bacterial associates in pea aphids (Oliver et al. 2003). On average, 50% of individuals within an aphid population are infected with *H.* defensa (Russell, 2013, Oliver et al. 2014, Zytenska 2016). The pea aphid-*H. defensa* system is ideal for studying insect-symbiont interactions because the pea aphid can reproduce asexually, and single matrilines can be generated which are genetically identical and harbor only a single strain of *H. defensa*. Using these matrilines Oliver et al. (2003) found that there is variation among *H. defensa* haplotypes in their ability to protect their aphid hosts from the parasitoid *Aphidius ervi*. Further investigation of these isolated *H. defensa* strains revealed that they were infected with a bacteriophage, named APSE, and that different haplotypes of this virus correlated with differing levels of protection (Oliver et al. 2009, Degnan et al. 2008). Sequencing of these APSE haplotypes showed that they share a core genome with one large variable region of 3.7-5.8 Kb which contains putative eukaryotic toxin genes. This large region, known as the toxin cassette, accounts for most of the variability in APSE gene content and expression between haplotypes (Degnan et al. 2008, Chevignon et al. 2018).

To date, 9 APSE haplotypes have been found in *H. defensa* with some haplotypes providing nearly complete parasitoid protection to moderate levels of protection (Oliver et al. 2014). Additionally, uninfected *H. defensa* provide no protection to their aphid host (Weldon et al. 2013). Therefore, it has been hypothesized that protection is due to some phage encoded protein or factor.

Further strengthening this hypothesis is the fact that different haplotypes are nearly identical to one another outside of a variable region which carries different putative eukaryotic toxins. Levels of protection have also been shown to be consistent across haplotype and toxins carried by APSE. The best studied haplotypes of APSE have been those carried by North American pea aphids, APSE2, APSE3, and APSE8 (Oliver et al. 2009, Weldon et al. 2013, Martinez et al. 2016, Brandt et al. 2017, Chevignon t al. 2018). Among these haplotypes APSE3 confers nearly complete resistance to *A. ervi* and encodes a putative Serine-Aspartic acid-repeat toxin gene (YDr). In contrast, ASPE2 and APSE8 confer moderate levels of protections (i.e. 40-60% aphid mortality upon parasitism) and carry a cytolethal distending toxin subunit B (cdtB) gene in their toxin cassette (Degnan et al. 2008, Oliver et al. 2009, Martinez et al. 2016).

While cdtB has been well described as a eukaryotic toxin with DNase like activity YDr has had fewer studies linking it to eukaryotic cell death and with no clear consensus on its mode of action. However, prior studies do link YDr proteins produced by bacteria in the genus *Photorhabdus*, which are symbionts of the entomopathogenic nematodes, with eukaryotic toxicity and insecticidal activity (Sergeant et al. 2003). YDr proteins contain five major domains: a trafficking domain, the large titular YD-repeat domain, a transmembrane domain, a releasing peptidase, and a toxin domain. The toxins found in YDr complexes are diverse and include deaminases, nucleases, and peptidases (Zhang et al. 2012). The two most common functions of YDr proteins are in self-recognition or as a part of the ABC complex of secreted toxins that contribute to eukaryotic pathogenicity (Zhang et al. 2012, Koskiniemi et al. 2013, Hachani et al. 2014). In this way some YDr proteins are remarkably similar to teneurins and has led to the

hypothesis that when similar YDr proteins interact the C-terminus toxin domain is not released but are released in heterotypic interactions (Ferralli et al. 2018). Structural analysis of these proteins reveal that the YD repeat domain forms a hollow barrel structure that surrounds the toxin domain until unknown structural changes occur and the toxin is released (Busby et al. 2013, Feralli et al. 2018). In this way the bacteria are protected from the toxin until structural changes occur and the toxin is either released into extracellular space or after uptake of the YD barrel into a target eukaryotic cell (Busby et al. 2013).

Presently, we still do not know how or even if the YDr protein expressed by APSE3 infected *H. defensa* kill *A. ervi* larvae. Recent advances in our ability to manipulate complex symbiotic relationships like this have us poised to experimentally link YDr proteins and wasp embryo mortality or alternative modes of action responsible for host aphid protection. Using our previously established in vitro conditions for bioassays we aimed to better understand the mechanism and underlying factor(s) responsible for *A. ervi* mortality in this complex system.

4.3 Materials and Methods

(a) Insect rearing and *H. defensa* cultures. *A. pisum* lines of identical genetic background were previously established from single parthenogenetic females that hosted the phage-free A2C, APSE3-infected AS3, APSE2-infected NY26 and APSE8-infected ZA17 strains of *H. defensa* (Oliver et al. 2009; 2012; Martinez et al. 2014; Doremus & Oliver 2017). *A. ervi* used in the study were reared in the laboratory as a large population founded from commercially purchased (Arbico)

and field-collected wasps on an *A. pisum* line with no facultative symbionts; adult wasps were provided continuous access to honey and water at 20° C under a 16 light: 8 h dark photoperiod (Martinez et al. 2014). In vitro cultures were established from these aphid lines and maintained in 25 cm^2 culture flasks (Falcon) under ambient atmosphere at 27° C in 5 ml of TC100 medium (Sigma) plus 10% of fetal bovine serum (FBS) (Hyclone) (Brandt et al. 2017). Cultures were passaged weekly after growing to an average maximum density of 1 x 10^{10} bacteria per ml.

(b) *A. ervi* assays. *A. pisum* without facultative symbionts were individually parasitized by *A. ervi.* Aphids were surface sterilized as previously described in Brandt et al. 2017 2 hours post-parasitism and then dissected to isolate either newly laid *A. ervi* eggs or 1st instar larvae. One egg or larva was placed per well of a 96-well culture plate that contained 100 µl TC100 + 30% FBS. AS3 from 3-day-old established cultures were then added to wells at a density of $1x10^6$ bacteria per ml as a positive control for all bioassays. *A. ervi* eggs were allowed to develop for 72 h at 27°C and embryonic stages of *A. ervi* were recorded every 24 h.

(c) Western blots. *H. defensa* was collected either from 3-day-old in vitro cultures or from single adult aphids. Adult aphids were surface sterilized and dissected into 100 µl sterile PBS before the carcass was removed and aphid cells separated by brief spinning on a microcentrifuge. The resulting supernatant and 100 µl of culture medium containing *H. defensa* were then centrifuged at 8000g for 10 min at 4°C to pellet *H. defensa*. The supernatant was removed and placed into clean 1.5 ml centrifuge tubes and put on ice for later processing. Resulting bacterial pellets were

then washed 3 times with sterile PBS and repelleted by centrifugation before being resuspended in 100 µl PBS. 100 µl of 2x SDS Page buffer was then added to each sample and boiled for 5 min. to denature proteins. 20 µl of samples and 5 µL of PageRuler Plus Prestained Ladder (ThermoScientific) were then loaded into 1-4% gradient PROTEAN precast mini polyacrylamide gels run on Mini-PROTEAN® Tetra Vertical Electrophoresis Cell(BioRad) at 200 V for 1 hour. Two primary antibodies for YD-toxin were separately produced with two peptides (KETWSEPGEKGGTAC and ASEKTVREGKTHLSC) synthesized and purified by GenScript USA. Two-hundred µg of the peptide in 0.5 ml of phosphate buffer (pH 7.4) was mixed with 1.0 ml of Freund's complete adjuvant. New Zealand white rabbits were immunized three times for 2 weeks with the prepared antigen. The rabbit serum was collected 7 days after the final immunization and stored at -80°C. Proteins were then transferred to PVDF transfer membranes using a Mini-Trans blot Cell (Biorad) at 30 V constant 0.9 mA overnight in 4°C cold room. Following transfer membranes were blocked with 2% instant skim milk + washing solution for 1 hour before primary antibody against YDr (antirabbit) was added at 1:5000 ratio and incubated overnight at 4° C while gently shaking. Membranes were then washed 3x with washing solution before secondary antibody was added at 1:10,000 and incubated at 4°C for 2 hours before washing membrane 3 more times. Bands were then visualized using Clarity ECL blotting substrates (Biorad) at 5 min exposure times.

(d) Molecular weight cutoff filters. Fifteen ml of 3-day-old *H. defensa* cultures were centrifuged at 8000xg for 10 min at 4°C to pellet bacteria. Supernatants were then transferred to 100 kDa

Molecular weight cutoff Amicon Ultra centrifuge filter tubes (Millipore) and centrifuged at 4000xg using an Eppendorf 5804 swinging bucket rotor for 1 hour. This process was repeated until only 1 ml of supernatant was left above the filter at which point it was removed and placed in a new 1.5 ml centrifuge tube for later use in bioassays. Similarly, 4 ml of filtrate was reserved for use in bioassays. The remaining filtrate was moved to 30 kDa Amicon filters and the above process was repeated. Finally, the last remaining filtrate was run through a 10 kDa Amicon filter following the same protocol. Concentrated samples were resuspended into new TC100 + 30% FBS to a physiological dose before being assayed. Samples that were not concentrated were replenished with FBS as filtering removed this important component after the first centrifugation.

(e) Boiling and Proteinase K treatment. The resulting ≤30 kDa MW fraction from filtering was then subjected to 2 different treatments. The first sample was boiled for 5 min and then centrifuged at 4000xg for 5 min to remove denatured and coagulated proteins and the supernatant was moved to a clean 15 ml centrifuge tube where new FBS was added to replenish what was lost through filtering and boiling. The second set of samples were incubated with Proteinase K for 1 hour at 50°C. Samples were then filtered through another 30 kDa Amicon centrifuge filter to remove the proteinase before being replenished with FBS and bioassayed.

(d) High Performance Liquid Chromatography. Another set of ≤30 kDa samples were generated and subjected to fractioning by HPLC. Samples were first lyophilized before being rehydrated with 0.1% TFA in 80% acetonitrile/pH2O. Semi-prep-High Performance Liquid

Chromatography was then performed using a Phenomenex Jupiter C18 300 Å 10 mm x 250 mm column with guard cartridge which was equilibrated prior to sample injection. Two ml/min fractions were collected into large polypropylene tubes starting at 0 min. for use in bioassays. Eluates were then pooled into groups for bioassay at physiological doses as well as a single 1:10 dilution rehydrated with fresh TC100 + 30% FBS. The pooled group which showed the most activity was then selected to test single 2 ml/min fractions for activity against *A. ervi* development.

(e) LC-MS analysis. Four 2ml/min fractions were selected for LC-MS analysis. Two of these fractions showed activity in our bioassays while the other two showed no activity and were used to narrow mechanism candidate lists. LC-MS was performed by the University of Georgia Proteomics and Mass Spectometry facility in Athens, Ga on an Orbitrap Elite LC-MS machine (ThermoScientific). Proteomic data was then analyzed using the Thermo Scientific Protein Discover 1.4 software utilizing the Sequest Database and in-house MASCOT database search engines with our newly constructed APSE3 and *H. defensa* strain AS3 genome manually added to each engine. Results were then collated and sorted by presence in active and non-active fractions and mapped back to full protein sequences.

(f) Recombinant HP12 and YDr. We generated recombinant HP12 by bacterial expression using the full-length ORF which was PCR-amplified using the primers APSE_HP12_FWD (5' - ATG ACT GAA GCA ACA CAA TCC G-3') and APSE_HP12_FWD (5'- TTA TTT ATA TCT ATC

GGT TAT TTT GAT ATA TAG-3'). The resulting product was then directly cloned with pET-APSE_HP12_FWD (5'-GAC GAC GAC AAG ATG ACT GAA GCA ACA CAA TCC G-3') and pET-APSE_HP12_RVS (5'-GAG GAG AAG CCC GGT GCT TTA TAT CTA TCG GTT ATT TTG ATA TAT AG-3') in frame with the C-terminal His-tag of the vector pET30 Ek/LIC (Novagen) using T4 polymerase. The pET30-HP12 construct was confirmed by DNA sequencing, and then expressed by transforming into *Escherichia coli* BL21 (DE3) cells cultured in SOC medium (0.5% Yeast Extract, 2% Tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM Glucose), supplemented with 10 mg/ml of kanamycin to an O.D. of 1.0 at 37 °C. One milliliter of 0.1 mM isopropyl-b-D-thiogalactopyranoside (IPTG) was added to 400 ml of culture and incubated for an additional 18 h at 20 °C. Bacterial cells were harvested by centrifugation at 5000 x g for 10 min and used immediately or stored at -80 °C.

For bacterial expression of truncated YD-toxin, the sequence corresponding to toxin domain (amino acid 1532 to 1683) was PCR-amplified using the primers of YDrpt_APSE_TXN_FWD (5'-CTG GGT GCG GAT ACA AAA ATA AAC TGT G-3') and YDrpt_APSE_RVS (5'-TCA CAA ATA AGA TAT AGG TTT TAT CAC ATC GAA GA-3'). The resulting product was then directly cloned into pET30 Ek/LIC vector (Novagen) with two primers, pET-YDrpt-TOXIN_FWD (5'-GAC GAC GAC AAG ATG CTG GGT GCG GAT ACA AAA ATA AAC T-3') and pET-YDrpt-TOXIN_RVS (5'-GAG GAG AAG CCC GGT CAC AAA TAA GAT ATA GGT TTT ATC ACA TCG-3') using T4 polymerase. The pET30-YDtxn construct was confirmed by DNA sequencing, and then expressed by transforming into *Escherichia coli* BL21

(DE3) cells cultured in SOC medium (0.5% Yeast Extract, 2% Tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM Glucose), supplemented with 10 mg/ml of kanamycin to an O.D. of 1.0 at 37 °C. One milliliter of 0.1 mM isopropyl-b-D-thiogalactopyranoside (IPTG) was added to 400 ml of culture and incubated for an additional 18 h at 16 °C. Bacterial cells were harvested by centrifugation at 5000 x g for 10 min and used immediately or stored at -80 °C.

Bacterial pellets were then resuspended in lysis buffer (50 mM NaH2PO4, 300 mM NaCl, 10 mM imidazole) on ice for 1 h followed by addition of lysozyme (1 mg/ml) in 50 mM Tris buffer (pH 8.0) and sonication. After centrifugation of the lysate at 10,000 x g for 25 min, supernatants were bound to a Ni-NTA matrix (5 Prime) pre-equilibrated with lysis buffer. Bound proteins were washed 5x with wash buffer (50 mM NaH2PO4, 300 mM NaCl, 20 mM imidazole) followed by elution with three column volumes of elution buffer (50 mM NaH2PO4, 300 mM NaCl, 40-300 mM imidazole). The eluted proteins from wash fractions were loaded onto 4~20% SDS-PAGE gels and visualized by coomassie blue staining. Proteins eluted using 300 mM imadizole were desalted and concentrated by Centricon 30 (Millipore, USA). Protein sequence identities were confirmed using LC-MS by the UGA Proteomics and Mass Spectometry Facility. Recombinant HP12 was confirmed to have >99.99% identity, but we were unable to purify any products that aligned to truncated YDr.

4.4Results

(a) YD-repeat is found only inside bacteria. Our previous work using *A. ervi* bioassays revealed that *Hamiltonella defensa* infected with APSE3 was capable of secreting some factor that killed wasp embryos prior to their development to the morula stage (Brandt et al. 2017). In order to test the long-standing hypothesis that YD-repeat toxicity is the underlying cause of *A. ervi* mortality we first created a polyclonal antibody against the N-terminus end of YDr. *H. defensa* cells were separated from aphid hemolymph and culture medium and both bacterial pellets and resulting supernatants were tested for the presence of YDr. Interestingly, western blots showed no signal from either aphid hemolymph or culture media while bacterial pellets of *H. defensa* carrying APSE3 had a strong single band at around the predicted molecular weight of YDr, 180 Kda (Fig. 4.1). These results suggest that YDr is not secreted into the medium or aphid hemolymph and are present as membrane bound proteins. Alternatively, YDr could act very similar to other proteins in its family and was membrane bound until the c-terminus toxin domain (CTD) was cleaved from the protective barrel domain and transported outside of the bacteria where it would interact with the developing wasp egg. We were unable to test this possibility as our antibody did not recognize the c-terminus end of YDr.

(b) Factor(s) responsible for killing *A. ervi* are ≤30 Kda. Considering we were unable to recover a signal from YDr in the aphid hemolymph and culture medium we moved on to better understanding the factor or factors involved in *A. ervi* mortality. Previous work with in vitro assays showed that the AS3 line of *H. defensa* was able to condition medium to have lethal effects on *A. ervi* eggs. That is to say that APSE3 infected *H. defensa* were able to secrete some factor that was

toxic to *A. ervi* (Brandt et al. 2017). This led us to believe that we could isolate the activity for better characterization of the factor(s) responsible and determine if the CTD toxin of YDr was possibly involved.

Using our previously reported genome of APSE3 and protein prediction software, RaptorX, we identified that our CTD toxin was approximately 30 kDa in weight barring any unpredicted modifications (Kallberg et al. 2012, Chevignon et al. 2018). We first obtained conditioned culture medium from 6-day-old cultures of APSE3 infected *H. defensa* (strain AS3) and separated bacterial cells by centrifugation. We then took the resulting supernatant and ran it through a series of molecular weight cutoff (MWC) centrifuge filters and tested each fraction for *A. ervi* killing activity. Single *A. ervi* eggs were placed in a well with 100 µl of medium and their development was measured over 72 hours. Results show that large molecular weight fractions (≥100 kDa) had no effect on wasp development while the 30-10 kDa fractions and ≤ 10 kDa fractions both significantly reduced the proportion of developed wasp embryos (Fig. 4.2). Similar experiments were conducted with uninfected *H. defensa* (strains A2C and AS3-N) which had no effect on *A. ervi* development in any of the MWC fractions (Supp. Fig. 4.1).

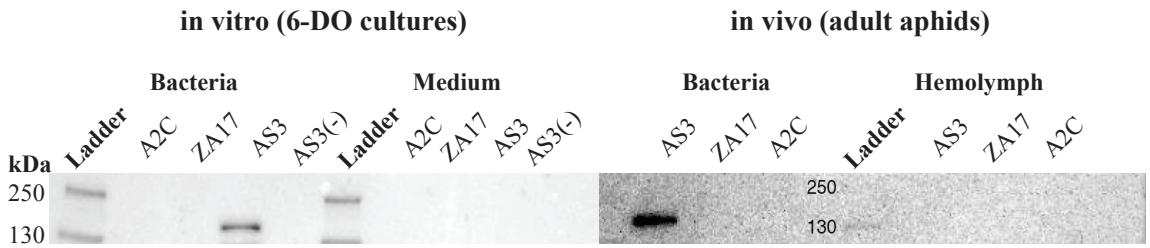


Figure 4.1. YDr is present in *H. defensa* as a membrane-bound protein. Western blot using polyclonal antibody against N-terminus of YDr shows that YDr is only present in *Hamiltonella defensa* carrying APSE3 both in vitro (left) and in vivo (right). YDr is only detected in bacterial pellets (Bacteria) and cannot be detected in either culture medium or aphid hemolymph.

(c) Wasp killing is reduced by heat treatment and proteinase activity. The leading hypothesis for the mechanism of aphid protection is that toxins produced by APSE3 infected *H. defensa* directly kill or interfere with *A. ervi* development early on in their development (Oliver et al. 2009, Weldon et al. 2013). Most data points to the putative toxin YDr being the causative agent given that there is little difference between haplotypes of APSE outside of their toxin cassette (Degnan and Moran 2008, Chevignon et al. 2018). Our previous work with the line AS3 and APSE3 also supports that ASPE3 infected *H. defensa* secrete factors into culture medium/aphid hemolymph that interferes with *A. ervi* development and eventually kills the developing embryo (Brandt et al. 2017). However, there has been little other information directly pointing to proteins secreted by *H. defensa* are responsible for this activity.

Therefore, we used our ability to generate conditioned medium capable of killing developing *A. ervi* to test parameters that would suggest secreted proteins were responsible for wasp mortality. Specifically, we tested the heat stability of the killing activity as well as its sensitivity to proteinase K treatment. Boiling samples has been shown to degrade smaller proteins

and denaturing and possibly inactivating larger proteins. Proteinase K is a broad-spectrum protease which has been shown to inactivate nucleases and destroy proteins in cell lysates. However, if neither of these treatments reduced *A. ervi* mortality in our bioassays then factors other than proteins could be responsible. Results show that boiling the active MWC fractions increased the proportion of *A. ervi* that developed to the 128-cell stage (Fig. 4.3A). Likewise, Proteinase K treatment increased the proportion of wasps developing past this critical stage (Fig. 4.3B). Sensitivity to Proteinase K treatment and boiling suggests that protection and therefor, wasp mortality, is due to toxic proteins secreted or released by APSE3 infected *H. defensa*.

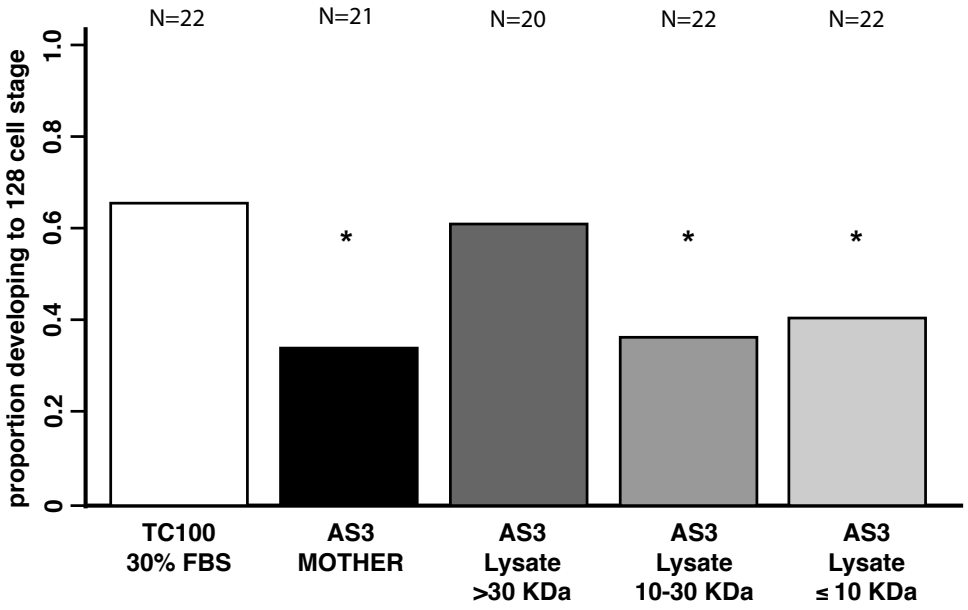


Figure 4.2 APSE3 infected *H. defensa* produced factors ≤30 kDa that prevented *A. ervi* development. Proportion of *A. ervi* eggs that successfully developed past the 64-cell stage in the presence of molecular weight fractions from APSE3 infected *H. defensa* conditioned medium. Fractions containing factors greater than 30 kDa (AS3 Lysate > 30 kDa) had no effect on proportion of *A. ervi* eggs developing to the 128-cell stage while fractions ≤30 kDa (AS3 Lysate 10-30 kDa, AS3 Lysate ≤10 kDa) significantly reduced the number of eggs which developed. Asterisks indicate that the treatment was significantly different from the control group (TC100 30% FBS). N represents the number of individuals in each treatment. X^2=16.93; df=4; p≤0.001 followed by post-hoc pairwise Fisher Exact Test. Asterisks indicate treatment is significantly different from control of TC100 +30% FBS.

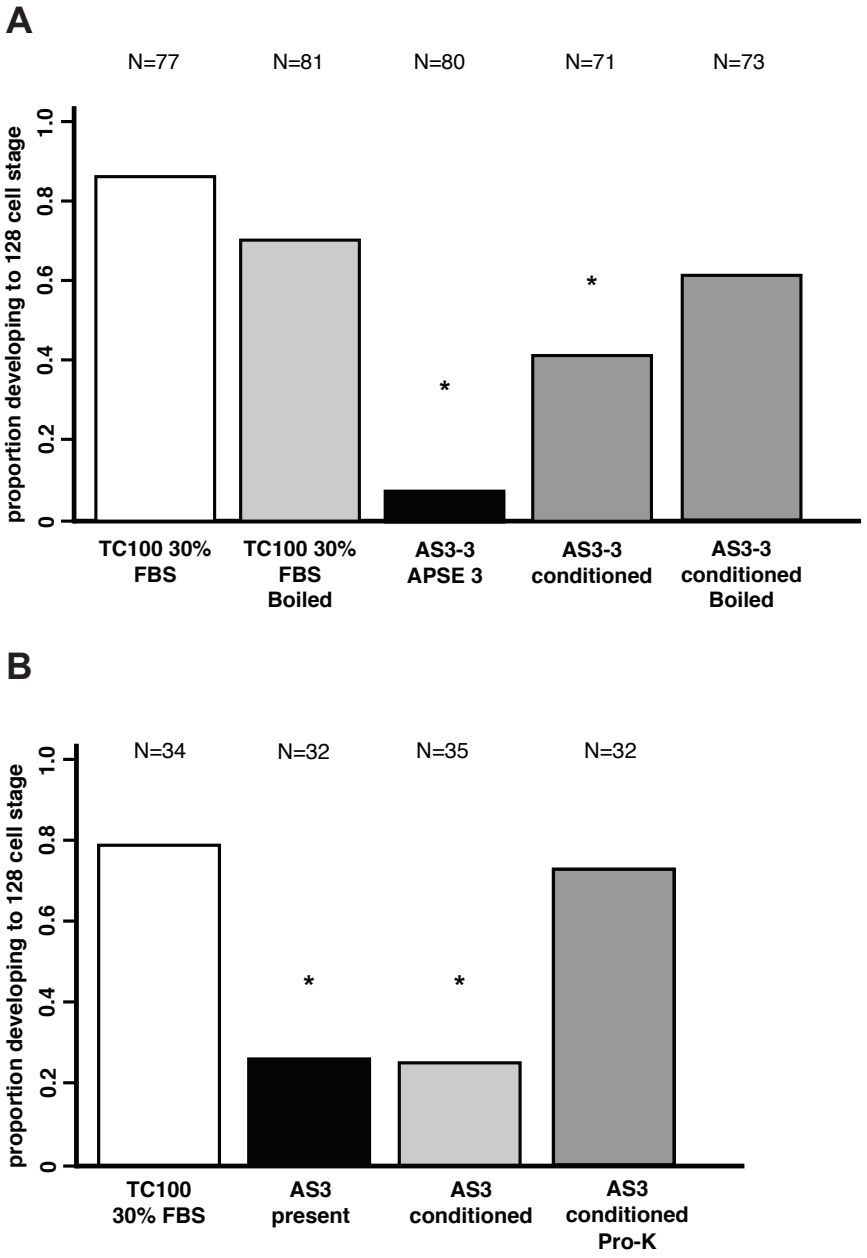


Figure 4.3 Boiling and Proteinase K treatment reduced *A. ervi* killing activity. Proportion of *A. ervi* developing to the critical 128-cell stage. Boiling the active 30 kDa MWC fraction for 5 minutes significantly reduced the killing activity and similar proportions of *A. ervi* were able to develop to or past the 128-cell stage (A). X^2=116.99; df=4; $p \leq 0.001$ followed by post-hoc pairwise Fisher Exact Test. Similarly, Proteinase K treatment of 50 µg/ml incubated at 50°C for 30 minutes reduced killing activity to that of control non-conditioned TC100+30% FBS medium (B). X^2=33.78; df=3; $p \leq 0.001$ followed by post-hoc pairwise Fisher Exact Test. Asterisks indicate significant difference from the control TC100 + 30% FBS.

(d) HPLC purification reveals HP12 as possible candidate for *A. ervi* mortality. Considering factor(s) produced by APSE3 infected *H. defensa* were most likely proteins or small peptides our next step in identifying the underlying mechanism was to further separate the active MWC fractions using HPLC and testing the resulting fractions against *A. ervi* development (Supp. Fig. 4.2). In summary, bioassays of grouped fractions found that a group of fractions (time 14-16 min) that centered around a large peak in the HPLC readout had the most activity against *A. ervi* development. Out of these 4 fractions we found that fractions 15 and 16 had the most activity against the developing parasitoids. All four fractions were then sent to the Georgia Proteomics and Mass Spectrometry Facility and run on the Orbitrap Elite (ThermoScientific) to identify potential proteins or peptides responsible for *A. ervi* death. A list of all peptide sequences and predicted alignment to the AS3 strain of *H. defensa* including APSE3 can be found in Supplemental Table 1. Hits that where unique to active fractions from AS3 and ASPE3 conditioned cultures aligned to various pseudogenes, structural phage proteins, proteins involved in *H. defensa* metabolism, and interestingly, a hypothetical protein with unknown function, HP12 (Supp. Table 4.2). We were never able to recover any hits the reliably and exclusively aligned with YDr. However, our results did reveal to us another possible candidate for testing with our in vitro assay: Hypothetical Protein 12 (HP12). Recombinant HP12 was generated and isolated from *E. coli* (K-12 strain) and was added to 100 ml culturing medium at concentrations of 1, 10, and 100 ng/ml. Bioassay results show that recombinant HP12 has no effect on *A. ervi* development from newly laid eggs to the 128-cell stage (Fig 4.4). Unfortunately, we were unable to generate any recombinant YDr protein due to its large size. Furthermore, we were unable to generate even the smaller predicted toxin

domain of YDr which is about 30 kDa in size. This could have been due to possible acute toxicity to *E. coli* or its predicted exceptionally hydrophobic nature. While we were unable to directly test a recombinant version of YDr it remains a strong candidate for the mechanism of *A. ervi* mortality given the number of hits we recorded in active HPLC fractions and that MWC fractions reveal that activity is found around 30 kDa or the approximate size of the predicted toxin domain in Ydr.

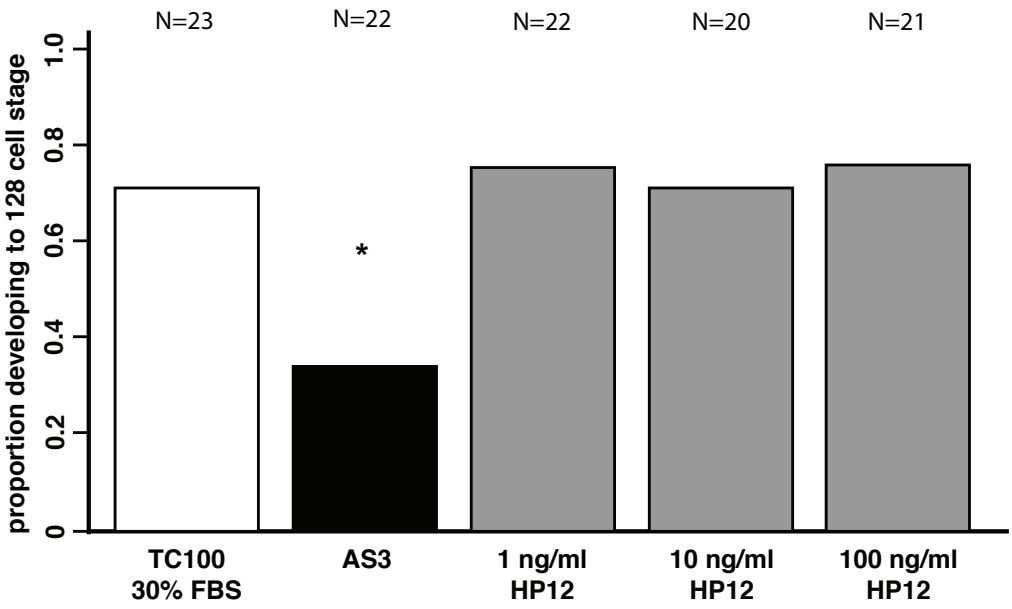


Figure 4.4 HP12 has no effect on *A. ervi* development. Proportion of *A. ervi* eggs developing to the 128-cell stage in the presence of 1, 10, or 100 ng/ml recombinant HP12. Even at high concentrations HP12 has no effect on *A. ervi* development during early embryogenesis and is unlikely to be directly responsible for APSE3-mediated defense in the pea aphid. TC100 + 30% FBS was used as a negative control while AS3 bacteria were used as a positive control for *A. ervi* development. X^2=18.73; df=4; p≤0.001 followed by post-hoc pairwise Fisher Exact Test. Asterisks indicate significant difference from control TC100 + 30% FBS only.

4.5 Discussion

The pea aphid, *Acyrthosiphon pisum*, has emerged as a leading model organism in the quest to understand the evolution and ecology of insect microbial symbionts. The pea aphid is an ideal organism because not only does its novel reproductive strategy make it easy to create clonal matrilines, but it also plays host to a primary nutritional symbiont and an array of diverse secondary symbionts (Moran et al. 2005, Brisson and Stern 2006). The best studied of these secondary symbionts protects its aphid host from the parasitoid wasp *Aphidius ervi* by killing the wasp larvae as it develops inside the aphid. This bacterium, known as *Hamiltonella defensa*, only confers this protection when it is infected with bacteriophage APSE (Oliver et al. 2009, Weldon et al. 2013). Furthermore, different haplotypes of APSE have been shown to confer different levels of protection of which APSE3 has been shown to be the most protective haplotype commonly occurring in North American pea aphids (Russell, 2013, Oliver et al. 2014, Zytenska 2016). APSE3 primarily differs from other, less protective, haplotypes in its toxin cassette (Degnan and Moran 2008, Chevignon et al. 2018). The toxin cassette of APSE3 is composed of a hypothetical protein, a holin, a lysozyme, and a putative eukaryotic toxin from the Serine-Aspartic acid repeat toxin family (YDr) (Chevignon et al. 2018). Given that almost all other genes and even gene expression is conserved across haplotypes of APSE it is believed that YDr is probably responsible for killing *A. ervi*. However, there has been little experimental evidence linking YDr produced by APSE3 to wasp death.

In our previous study we outlined methods for functional bioassays that tested *A. ervi* development in the presence of only APSE3 and *H. defensa*. Using these assays, we were able to show that not only did APSE3 and *H. defensa* kill developing wasp embryos in vitro, but also *H. defensa* conditioned medium still had activity when bacteria where removed from the culture. Here we expanded upon this work and showed that YDr is almost certainly responsible for host aphid protection. YDr is a large protein with four major domains: an unknown domain, the large titular yd repeat domain, a transmembrane domain, and a c-terminus toxin domain (CTD) (Fig. 4.5A). While YDr has much in common with the other Polymorphic toxins it shares a family with it is missing 2 features commonly found in this family: An N-terminus tracking domain and a releasing peptidase just before the toxin domain (Zhang et al. 2012). 3D protein predictions of YDr also highlight similarities to other polymorphic toxins showing that the YD-repeat domain forms a barrel-like structure which encompasses the CTD to prevent toxicity to non-targets (Fig. 4.5B). With this knowledge we set out to test the long-standing hypothesis that YDr is directly responsible for killing developing wasp embryos using both in vitro and in vivo approaches.

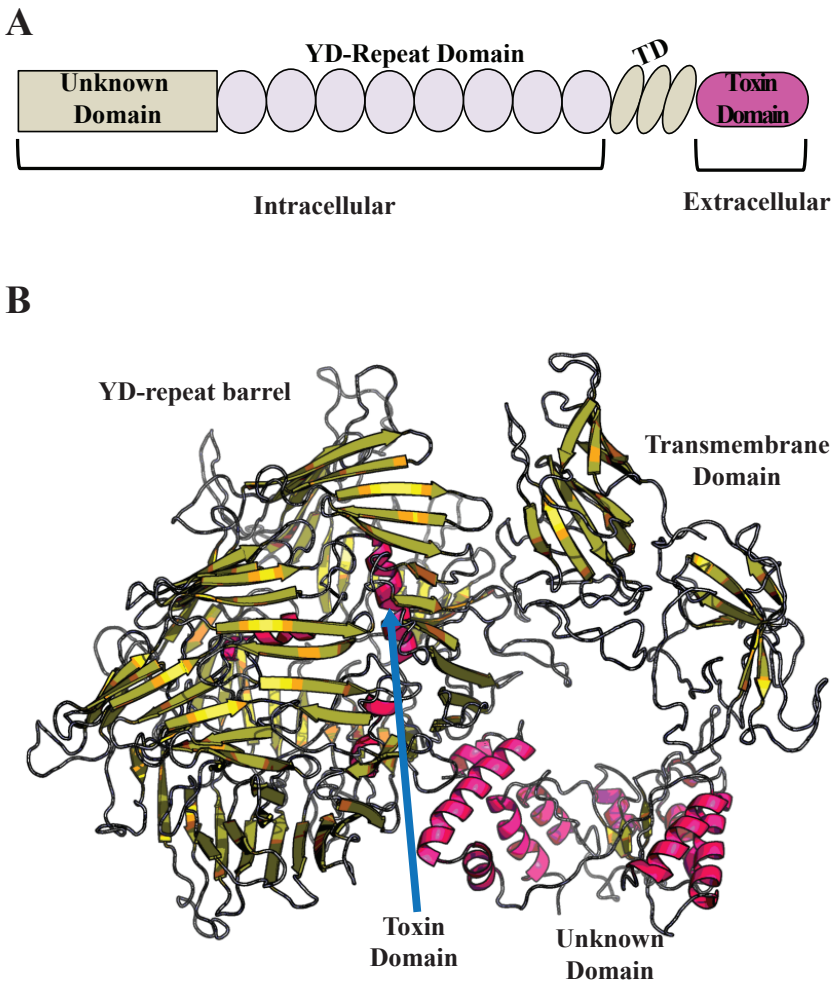


Figure 4.5. Schematic and 3D representations of YDr. Schematic representation of YDr showing the 4 main domains (A). 3D predicted structure using RaptorX (B). Arrow points to buried Toxin domain within the YD-repeat barrel

First, we developed a polyclonal antibody to YDr to determine if YDr is secreted by *H. defensa* into culture medium or aphid hemolymph. Western blots of 3 different strains of *H. defensa* carrying different APSE haplotypes revealed that only APSE3 carrying *H. defensa* produce YDr, but that YDr is membrane bound in both in vitro and in vivo systems. These results do not preclude YDr as being responsible for *A. ervi* death however because our antibody only picked up the N-terminus domain of YDr. As previously stated, it is commonly accepted that proteins of this family are membrane bound until activated by heterotypic interactions or upon cell

death. When activated the CTD is transported out of the cell and cleaved from the protective barrel domain (Zhang et al. 2012, Busby et al. 2013, Koskiniemi et al. 2013, Hachani et al. 2014, Ferralli et al. 2018).

Given these results we thought it prudent to better understand the nature of the factor that was secreted and killed the wasp embryo in vitro. As such, we designed several descriptive experiments to test the parameters of the active factor. Our first step was to determine the size of the factor responsible for wasp mortality. Using culturing medium conditioned by APSE3 infected *H. defensa* that was run through a series of centrifuge molecular weight cutoff filters we found that activity was only in fractions that held products of ≤30kDa. This is much smaller than the complete YDr complex, but still larger than the CTD which was estimated to be roughly 28 kDa in size. In order to test the theory that the smaller cleaved CTD could be responsible we first tested parameters that would preclude products other than proteins as the possible source of activity. Two common ways of doing this is to test the effects of heat and proteinase treatments on activity. Results showed that both boiling active fractions and Proteinase K treatment reduced activity in our bioassays and allowed a normal proportion of wasp embryos to develop.

Our final step was to identify the responsible factor(s) and test them directly in our bioassays. In order to complete this final step, we needed to create a purer sample for identification using an Orbitrap Elite LC-MS (ThermoScientific). Using HPLC we separated the ≤30 kDa fraction and tested each resulting fraction in vitro. These efforts resulted in 2 HPLC fractions showing activity and these plus 2 non-active fractions were used for identification of unique sequences to be analyzed. Proteomics analysis revealed two possible candidates: a hypothetical

protein of unknown function found in the APSE3 toxin cassette (HP12) and multiple fragments of Ydr. The next logical step was to create recombinant versions of these two proteins and directly test their effect on wasp development in vitro. While we were able to successfully clone and isolate recombinant HP12 we were unable to produce any Ydr. This was possibly due to YDr being a relatively large protein, so we instead attempted to produce only the CTD from YDr. Once again, we were unable to generate any recombinant protein possibly due to the extreme hydrophobic nature of CTD or due possible acute toxicity to *E. coli* used to produce the recombinant protein. Despite this setback, we were able to bioassay HP12 and determine that it had no effect on *A. ervi* development in vitro.

These results along with proteomic analysis that revealed only fragments of YDr as the only other plausible source of activity strongly suggest that it is responsible, at least in part, for the death of *A. ervi* in the presence of APSE3 infected *H. defensa*. While we would have liked to test this hypothesis directly, we believe that our results strongly support the standing hypothesis that YDr is directly responsible for the protective phenotype found in APSE3/*H. defensa* carrying aphids.

4.6 Acknowledgements

We would like to thank Kyungsun Kim and Jena Johnson for aide in rearing insects and experimental setup. We would also like to thank Chau-Wen Chou at the UGA Proteomics and Mass Spectrometry facility for all her hard work and advice with proteomics analysis. This work was supported by a grant from the US National Science Foundation (IOS 1256794) to KMO and MRS.

4.7 References

Douglas, A. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria Buchnera. *Annual Review of Entomology*. **43**, 17-37. (doi: 10.1146/annurev.ento.43.1.17).

Moran, N.A., McCutcheon, J.P., Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. *Annual Reviews of Genetics*. **42**, 165-190. (doi: Annurev.Genet.41.110306.130119).

Oliver, K.M., Degnan, P.H., Burke, G.R., Moran, N.A. 2010. Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annual Review of Entomology*. **55**, 247-266. (doi: 10.1146/annurev-ento-112408-085305).

Oliver, K.M. & Martinez, A.J. 2014. How resident microbes modulate ecologically-important traits of insects. *Current Opinion in Insect Science*. **4**, 1-7. (doi:10.1016/j.cois.2014.08.001).

Moran, N.A., Russell, J.A., Koga, R., Fukatsu, T. 2005. Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Applied and Environmental Microbiology*. **71**, 3302-3310. (doi:10.1128/AEM.71.6.3302-3310.2005).

Brisson, J.A. & Stern, D.L. 2006. The pea aphid, Acyrthosiphon pisum: an emerging genomic model system for ecological, developmental, and evolutionary studies. *Bioessays*. **28**, 747-755. (doi:10.1002/bies.20436).

Montllor, C.B., Maxmen, A., Purcell, A.H. 2002. Facultative bacterial endosymbionts benefit pea aphids, Acyrthosiphon pisum, under heat stress. *Ecological Entomology*. **27**, 189-195. (doi:10.1046/j.1365-2311.2002.00393.x).

Oliver, K.M., Russell, J.A., Moran, N.A., Hunter, M.S. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 1803-1807. (doi:10.1073/pnas.0335320100).

Scarborough, C.L., Ferrari, J., Godfray, H.C.J. 2005. Aphid protected from pathogen by endosymbiont. *Science*. **310**, 1781. (doi:10.1112/science.1120180).

Lukasic, P., van Asch, M., Guo, H., Ferrari, J., Godfray, H.C.J. 2013. Unrelated facultative endosymbionts protect aphids against a fungal pathogen. *Ecology Letters*. **16**, 214-218. (doi:10.111/ele.12031).

Asplen, M.K., Bano, N., Brady, C.M., Desneux, N., Hopper, K.R., Malouines, C., Oliver, K.M., White, J.A., Heimpel, G.E. 2014. Specialisation of bacterial endosymbionts that protect aphids from parasitoids. *Ecological Entomology*. **39**, 736-739. (doi:10.1111/een.12153).

Heyworth, E. & Ferrari, J. 2015. A facultative endosymbiont in aphids can provide diverse ecological benefits. *Journal of Evolutionary Biology*. **28**, 1753-1760. (doi:10.1111/jeb.12705).

Russell, J.A., Weldon, S., Smith, A.H., Kim, K.L., Hu, Y., Lukasik, P., Doll, S., Anastopoulos, I., Novin, M., Oliver, K.M. 2013. Uncovering symbiont-driven genetic diversity across North American pea aphids. *Molecular Ecology*. **22**, 2045-2059. (doi:10.1111/mec.12211)

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

Zytynska, S.E. & Weisser, W.W. 2016. The natural occurance of secondary bacterial symbionts in aphids. *Ecological Entomology*. **41**, 13-26. (doi:10.1111/een.12281).

Degnan, P. & Moran, N.A. 2008. Evolutionary genetics of a defensive facultative symbiont of insects: exchange of toxin-encoding bacteriophage. *Molecular Ecology*. **17**, 916-929. (doi:10.1111/j.1365-294X.2007.03616.x).

Chevignon, G., Boyd, B.M., Brandt, J.W., Oliver, K.M., Strand, M.R. 2018. Culture-Facilitated Comparative Genomics of the Facultative Symbiont Hamiltonella defensa. *Genome Biology and Evolution.* **10**, 786-802. (doi: 10.1093/gbe/evy036)

Weldon, S., Strand, M. & Oliver, K. 2013 Phage loss and the breakdown of a defensive symbiosis in aphids. *Proceedings of the Royal Society of London B: Biological Sciences* **280**, 20122103. (doi: 10.1098/rspb.2012.2103).

Martinez, A.J., Doremus, M.R., Kraft, L.J., Kim, K.L. & Oliver, K.M. 2017 Multi-modal defenses in aphids offer redundant protection and increased costs likely impeding a protective mutualism. *Journal of Animal Ecology*. (doi:10.1111/1365-2656.12675).

Brandt, J.W., Chevignon, G., Oliver, K.M., Strand, M.R. 2017. Culture of an Aphid Symbiont Demonstrates its Direct Role in Defense Against Parasitoids. *Proceedings of the Royal Society of London B: Biological Sciences*. **284**, 20171925. (doi: 10.1098/rspb.2017.1925).

Sergeant, M., Jarrett, P., Ousley, M., Morgan, J.A.W., 2003. Interactions of Insecticidal Toxin Gene Products from *Xenorhabdus nematophilus* PMFI296. *Applied and Environmental Microbiology*. **69**, 3344-3349. (doi: 0.1128/AEM.69.6.3344–3349.2003).

Zhang, D., de Souza, R. F., Anantharaman, V., Iyer, L. M. and Aravind, L. 2012. Polymorphic toxin systems: comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biology Direct*. **7**, 18. (doi: 10.1186/1745-6150-7-18).

Koskiniemi, S., Lamoureux, J.G., Nikolakakis, K.C., t'Kint de Roodenbeke, C., Kaplan, M.D., Low, D.A., Hayes, C.S. 2013. Rhs proteins from diverse bacteria mediate intercellular competition. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 7032-7037. (doi: 10.1073/pnas.1300627110)

Hachani, A., Allsopp, L.P., Oduko, Y., Filloux, A. 2014. The VgrG proteins are "A la carte" delivery systems for bacterial type VI effectors. *Journal of Biological Chemistry*. **289**, 17872-17884. (doi:10.1074/jbc. M114.563429).

Feralli, J., Tucker, R.P., Chiquet-Ehrismann, R. 2018. The teneurin C-terminal domain possesses nuclease activity and is apoptogenic. *Biology Open*. **7**, bio031765. (doi: 10.1242/bio.031765).

Busby, J.N., Panjikar, S., Landsberg, M.J., Hurst, M.R.H., Lott. J.S. 2013. The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. *Nature*. **501**, 547-550. (doi: 10.1038/nature12465).

Martinez, A.J., Weldon, S.R. & Oliver, K.M. 2014 Effects of parasitism on aphid nutritional and protective symbioses. *Molecular Ecology* **23**, 1594-1607. (doi:10.1111/mec.12550)

Källberg, M., Wang, H., Wang, S., Peng, J., Z., Wang, Z., Lu, H., Xu, J. 2012. Template-based protein structure modeling using the RaptorX web server. *Nature Protocols* **7**, 1511–1522. (doi: 10.1038/nprot.2012.085)

APPENDIX A

SUPPORTING INFORMATION FOR CHAPTER 2

A.1. Supplemental Tables and Figures

Table S2.1. Insect culture media and established insect cell lines tested for primary culture of *H. defensa*[a]

Medium	*H. defensa* strain tested
Grace's	A2C
Grace's + 10% FBS	A2C
Leibowitz (L15)	A2C
Leibowitz (L15) + 10% FBS	A2C
Leibowitz (L15) + 10% FBS and C6/36 cells (*Aedes albopictus*)	A2C
Schneider's	A2C
Schneider's + 10% FBS	A2C
Schneider's + 10% FBS and S2 cells (*Drosophila melanogaster*)	A2C
Sf900II	A2C
Sf900II + 5% FBS	A2C
Sf900II + 5% FBS and CiE1 cells (*Chrysodeixus includes*)	A2C
Sfx	A2C
Sfx and S2 cells	A2C
Sfx + 2% FBS	A2C
Sfx + 2% FBS and Aag2 cells (*Aedes aegypti*)	A2C
TC100	A2C, AS3, NY26, ZA17
TC100 +10% FBS	A2C, AS3, NY26, ZA17
TC100 +10% FBS and Sf-21 cells (*Spodoptera frugiperda*)	A2C, AS3, NY26, ZA17
TC100 +10% FBS and TN5 cells (*Trichoplusia ni*)	A2C, AS3, NY26, ZA17

[a] Grace's, Leibowitz, Schneider's, Sf900 and Sfx media were purchased from Gibco TC100 was purchased from Sigma, and fetal bovine serum was purchased from Atlanta Biologicals. The BTI-TN-5B1-4 (TN5) cell line established from the moth *Trichoplusia ni* was maintained in TC100 plus 10% (FBS) [1], the C6/36 cell line from the mosquito *Aedes albopictus* was maintained in Leibowitz (L15) medium (Sigma) plus 10% FBS [2], the S2 cell line from Drosophila melanogaster was maintained in Sfx medium [3], the Sf-21 cell line from the moth *Spodoptera frugiperda* was maintained identically to TN5 cells [4], the CiE1 cell line from the moth *Chrysodeixus includens* was maintained in Sf900 medium plus 5% FBS [5], and the Aag2 cell line from the mosquito *Aedes aegypti* was maintained in SFX medium supplemented with 2% FBS [2]. Primary cultures for the A2C strain of *H. defensa* were initiated as by collecting haemolymph from *A. pisum* as described in the Methods followed by transfer to different media with or without FBS and particular insect cell lines. Results identified TC100 plus 10% FBS and TN5 cells as the preferred conditions for growth of A2C *H. defensa* (see Results). Primary cultures of the AS3, NY26 and ZA17 strains *of H. defensa* were therefore initiated using only TC100 plus 10% FBS and TN5 cells. Established cultures of each *H. defensa* strain were similarly maintained in TC100 plus 10% FBS with or without TN5 cells (see Results).

Table S2.2. qPCR and diagnostic primers.

Primer Name	Strain specificity	Primer sequence	Gene name	Gene function	Primer position	Fragment size
qPCR_DNAK_F	*H. defensa*	CGGTTTATGATCTCGGAGGA	/	DNAK	/	94bp
qPCR_DNAK_R		TTGGTTGCCAACACTTCAAA		DNAK	/	
DNAK2_F	*H. defensa*	TGAAACACTTGTCGGTCAGC	/	DNAK	/	1599bp
DNAK2_R		TCACCTTTTATGGCGGTTTC		DNAK	/	
Buchnera dnaK-F1	*B. aphidicola*	ACGTTCACCCTGAAGCACGTGTA	/	DNAK	/	217bp
Buchnera dnaK-R1		TGGAGCTGCAGTACAGGGAGGA		DNAK	/	
ZA17_F1	ZA17	TACGTACTGTGATGCGCTGG	BJP43_04620	Hypothetical Protein	896465-896484	1000bp
Za17_R1		AATTGAGCAAGTGAAGCGTA	BJP43_04625	IS6 family transposase	897445-897464	
NY26_F1	NY26	TTTTTGGGGCGCAGACATTG	BJP44_10070	RTX Toxin	1784243-1784262	1036bp
NY26_R1		GTCACTCAGCTGCTCCTACC	BJP44_10065	IS630	1783227-1783246	
YD_APSE-3.2_F	APSE-3	GAAAATCGCTGCAACAGACA	BJP42_07995	YD Repeat	1450850-1450869	1799bp
YD_APSE-3.2_R		GAGGTGGTAGGTCCCACTGA			1452629-1452648	
cdtB_APSE-2.2_F	APSE-8/APSE-2	ATGGTGCTCGTTGGAATGAT	/	CdtB	/	801bp
cdtB_APSE-2.2_R		ATGGTGCTCGTTGGAATGAT			/	
Head-protein_APSE.2_F	APSE	AGAGACGGGTGACATTACGG	/	Major Head Protein	/	1048bp
Head-protein_APSE.2_R		GCAGACATAAGCAGGCAACA			/	
APSE-2_3000_10000_2_F	APSE-2 / APSE-3 / APSE-8	CCATCCAGCCCACCATTGAT	/	/	/	APSE-2: 3218bp / APSE-3: 1069bp / APSE-8: 2300bp
APSE-2_3000_10000_4_R		GGCTGGCTCAGGATTTGTCT			/	

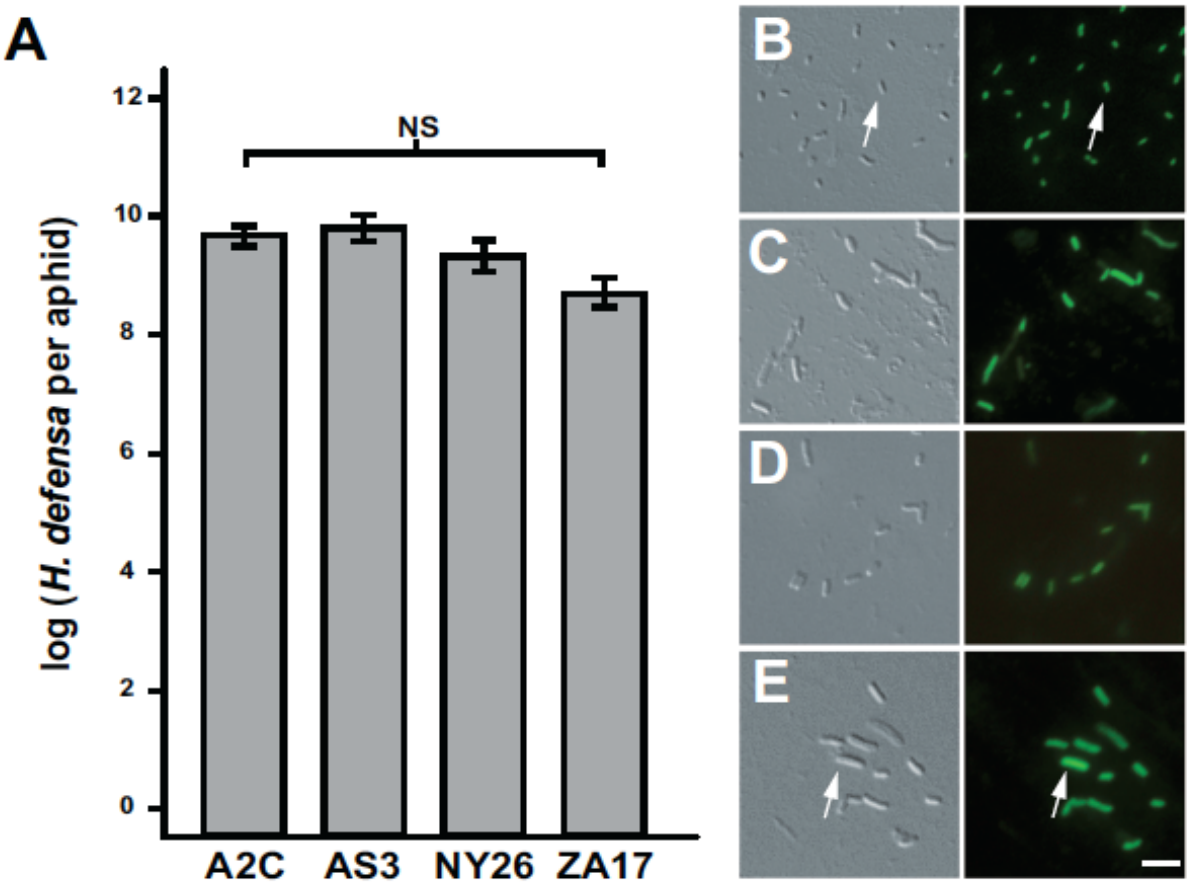


Figure S2.3. Most *H. defensa* in adult *A. pisum* reside extracellularly in haemolymph. (A) Genome copy number (± 95% confidence intervals) per aphid for the A2C, AS3, NY26, and ZA17 strains of *H. defensa* in haemolymph. ANOVA detected no significant differences (NS) in titre between strains ($F_{3,24}$=2.3; p.0.5). (B-E) Differential interference contrast (left) and corresponding epiflourescent (right) micrographs of the A2C (B), AS3(C), NY26 (D), and ZA17 (E) strains in haemolymph stained with acridine orange. Note that most bacteria are rod-shaped (arrows) but vary in size. Scale bar in (E) equals 10 µm.

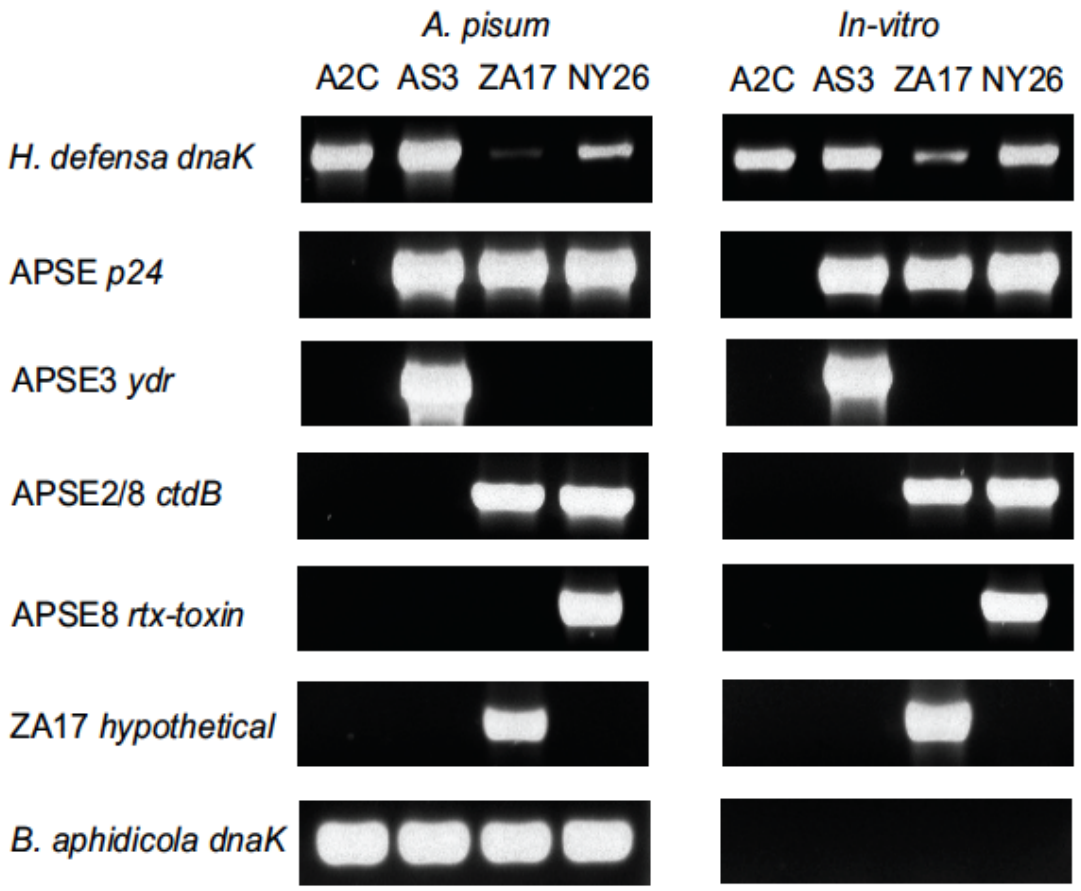


Figure S2.4. Diagnostic PCR assays distinguished the A2C, AS3, NY26, and ZA17 strains of *H. defensa* in *A. pisum* (left) and established in vitro cultures (right). From top to bottom shows PCR products generated using DNA from *A. pisum* or in vitro cultures and primers that amplify unique domains in : 1) the *H. defensa dnaK* gene, 2) the APSE *p24* gene, 3) the *ydr* gene in APSE3 that persistently infects AS3, 4) the *cdtB* gene in APSE2 and APSE8 that persistently infect NY26 and ZA17 respectively, 5) a gene encoding a hypothetical protein unique to ZA17, 6) the *rtx-toxin* gene in APSE8 that infects NY26, and 7) the *B. aphidicola dnaK* gene. A2C is identified by an amplicon for the *H. defensa dnaK* gene but the absence of any APSE-associated amplicons. AS3 is identified by amplicons for the *H. defensa dnaK, p24,* and *ydr* genes. NY26 is identified by amplicons for *H. defensa dnaK, p24, cdtB,* and *rtx-toxin*. ZA17 is identified by amplicons for *H. defensa dnaK, p24, cdtB,* and the hypothetical protein. An amplicon for the *B. aphidicola dnaK* gene was generated using DNA from *A. pisum* but not established cultures of *H. defensa*.

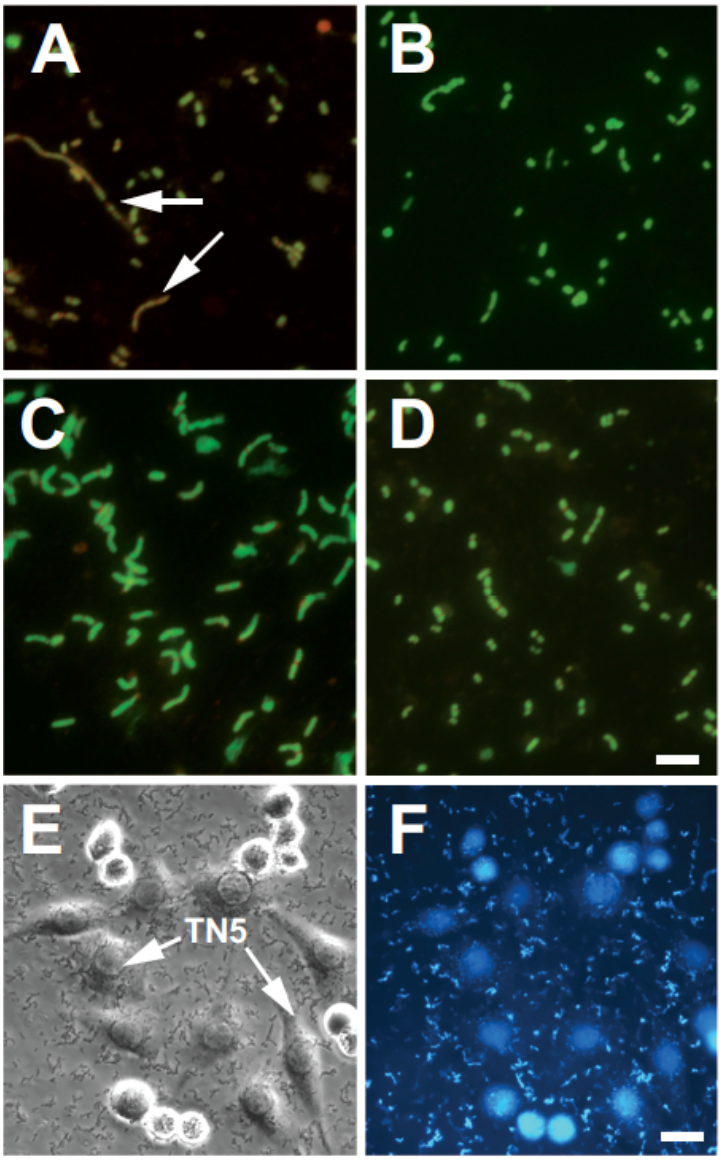


Figure S2.5. Epiflourescent micrographs of the A2C (A), AS3 (B), NY26 (C), and ZA17 (D) strains of *H. defensa* stained with acridine orange. Several of the Bacteria in (A) form chains (arrows) while chains or pairs of bacteria are visible in (B-D). Scale bar in (D) equals 10 µm. A lower magnification phase-contrast (E) and corresponding epiflourescent micrograph (F) of the AS3 strain with TN5 cells after staining with Hoechst 33342. DNA in bacteria and the nuclei of TN5 cells are blue. Bacteria are also extracellular, Scale bar in (F) equals 35 µm.

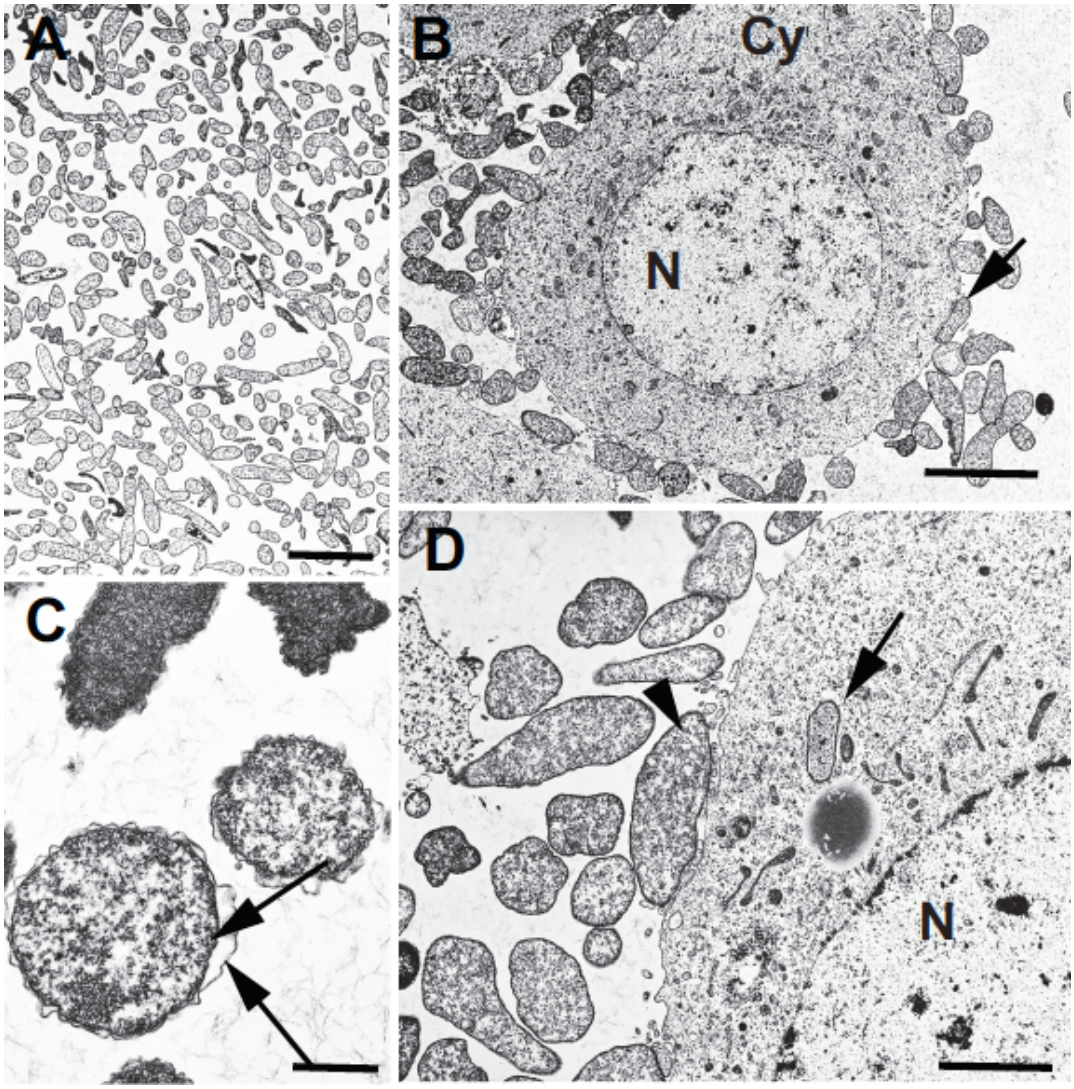


Figure S2.6. Transmission electron microscopy images of *H. defensa* in established cultures. (A) Lower magnification image of the AS3 strain. Note the pleomorphic size and shape of bacteria. Scale bar equals 10µm. (B) Lower magnification of ZA17 in contact with or close proximity (arrow) to a TN5 cell. The cytoplasm (Cy) and nucleus (N) of the TN5 cell are indicated. Scale bar equals 6µm. (C) Higher magnification image of A2C that near coccoid in shape with granular cytoplasm and cell envelopes that consist of two membranes (arrows). Scale bar equals 2 µm. (D) Higher magnification image of ZA17 and TN5 cell. Several bacteria are in close proximity to the surface of the TN5 cell (arrowhead)while one bacterium is in the cytoplasm (arrow). Scale bar equals 4 µm.

A.2. Supplemental References

[1] Beck, M. & Strand, M.R. 2003 RNA interference silences Microplitis demolitor bracovirus genes and implicates glc1.8 in disruption of adhesion in infected host cells. *Virology* **314**, 521-535. (doi:10.1016/s0042-6822(03)00463-x).

[2] Walker, T., Jeffries, C.L., Mansfield, K.L. & Johnson, N. 2014 Mosquito cell lines: history, isolation, availability and application to assess the threat of arboviral transmission in the United Kingdom. *Parasites & Vectors* **7**. (doi:10.1186/1756-3305-7-382).

[3] Vogel, K.J., Brown, M.R. & Strand, M.R. 2015 Ovary ecdysteroidogenic hormone requires a receptor tyrosine kinase to activate egg formation in the mosquito Aedes aegypti. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 5057-5062. (doi:10.1073/pnas.1501814112).

[4] Welburn, S.C., Maudlin, I. & Ellis, D.S. 1987 In vitro cultivation of *Rickettsia*-like-organisms from *Glossina* spp. *Annals of Tropical Medicine and Parasitology* **81**, 331-335.

[5] Johnson, J.A., Bitra, K., Zhang, S., Wang, L.H., Lynn, D.E. & Strand, M.R. 2010 The UGA-CiE1 cell line from Chrysodeixus includens exhibits characteristics of granulocytes and is permissive to infection by two viruses. *Insect Biochemistry and Molecular Biology* **40**, 394-404. (doi:10.1016/j.ibmb.2010.03.005).

APPENDIX B

SUPPORTING INFORMATION FOR CHAPTER 3

Supplemental Table 3.1: *H. defensa* and APSE specific primers

Amplified region	Primer Name	Primer sequence	Fragment size
H. defensa	DNAK_F	CGGTTTATGATCTCGGAGGA	94bp
	DNAK_R	TTGGTTGCCAACACTTCAAA	
APSE	Head-prot_APSE_F	TGGCAACGTATGAAGGGTTT	101bp
	Head-prot_APSE_R	GCAGACATAAGCAGGCAACA	
attL	Integrase_APSE-3_F (R)	CGATATGCACACTTAGCGTCT	245bp
	Integration _5'_APSE_F	AAATTTTCACTGCGCCCTTA	
	Int-5'_204_F	CTCAATTCGATATCTTGATACAAGAATTTG	359bp
	Int-5'_400_F	TTCGATAAGCTCAATGTCACCTTC	555bp
	Int-5'_600_F	AAAGGAGGGGAATTTTATAAAAATAATCGA	755bp
	Int-5'_1000_F	CTAAACCAGACAGATGTGTTTCACC	1155bp
attR	Integration _3'_APSE_F	AGCGAGGAAAGATGGGGTAT	96bp
	Non-Inte_Inte _APSE_R	ATTGGTACGCCCTACTGGATT	
	Int-3'_147_R	GGCACTTTTAAATAATTCCTAGGAGAC	147bp
	Int-3'_200_R	ACTGAGCATGCCAAGAAAATTGAC	200bp
	Int-3'_300_R	GATACAATCAGGTGTTCCACTCTCT	300bp
	Int-3'_400_R	TCAGGGTTGATGATAATAGAGCTTGG	400bp
	Int-3'_700_R	CGAATCGACTGATTTCAGTGGTAA	701bp
	Int-3'_979_R	TACAGAAGATGAAGTCACCCATGT	979bp
	Int-3'_1716_R	GAACTGTATCAGCAATTTAACGCAAG	1716bp
attB	Int_5'_516_H	CTGCGGGCATCAGTATTTTT	1943bp
	Int_3' _1427_H	TGAAAGAAGCCTGCCAATTT	
	Int-5'_400_F	TTCGATAAGCTCAATGTCACCTTC	
	Int-3'_300_R	GATACAATCAGGTGTTCCACTCTCT	604bp
attP	Int_5'_953_A	CAGCGGTACCCGATTAAAAG	2048bp
	Int_3'_1095_A	CGGGCATAATTATTCAGTGG	
	Integration _3'_APSE_F	AGCGAGGAAAGATGGGGTAT	
	Integrase_APSE-3_F (R)	CGATATGCACACTTAGCGTCT	251bp

Supplemental Table 3.2. CDS gene names and predicted functions

Full name	Abbrv.	predicted function
P38: Phage integrase;	P38	Phage integrase
P40: Excisionase;	P40	Excisionase
P41: Dead box helicase;	P41	Dead box helicase
P42: Phage DNA binding protein Roi;	P42	Phage DNA binding protein Roi
P43: Antirepressor;	P43	Antirepressor
C4: ncRNA;	C4	ncRNA
isrK: ncRNA;	isrK	ncRNA
G: KilA-N domain protein (DNA-binding);	G	KilA-N domain protein (DNA-binding)
H: Hypothetical protein;	H	Hypothetical protein
P44: Nuclease;	P44	Nuclease
P45: DNA polymerase I ;	P45	DNA polymerase I
I: Transcriptional regulator;	I	Transcriptional regulator
P47: Phage protein;	P47	Phage protein
P49: Phage protein;	P49	Phage protein
Lys: tRNA Lys;	Lys	tRNA Lys
P50: Phage protein;	P50	Phage protein
P51: Phage protein;	P51	Phage protein
P53: Phage protein;	P53	Phage protein
HP4: Hypothetical protein;	HP4	Hypothetical protein
HP5: Hypothetical protein;	HP5	Hypothetical protein
P1: Phage repressor;	P1	Phage repressor
P2: Transcriptional regulator;	P2	Transcriptional regulator
P3: ATPase;	P3	ATPase
P4: Phage protein;	P4	Phage protein
P5: Antitermination protein Q;	P5	Antitermination protein Q
HP12: Hypothetical protein;	HP12	Hypothetical protein
YD-repeat toxin: YD-repeat toxin and RHS-family protein;	YD-repeat toxin	YD-repeat toxin and RHS-family protein
K: Phage 21-like group II holin;	K	Phage 21-like group II holin
F: Lysozyme;	F	Lysozyme
HP11: Hypothetical protein;	HP11	Hypothetical protein
A: Hypothetical protein;	A	Hypothetical protein
B: Hypothetical protein;	B	Hypothetical protein
CdtB: Cytolethal distending toxin subunit B;	CdtB	Cytolethal distending toxin subunit B
D: Hypothetical protein;	D	Hypothetical protein
P11: Phage holin;	P11	Phage holin
E: Phage holin;	E	Phage holin
HP7: Hypothetical protein;	HP7	Hypothetical protein
HP8: Hypothetical protein;	HP8	Hypothetical protein
Shiga-like toxin: Shiga-like toxin;	Shiga-like toxin	Shiga-like toxin
HP9: Hypothetical protein;	HP9	Hypothetical protein
P13: Lysozyme;	P13	Lysozyme
P14: Phage exported protein;	P14	Phage exported protein
P16: Phage endolysin;	P16	Phage endolysin
P17: Phage protein;	P17	Phage protein
P18: Phage terminase, large subunit;	P18	Phage terminase, large subunit
P19: Phage portal protein;	P19	Phage portal protein
P23: Phage capsid and scaffold protein;	P23	Phage capsid and scaffold protein
P24: Phage major capsid / head protein;	P24	Phage major capsid / head protein
HP19: Hypothetical protein;	HP19	Hypothetical protein
P27: Packaged DNA stabilization protein;	P27	Packaged DNA stabilization protein
P28: DNA stabilization gp10;	P28	DNA stabilization gp10
P30: Phage packaged DNA stabilization protein;	P30	Phage packaged DNA stabilization protein
P31: Hypothetical protein;	P31	Hypothetical protein
P32: Phage DNA transfer protein;	P32	Phage DNA transfer protein
P33: Phage DNA transfer protein;	P33	Phage DNA transfer protein
P35: Phage DNA transfer protein;	P35	Phage DNA transfer protein
P36: Phage tail fibers;	P36	Phage tail fibers
P37: Phage tail fiber assembly protein;	P37	Phage tail fiber assembly protein

Supplemental Table 3.3: Summary of RNAseq data

	AS3	Replicate 1		Replicate 2		Replicate 3	
		Reads	Percentage	Reads	Percentage	Reads	Percentage
In vivo	*A. pisum* genome	13,247,130	54.99%	20,855,026	66.14%	13,395,086	51.18%
	A. pisum mitochondria	247,004	1.03%	419,176	1.33%	457,108	1.75%
	B. aphidicola	157,910	0.66%	120,608	0.38%	227,618	0.87%
	ApV	1,840,504	7.64%	2,054,494	6.52%	1,900,434	7.26%
	AS3 Chromosome	6,684,600	27.75%	6,468,398	20.51%	8,271,894	31.60%
	pHDAS3.1	589,864	2.45%	606,552	1.92%	747,088	2.85%
	pHDAS3.2	60,176	0.25%	63,430	0.20%	78,890	0.30%
	Unmapped	1,261,238	5.24%	944,164	2.99%	1,096,324	4.19%
	Total	24,088,426	100%	31,531,848	100%	26,174,442	100%
	ZA17	Replicate 1		Replicate 2		Replicate 3	
		Reads	Percentage	Reads	Percentage	Reads	Percentage
	A. pisum genome	18,075,160	75.91%	20,036,088	80.90%	22,905,970	72.89%
	A. pisum mitochondria	705,842	2.96%	255,798	1.03%	838,300	2.67%
	B. aphidicola	641,654	2.69%	206,544	0.83%	723,286	2.30%
	ApV	2,434,620	10.23%	2,110,582	8.52%	3,646,758	11.60%
	ZA17 Chromosome	1,220,360	5.13%	1,133,862	4.58%	1,453,580	4.63%
	pHDZA17.1	22,024	0.09%	21,526	0.09%	24,740	0.08%
	pHDZA17.2	11,090	0.05%	11,198	0.05%	13,668	0.04%
	pHDZA17.3	18,930	0.08%	17,864	0.07%	23,470	0.07%
	Unmapped	680,384	2.86%	973,808	3.93%	1,796,588	5.72%
	Total	23,810,064	100%	24,767,270	100%	31,426,360	100%
In vitro subsampling (from Chevignon et al 2017)	AS3	Replicate 1		Replicate 2		Replicate 3	
		Reads	Percentage	Reads	Percentage	Reads	Percentage
	AS3 Chromosome	7,318,614	92.92%	7,305,816	92.76%	7,317,726	92.91%
	pHDAS3.1	386,956	4.91%	343,338	4.36%	409,970	5.21%
	pHDAS3.2	46,214	0.59%	35,912	0.46%	52,830	0.67%
	Unmapped	124,408	1.58%	191,126	2.43%	95,666	1.21%
	Total	7,876,192	100.00%	7,876,192	100.00%	7,876,192	100.00%
	ZA17	Replicate 1		Replicate 2		Replicate 3	
		Reads	Percentage	Reads	Percentage	Reads	Percentage
	ZA17 Chromosome	1,255,718	94.54%	1,263,070	95.09%	1,246,368	93.84%
	pHDZA17.1	23,414	1.76%	20,198	1.52%	34,138	2.57%
	pHDZA17.2	5,828	0.44%	6,246	0.47%	6,718	0.51%
	pHDZA17.3	10,980	0.83%	7,180	0.54%	9,948	0.75%
	Unmapped	32,280	2.43%	31,526	2.37%	31,048	2.34%
	Total	1,328,220	100.00%	1,328,220	100.00%	1,328,220	100.00%

APPENDIX C

SUPPORTING INFORMATION FOR CHAPTER 4

Supplemental Table 4.1. Peptide hits from HPLC fractions.

Sequence (subspecies)	#PSMs	proteins	MH+ (Da)	ΔM (ppm)	Modifications	Fraction of Origin
TDAPS	77	1	490.21478	0.8		**fraction 14**
SEAxS	77	1	490.21478	0.8	x4(p)	
DTVEN	82	1	577.24664	0.38		
ESVEN	82	1	577.24664	0.38		**fraction 14**
TDVEN	82	1	577.24664	0.38		
VESEN	80	1	577.24683	0.7		
ESASP	77	1	490.21466	0.55		**fraction 14**
LSDEN	124	4	577.24664	0.38		**fraction 14**
TVDEN	16	1	577.24664	0.38		
SLDEN	10	2	577.24664	0.38		
ETGPS	66	1	490.21472	0.67		**fraction 14**
ASEPS	67	0	490.21472	0.67		
ADTPS	68	0	490.21472	0.67		
TEGPS	67	0	490.21472	0.67		
						fraction 14
DLSNx	17	1	577.24658	0.28		
SDIEN	46	0	577.24664	0.38		
DSLEN	218	0	577.24670	0.49		
Sequence (subspecies)	**#PSMs**	**proteins**	**MH+ (Da)**	**ΔM (ppm)**	**Modifications**	
LxEGG	278	2	489.23041	0.05	X2(N)	**fraction 15**
LSAGS	101	1	434.22440	-0.38		**fraction 15**
LSGAS	94	0	434.22440	-0.38		
TDVEN	147	1	577.24670			**fraction 15**
DTVEN	148	1	577.24670			
ESVEN	147	1	577.24670			
SVEGA	72	0	462.21967	0.4		**fraction 15**
SVEAG	72	0	462.21967	0.4		
AGSIS	92	1	434.22482	0.6		**fraction 15**
GASLS	85	2	434.22482	0.6		
GLGEN	157	0	489.23059	0.43		**fraction 15**
VGAEN	72	0	489.23059	0.43		
VSEAG	51	0	462.21951	0.07		**fraction 15**
VSEGA	51	0	462.21951	0.07		
VESEN	147	1	577.24640	-0.04		**fraction 15**
ASGLS	79	0	434.22461	0.11		**fraction 15**
EDDPE	81	0	604.20947	-0.39		**fraction 15**
AGISS	20	1	434.22485	0.67		**fraction 15**
GAISx	80	1	434.22485	0.67	X5(S)	
						fraction 15
QESPT	94	1	561.25140	-0.17		

Supplemental Table 4.1. Peptide hits from HPLC fractions (cont.)

GEGTPT	148	1	561.25140	-0.19		**fraction 15**
TVDEN	105	1	577.24652	0.17		**fraction 15**
LSDEN	103	0	577.24670	0.49		
SLDEN	24	2	577.24670	0.49		
TDDIN	62	0	577.24658	0.25		**fraction 15**
DISDQ	48	0	577.24713	1.23		
SDIEN	26	0	577.24640	-0.04		
DSLEN	236	0	577.24670	0.49		**fraction 15**
DLSNx	47	1	577.24609	-0.57	X5(E)	
LNGEG	20	1	489.23056	0.18		**fraction 15**
INAGD	103	0	489.23047	0.18		
INGGE	63	0	489.23047	0.18		
LGGQD	52	0	489.23047	0.18		
LNADG	139	0	489.23047	0.18		
EPTNT	42	0	561.25165	0.26		**fraction 15**
TNTPE	59	0	561.25159	0.15		
AGDIN	72	1	489.23065	0.55		**fraction 15**
QDLGG	24	0	489.23053	0.30		
DQLGG	36	1	489.23059	0.43		**fraction 15**
GEGIN	41	0	489.23074	0.74		
EGGLN	23	0	489.23059	0.43		
GDAIN	43	0	489.23059	0.43		
Sequence (subspecies)	**#PSMs**	**proteins**	**MH+ (Da)**	**ΔM (ppm)**	**Modifications**	
VESEN	133	1	577.24652	0.17		**fraction 16**
TDAPS	130	1	490.21457	0.36		**fraction 16**
SEAxS	130	1	490.21616	3.60	X4(p)	
ESASP	130	0	490.21478	0.80		**fraction 16**
GEGTPT	118	1	561.25104	-0.84		**fraction 16**
TDVEN	134	1	577.24652	0.17		**fraction 16**
QESPT	68	0	561.25183	0.59		**fraction 16**
DTVEN	134	1	577.24640	-0.04		**fraction 16**
ESVEN	134	1	577.24640	-0.04		
DTAVT	27	1	506.24579	0.19		**fraction 16**
ESAVx	27	1	506.24579	0.19	X5(T)	
TDAVT	32	1	506.24579	0.19		
AVEST	57	1	506.24591	0.43		
IGEST	56	1	506.24591	0.43		
AESVT	25	1	506.24585	0.31		
EASVT	24	1	506.24585	0.31		
SIDDK	28	1	577.28290	0.15		**fraction 16**
LSDDK	26	2	577.28290	0.15		

Supplemental Table 4.1. Peptide hits from HPLC fractions (cont.)

TTPST	2	1	506.24594	0.51		**fraction 16**
ADLST	12	0	506.24585	0.31		
AIDST	11	0	506.24585	0.31		
DALST	8	0	506.24585	0.31		
GELST	8	0	506.24585	0.31		
IDAST	9	0	506.24585	0.31		
IEGST	6	0	506.24585	0.31		
LADST	6	0	506.24585	0.31		
TSTPT	10	0	506.24570	0.03		
TVDEN	83	1	577.24670	0.49		**fraction 16**
TDDIN	66	0	577.24658	0.28		
LDDTGG	12	0	577.24652	0.16		
LSDEN	116	1	577.24652	0.17		**fraction 16**
DSLEN	186	0	577.24652	0.17		
DISDQ	16	1	577.24677	0.60		**fraction 16**
DLSNx	15	1	577.24677	0.60	x5(e)	
SDIEN	50	0	577.24658	0.28		
ADTPS	66	0	490.21454	0.30		**fraction 16**
ASEPS	65	0	490.21454	0.30		
ETGPS	66	0	490.21454	0.30		
TEGPS	102	0	490.21454	0.30		
TDPAS	54	0	490.21466	0.55		**fraction 16**
Sequence (subspecies)	**#PSMs**	**proteins**	**MH+ (Da)**	**ΔM (ppm)**	**Modifications**	
VESEN	77	1	577.24670	0.49		**fraction 17**
GEGTPT	61	1	561.25171	0.35		
TDVEN	77	1	577.24670	0.49		
SEAxS	71	1	490.21481	0.86	x4(A)	
TDAPS	71	1	490.21481	0.86		
DTVEN	77	1	577.24670	0.49		
ESVEN	77	1	577.24670	0.49		
QESPT	61	1	561.25165	0.26		
ESASP	71	0	490.21457	0.36		
xISES	34	1	550.23584	0.58	x1(d)	
LDSES	17	1	550.23578	0.47		
DISTD	14	3	550.23584	0.58		
IDSTD	31	2	550.23584	0.58		
SLDTD	8	3	550.23590	0.69		
TVDTD	1	0	550.23584	0.58		
ADTPS	19	0	490.21481	0.86		

Supplemental Table 4.1. Peptide hits from HPLC fractions (cont.)

ASEPS	21	0	490.21481	0.86		
ETGPS	19	0	490.21481	0.86		
TEGPS	19	0	490.21481	0.86		
TVDEN	6	1	577.24664	0.38		
DISDQ	22	0	577.24664	0.38		
DLSNx	21	0	577.24664	0.38	x5(e)	
xSLEN	221	0	577.24658	0.28	x1(d)	
ISDEN	123	0	577.24658	0.28		
SDIEN	38	0	577.24652	0.17		
AIDST	9	1	506.24594	0.49		
ADLST	9	0	506.24594	0.49		
DALST	8	0	506.24594	0.49		
EAVST	7	0	506.24594	0.49		
GELST	6	0	506.24594	0.49		
IDAST	6	0	506.24594	0.49		
IEGST	5	0	506.24594	0.49		
LADST	4	0	506.24594	0.49		

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

# ENTRY	SEQUENCE PREDICTION	ALT. PREDICTION	ALIGNED PROTEIN NAME AND FULL SEQUENCE
>1		>01	>BJP42_12165_Tn3_pseudo
	TDAPS	ADTPS	MPADFLNADQRASYGQFSGEPNDVQLSRFFLLDEADIDFINNRRGRTNRLAVALLIGCVRFLGTWPHNLSSIPANVQWFVARQLGISDTG VLSEYSRRETTLREHQALIRRQYGYRDFAWPWTFRLSRLLFTRSWLSNERPGLLFDLATSWLLQNKILLPGVTTLTRLISEIREKSADRLW SRLSGLASGEQCSLLEELLQVPDGVRTSRFEQLRKGPVAISGPAFNQAVARYLKLKAFGMQDLDFTGIPPVRFNALARYAEMISVYKIAR MPPTRRIAMLVAFVRSCEVSALDDALDVLDGVIADIGREAKKIGQKKRLRTLKDLDKSALELAHICSVLLDENIDSELLRSTIFKKFPPARL ADTITSINAIARPPNASFHDEMVEQYGRVRRFLPCLLENIEFSAAPAGETTLEAIRYLAAIRSTRRQHIDDAPVAIITGPWKRLCYGKDGHLS RQGYTLCVMNKLQDSLRRRDIYVAQCERWGDPRAKLLQGQDWHTNRVQVYRSLGHPFNAVEVVNALNRQLDTVYRQVAKNFSDNQA VSLDFTGKRTKLTIAHLNGLDEPPTLKLLSKRISDLLPAVDLTELLLEINAHTGFAYEFTHASEAGARMDDLTVSICAVLLAEACNIGIDPFI RPNIPALTRYRLSWTRQNYLRAETLVKANARLVDYQSTLPLAQKWGGGEVASADGMRFIAPVRTVHAGPNRKYFGSSRGITWYNFISDQ YSGFHGVVVPGTLRDSIFVLEGLLEQQTGLNPTEIMTDTAGSSDLIFGLFWLLGYQFSPRLADVGASVFWRVDEEADYGVLNDLARNTA NTRKIEQHWDDMIRMAGSLTLGKIRASVAIRSLLSSDRPSGLTQAIIEAGKINKTCIYLTILMMKIIAAAF*LNSTGEKADTPSPEPSVMDKK VKSENAIRTGRRISWVRWGW*PMQLFSGIPSTCKLR*IICAMKVKPSMKRKKPDFRH*DMRI*TCWGTTDSRLQSR*QKVNLDH*SKRKK RMNGRC
>2		>02	>BJP42_03515_RTX-toxin_pseudo
	DTVEN	AGDIN	MGNSVSARHTREANQQMTSSNSESSEIGNDLNQYQITNTENRDLIEKKILKLARVDGKSFNPEYLDFDSNSAMEDLFKKRTRLQKDAKLY FKEKSIQGRSEFHLSLFANQGKTAESISIKEFLRNNSGLVIGESHSALTSKKLLIDNMKILASKGVKVLYLEHLFSDLHQEYLDKYFSSGKM PKQLHEYLNILDRGHHTDRSGQYTFLKLVEAAKNAKIKIIAIDCAASYINEGINSRDKSKTRIQMMNYFAEKIISQDVKNNNKKWIAFVGN THTNTMHDTPGLAELTNVIGIRVEDSNENNLSIDPGKNIEGIENIQRPGKIFIKADLLLKINVNSHDQINIERLPIAPSIKHSHPVSDKSFSQAE TRFAGQIIVELEEDRVVSEAAKRLAGKHAERSLIIQLKKEGGYEVVKGNPYLLRGDLRWQVVGHSREGANEGYPGTFGGLSAQQLSNQLI ALHAQLKKQFRINPRPKYISLVGCSLAAASVEDKSYAQHFGKALKHQAGWQTDIGARRLPVTVNAEGRKLSVVNEAASEEMPDKVVLS WTPEGELVVQKTRFSDRASGVLAPETLSSWEPPLDSLVEPRSAVASLQTGLGDVPPALDRSSDAGEISLSRKALHQMGAMVEGQPLSPEH FTKENGVQKLRFDAQRLAFYLHHADGGELSTQQAVRFLKQQLDVKSPPSLLEGDDKASTMVALEQLNRINEHVQRGAPAPMTSSLDIPP ALWGSLKSVLAPAGGARLNRLSARAGFGMQGYGYLRGLRDLKNYRDLLNKGGLTETQKEELIFQHHLAIASFSSNIGIDATQSGLGQYG RRLVQNGISSGVGMNVARFGGPVLGALSSGFDMVQAYRAFSELPSTAEPKARQDLIVNGALSVAGATLNIGMAIAFAVGGQASLLAGPA GLIMGGGLVLGGMIYSAVRQVEELKKWVQLSVLEELDSGWRLFSGEYLAPELANQASRHQMLAHAQAEHQKAQTAQFQTLLLSHGQID KVYGSLGQILLTPQRYKKIVAHHAGMEEDRVHPDEITQGKRALSDKLNRLKSNLLISGEMTVVDSDYEYYSPTLTSRDDRSDADIMMTLS AGHRQSLGRFVDENTQSLTASLEMKGRHRAPSALGDFNGDGRPDIGYFSPEGLFFLFADENGAYGPVKKMDTKNTFYHGLYANHVRHL VGDINGDGLDDIMFFLDSGDAVQILLGQPTEDFVFKEIKGVKLPALSQAAPVLTDVNGDGHPDWVSFVKNSLYVHYGDAGKMLGMPEE VSHPDAPLHSAQYLHHLAGDINGDGLGDILSVTRDGQLETRLGSRTPGQPLRRLANQTQAAVQGFGGVFNPSQLQLHDMNGDGRSDLV VIQDDGSYTMSYGQAEGTLGEASDDKSREEQGAKVAYRPNLLAKQNEQKIVGIRQGAVHLELISLNEQGEVHAHPFESRRERDLIAYYK LGDGEDDMTGQPSRRNYFDVGGGPKSSRAGLAPILFYCKGSSRHRVNPAISTVRALSVLRRARKMTRTP**RRSSRLTAEAIRLTCTKERS NMSARIPR*RH*RISNMR*VTVRRMIF*LVTTGITIWRVPVGQIPCQVTVGMTF*A*AQGPRQEVRASTAIAFYKTTAPQTR*PK*RSKRPW VNMTPATSYWIITRPM**ACLSRTQSRGATWC*P*EMITRVSLR*DLKTCTVWRRERREPCSIIIFCTPGMACSSPGFLKRSCNRRMADGLC LP*F*ITFTKKSF*WAPAWPIIMW*RMRF*NGTSSFIVMSCLIIRTTL*YRIE*KSPGCCKRRYSPQTKTTWC*ALSAIPTLVSVSNISSKRRLT VLSR*STSMGMSPL*R*TKKVWSLLADLRERTKMTR*GYVMTER*KISAWI*GRALLCQ*ARSRSLT*QQEDCLNPWRVLRHHKVTHWR LRQSLYTDTGTHAGQVDTTVIHP*YAIFGVI*HF
>3		>03	>BJP42_05035_HP_pseudo

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

	ESASP	AGISS	MPIKILPPELANQIAAGEVVERPASVVKELVENSLDAGATKIDIDIHRGGIEQIQIRDNGCGISQEDLPLALARHATDKIDSLED*QRLLSMG FRREALAGISSVSRLQFTSCPAEQAEAWQAYAEGRDMKVTIKPASHPVGSTIEVLNLFYNTPARRKFLSTEKTQWQYID*VVRRLVLSRFD VSIALSHNGKLVRQYVRAQSGSQKEQRLLRLCGPPFLKQARALSFDHSDLSLKGWITEANGPKINEIQYFFINRRIVRDRVVNHAIRQAYE AISTGPKTLSYVLYLDIDPRQVDVNVHPSKQEVRFHQTRLVHDFIYQAITTVLNHSACLSSDTSDIEQNASAVKESPQVPKENRSAAGENH YTRPDRALSDPQILTKKTEEAALFFHDDSQTGHEKISLKKDILEKVNEHYFGRVLSIVSSCHALIEYGHHIALLSLPEADRLLKKVGLTPPSE KSRIQPLLIPVKFTLEKAQITAGERHQHLLTLIGIDLAIEKNAAVLRGISLVLRAYNLQKLIPELLVYLSTHDTLLPENLADWVASQLPSLHH EWTVSQAIQLLTDFERFCAESLKAPSEKLLQQIDIQPALARLAH*
>4		>04	>BJP42_06585_HP_pseudo
	LSDEN	AGSIS	
>5		>05	>BJP42_06825_tRNA_pseudouridine_synthase_A_pseudo
	ETGPS	AIDST	MISKFQYRLLGVAIMVILAIIFLPWLLDGKKEYDITSVSNSVLNSTPGMNPSVSSDSSALVQSTNVLSFEVEEKKISKEEAIDSTDQNKSPPS EKKLRSKQAMSEIQQEYVIQLGAFNDMARVNQIIAKLRLSDYQIIRFLLCLLCL*KEKSLGF*LVLILRKISLNLCDLSFIDSLV*TDKLKPIS
>6		>06	>BJP42_01080_Phosphoenolpyruvat__phosphomutase
	LSAGS	ASGLS	MSTLSPSYTIDSTNKIPLTHSAQLRQMLLSDELEFLMEAHNGLSARIVREAGFRGIWASGLSISAQFGVRDNNEVSWTQVVDMLEFMVDA SNLPILLDGDTGYGNFNNMRRLVRKLEQRGVAGVCIEDKQFPKTNSFINGERQPLADIEEFVGKIKAGKDTQTDSNFSIVARVEALIAGWG MDEALRRAEAYRLAGADAILIHSKLSKSDEIVAFAREWADRGTPLVIVPTRYYSTATDVFRQAGISTVIWANHSVRSAVSAMQSVVKEIY ETQTLINVEDRIASVNEIFRLQNAAEYTAAEKIYLSASSSPRVGLVLAASRGARELEPITHDRPKVMLPIAGKSLLRWLVDGFKKEGINDIT VVGGYRAENIDNAGIRLVVNQNYATTGELVSLTCAIDKINSEVTICYGDLLFRHYVLRDLLENTAEFCVVVDSSQTSADNRTVRNFVYCT EKDNRGMLGQKVLLKHVTNDLNAVPYEDPLVQGRWIGLLNVSGKGLNRLKIVLTNLRMQPEFNTFGMPELLNALIADGAAIEVLYIHGH WRSVNDLNDFTNAQTPLKNTENLSYKPEIDQ
			>BJP42_04955_mobilization_protein
			MSQSEKRQRTQLLLGILCTPEEKKLIQEKAEASGLSVGEFLRRCALGRRITPKTDVKLISELSKTGLLQKQLFNEGKGVHSQEYSDILVALK KAILKIDFKE
			>BJP42_04955_mobilization_protein
			MSQSEKRQRTQLLLGILCTPEEKKLIQEKAEASGLSVGEFLRRCALGRRITPKTDVKLISELSKTGLLQKQLFNEGKGVHSQEYSDILVALK KAILKIDFKE
>7		>07	>BJP42_04250_MxiG
	TDVEN	DISDQ	MIETQNNQEIQEPQMIIRLLNGVLKGYEFELGKGKTLFVAGKEEAILKKENEAKLPENTIFIPFDVEEFQFEIEVDSSNKILLREFIHHASKEK TIDFNHIVSVETLNFAIKAINTEWDEAILTGKKKEMAVSKQSKKSFVRLLMTLLISGFVVGLIFFMLSYFKTTDEEKKEKEKTKEISQISYVL GDYQGKYEILGAHDHWMYVFAKNEKEATWARQSLTRADIKNPIRVINLQDEEKRVAAWLSENWSWVKYHRVKIETPLVPVLLISQERT VLDAQKSLNLMQSLKEQMPYAKEIKIIDISDQEVASQAEAGIQKISVPYIRQDHGNSVTFVIKGALEDGHIQQLRRFVVNYNKAWGQQYV QFSIELQDDWLKGKSLKYGHQGSYVKVSPGYWYFPKP
>8		>08	>BJP42_00620_GTP_cyclohydrolase_II
	SVEGA	DISTD	MNLKRISEAKLPTPWGDFLMIGFKELDKEQDHLALVLGDISTDAPVLSRIHSECLTGDALFSLRCDCGFQLKAAFKQICEEKRGLLIYHRQ EGRNIGLLNKIRAYALQDQGADTVEANHQLGFATDERDFTVCANIYKILDVKAIRLLTNNPEKIKILKKEGINVVQRVSLIVESNPNNQDY LSTKKNKMGHFFKK
			>BJP42_00890
			MMKFNHKILKRLTLALFIFSVMACTSHKNTNHSEMGAETTDISTDMNGNSGVSSADERARAQIEGLQKTSVVYFDFDKYNIKPNFAEILN AHADFLLNNPSFKVDIEGHTDQSGTSEYNMALGERRANAVKMYLQGKGVSAQQMSIVSYGKEKPVYRGNNKEDYSQNRRAVLIY
>9		>09	>BJP42_08760_RTX-toxin

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

	AGSIS	DQLGG	MTQEQFPRNSFIPPHQSDLPVHRTKRTPTDFVPVTQWDGLYETVMTRSDGGLTRYAGQLIIQLENDPIILDAALRLAEKHPGSVLVQLDQR NHYRVLSGCPKALQGKLRWQVVGHGNPARANTSPSIRRMSGMNPGGLAELLTHFNQHFSKKYHIDSTPNRISLVGCRLINYTPPNFATQF ARAMKRKGILADISARTEFMAVADGSGSGKKFTTQTLSEIGAQTAGKTLLLGWNKNGELIQKKHSSAFRLLSPVLDAWAPRYFTNPKIQN KRFRYLVYPSSITDLMSRKDIFSKNIQDKLRQSDLISASILSNETLRGYTVGVILEVPEQNILAADPDETSLEANFKGVIKKSGGERLALTKNI PVSELTRNGMLSKQLFSYPQRVATPDDIINNTLSQNTVLITGRPDINIYPHYPVTKDIKITGLFLISSDDLKNTGGASHFSSSKTLTSKNFRLF SEAARDIAQELDIELKVIKVRQWVDGFTGHTRNITSEMNQKSSRNLSDMARAQKHMTSSLTDKLKQILILVDHLSSRRESLLSLGHFQLLT EFFHLPSGELDRKKLYQTLSDPLLYSEFTSQVQSLMQLSDAHDQLGGLTGQLALKRTADWDNFVVENLLNWNERIKLYNNTAEFRAEIQ PSVQAIYIDDGSDLAQQKARKVSAAYLSLLKTGDDALLVKLLNSFQSHAEISAQKIFFPLDDKEAESFRRVIHQMEDASVFVHKGAQNLS QFFNTLAHQTAGYYQLRMGQHILTIAKRQSVDKKIHWSLYDANFGEIRVTTSDVQKASHPMRVLLHDYFTRLSGTPQQGGAVVLDVYQ LEPTQMLQSASFKKLQSLMSTTDGSLSIPLKPVQTQKNLLNPRFLKLSRTIGMFGEASQGLSWLQSVTCLSHYWRRRSSDELMEPEKQAL DFQAKLAMSGLLYDLGSILLTLGFRQLGARIMQKISYQPVLGVSSRIRGLHYKAGVTVARYGGAFLNMFGAAFDIYQAQKAATELKTAT APDARQDLQVSLALSSLSASISLGSGLAFLALTGQIALMAGAIGIALGMLTAVAGGIYFSVRQVQEIERYTTLTGLQKLRTGWLKFWGAE VDIDITNQVTKGQAAIEAKKTLLQQCQTHFQTLLDSDTEIETVYYSTGNIMLSEHPYQKLTGKKWTMLIRLGTRYLLEDIKDKILPSQAGAI LLEYREKKVQNDLPLYMDLTLQDSQYKFYDIKALSPTDDKISLDTEAHSDGILSLNRLVTPSSLSSSDMQSNPPNATNPAKIYFFLGDGDD EVVGHPEKRNMFDIGDGVKNFKGGNQADMFIFNGTSAPAQPSVFDGMKGVNSLVANKKPSQGGYRVNLEEGTFSFADFQKIATLKNID HVETCAETDDVILGNAQANILNGKGGYDILKGFDGNDILVAQTGELYGGKGTDCYRVLQNTGSDPATLIIKEEDAPESLSPVFLEYAVTQI ISIKKLKQRLFIQLRNDNQSITTLTLYDLYVSSVDEQTRRLTSQYIFYTKDGVIFTGFPPEIAVNPDEEEPEPLTLMAQYAPKFDQTFHFPKT DSEKAPVIIQKMQHQKTPGVLDVRGETKLLPDFLRLLPQETGFDDNLQGDERRNTLHSKKGQDILEGKGERDVYIIEDRPMGTEVSINNH DQPETGEPGQDILLLPFLLDDLEIRQEKKHVILSHLDSPTEHVTIRLMNFMEDKSYRHLSIMDKEGRMHTVRLNSSNEPDLIVDQLIPDMTV SDGHHILSQLRQTNAGFSEVEQVFNIPGDKILQTLQAKMIPVISS
>10		>10	>BJP42_09905_Asparaginyl-tRNA_synthetase
	GLGEN	DSLEN	MKTVPLVDVLQGHIDLGHTIIVRGWVRTRRDSKSGISFVALYDGSCFDTLQIVIEDSLENYHDDILHLTTGCSIEVSGTVVASQGKGQKFEI RANHLKVIGWVEDPETYPMSPKRHSVEYLREVAHLRPRTNLIGAVTRIRHHLAQAIHRFFHEKGYFWIATPIITTSDTEGAGEMFRVSTLD LANISKTADQKIDFSEDFFGKEAFLTVSGQLNAETYACALSKVYTFGPTFRAENSHTSRHLAEFWMVEPEVAFATLGDVALLAEEMLKY VFSAILKEREDDLDFFTERVDKKAKKRLEEFISSDFARIDYTEAIEILRHCNHSFVHPVSWGIDLSSEHERYLAETHFKAPVIVKNYPKDIKA FYMRLNEDGKTVAAMDVLAPGIGEIIGGSQREERLDNLDQRLTEMGLNREDYAWYRDLRRYGSVPHAGFGLGFERLVAYVTGVENVRE AIAFPRTPKNASF
			>BJP42_10085_Putative_inner membrane_or_exported_protein
			MEEESNPWPSFVDTFSSMLCIFIFIMLIFLLNNMLVVYESAIGGEGKVIYQKQSDSIELQQENSLNTLNKTEVEDSIQKIQPGIENGHVNTAQ DKVLVNKKNTQNELSQIQIENENIVKGVPKQVKNRNNDIAHNKDKTRVTITGNTLTIEYEGEVNTYLQEDITKMMNWIKSTKKNSFEIEIFI PKSKISFSDSLRLAYEKGLILMREIKSVDPSLMLNININTNSDSLENKVVIVPKN
>11		>11	>BJP42_10085_Putative_inner_membrane_or_exported_protein
	VSEAG	DTAVT	MEEESNPWPSFVDTFSSMLCIFIFIMLIFLLNNMLVVYESAIGGEGKVIYQKQSDSIELQQENSLNTLNKTEVEDSIQKIQPGIENGHVNTAQ DKVLVNKKNTQNELSQIQIENENIVKGVPKQVKNRNNDIAHNKDKTRVTITGNTLTIEYEGEVNTYLQEDITKMMNWIKSTKKNSFEIEIFI PKSKISFSDSLRLAYEKGLILMREIKSVDPSLMLNININTNSDSLENKVVIVPKN
>12		>12	>BJP42_11005_VapC_toxin_protein
	VESEN	DTVEN	MKKYMLDTNICIYLIKKCPPEIVKKFDNYRKGEIIISAITWAELCCGIKKEGRDTVENLLSVLDVEPFGIKQADKYGELTHFFPNRKANLDR MIAAHAISLNVTLVTNNIDDFCSYTRVGLSIENWVSSHK
>13		>13	>BJP42_12325_RTX-toxin_pseudo

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

	ASGLS	EDDPE	MTQKPSPSDFFITHHQSDVPIRRTKRTPPDFVTVDQWEGLSQPVMTRSDGGSTRYAGQLIIQLEDKPGVLKYAVRLAEKHPGSVLVQLDK KIIIACYPALQKLFKANCVGN**DTPIHLLKKHPLIFGPCLA*HLLI*PTF*KFLNNVLIKNILSTVLQTGSV*CHVD*SITSLKILPLNLLMQ*N NRVFRLKYRPEYHWCLFLYMVMEKNSPVK*HQ*IRLRPQEKPYY*AGTKKES*YIKSTAVLFGCCILY*RIGHQTILQILK*KVRNLSISFIL LQLKI*YTAKTFFQKTFRIKSESGI*CLPVSSQMRPSEAIQSG*YWRSLNKIFWPPLQMKAFLSPM*TVL*KNPGRTSLQ*KQIFFKNSMRLVI PRTP*KLMVCYRINYFTTDKESADLMILSKKHRGEIQCW*RDDLISIFIPIIL*QKTLKSPGCF*YILIPLKTQGRRPVFLLQSLFL*KI*GYYVR RLEILPKN*ALN*RSSKYIGTLMMHRVIIKIFPQNQSKDINQSLRNDPWSKKSRVFFNRELKTSLDLGQ*PLL*QVDIL*PQ*SQNINGFFSSPL RRTRLAKIIPGIVRSNVLFSIYDSNTKLDAIK*HSWSIHELNRPRGFKSHRSMGQLCCRKYGALERTDKIVQ*NPWIWPSNSS*HTSDVY*L WQ*LGAKKSEKSEFSLFKLTQAG**GFTCKITRCISISC*DKRTKNIFSRGCSRPQKFWAAD*PAGRKFCFYCKRSPEPRSIF*HPG**KGGIL PATYGGTYFSDSETSGRR*KNPLVRL*CEFW*NIDDQ**ISKYQ*PSEINVTGLFHALKECTSRKRDRGI*CLSVQF*SHV*NLFF*TIKRIYQQ Y*WLVAHFIDP*IF*NTQKQG*HTTSWQAISYSRYVGKSTSRTLLDPIFNIPISLLAAALFRPINGIPKTGPRLRSSTCHGGSYVRCGEYLFRIG F**TGISTDSKVLFSVRWGVCF*NPAVAV*SRTQAGSIWRPFLKYLGGRFRYLSSTKNCH*T*NSD*TGYSTRLKGQYSPFFFRRFYQSHQR PDILSPKRKNGINGRHDGDSLGNVDRCSGGIYFSVREVQQIERYATLTGSQKLRTGWLKFLGSEIDTEVANQVTKGQAEKQAKENIHQQF KTHFQALLDSDNEIEAVYYTTGPIILSGCPYKKLTGRKRADLALTVSKLNRNIFTKYIKDKILPSQAGAILLDHRRKKEGQFSLYVDLTSEN SEYKFYDFEGLSATHDKISLDTDVLSEGILSVKRQSQLSSTSSPSVPPIKADPPDSTKPAKIYFFLGAGNDEVVGYKEKRNIFDIGDGTKNFK GGDQSDTFIFRGNTPPPKPSVLDGVRGSNLLIANKKPSQGGYRVDLQAGYFSFTEGDDSKKIATLKNIHHIETCPETDDVIWGNDQTNILN GKGGEDKLIGFDGNDILVAQAGKMEGGKGTDCYRVLQNTHGEPATLIIKEEDDPEEFSHVFLDYSVEQIISIRRLKQNVLIELRNNNDSITT LTLSYLYFLSVERPQRRLTTQYIFYTQEGVIFSGLPSEIAFNFNEEEPEPLTLIAQYAQKFDQKFTPPEYESEKYPLIIQKMAYQKNRGVIDIR GEKKILPDFLRLVPQETGFDDDLQGNELNNTLHSKTGEDILEGKEGRDMYIIEDRPTGTQVSIYNYDEPETGEPEQDIVLVPFLLDDIEIKQE KDHVILSHLDSPTEHVTIRLMNFMRDESDRHLSIMDKEGRVHTVKLNSNHEPDLVVDQLFPDMTVNNGDNILSQLRQNMTELSEVEKVF HVSSDQISQTISQTMIPSVSS*
>14		>14	>BJP42_07240_HP
	EDDPE	EPTNT	MSITQDKTRVIKKVFKKNGMENSAINSDDKYNRLLITNDHAFYELIRKKMFDGDITIKQYDYLYKDISEKMGQVALGKESFDIKYIINYSK SINQLDNTQSLFTEDELKNLISKNNQSNTDSNFDIRKTDKYIKTNNIKNVKISSYKSENVQKNNNDTLNKKIIFDSDGFYKFLIKKRELNLLV STYIHHINYIFSSKRLDSKDLIEKFNYSDKNITDFFKKKIPQNKNLLKEMAYISKKITEELEKMNRSDQKKEKINNFLNKSLSKTINQEQYN MLYEEISKNLVNTIMSIKDIVNIYSKHLISEPTNTEFKKLLKKQDEKIQPYIYETKTGLKFQKKISIHFNSKLVFNIDIFSGKEIYSDDLPNENEI RKSLTDVLILFDKTFGLKTDNIPFKGDANQANFRLFLFEDAQKYKEFFNRHEKNGFTNHVGFAQGAELSYISNMYVNTEKTLNDSLSTIK HEFVHALTFYATGKMGIKNVFLEGIAMYIAGLSERKTSTDFVSSFIHSVKGKVPKDLSLKNIMNPEYFKENQLNSYTIGPALIAYFEEKNLY FIENLFDAIKYDKKNSTSGHYFNQMMKNIYKEEENNNNGFQKWVDLKLEANSKRQKRSFDNQPNLLKQEWAGFDSEERIFTTNHSEYSP QKQDPIGLYSNRQP
>15		>15	>BJP42_09320_HP
	AGISS	ESASP	MENIQNEFQNAIKTEQFDDKTIQDLYKRTNKMFKEDSILRSYFPGYLKQKTNPLMDALENLKTLVTPTKDSIVDKMILGTFCIFNVVAGQS SGMTTSSTEGGALEGLVGAVTTMSFNGIPEHDAHAQENKVQDERKYAYKDAERGNKSIKREGLRKERDENLKIFNDSCEQQRQNTTNKI KQKEFNSAIEARGTTNTPEQITVNRQYSIPADNSVSDIIGDNNDSHTQANTSSVTTEKADNAVIKNKNVTKTFSNTASTSLTNNIYNPEPKD QLHFDSENMAVWESASPSPGQWFTGLVETQTGAIKIVPVNIFGGQNGKLNELTLNNTGANSINRYASGVPGERNGKSIEPSWLSHRNGST PHTAIAFKYGFNQESCLGFSLIKTTKNGEFAQIKFSSTSLNSKPDGEVKHSFSRATASDSKGFSQTAL
>16		>16	>BJP42_09415_ISNCY_ssgr_ISPlu15_pseudo
	QESPT	ETGPS	MKMFLPHDALFKQFLGDIALARNFFEVHLPADVLSQVCLSSLKRVSGSFVEPTLKQSHSDMVYSFETVTGKDGYLYCLLEHQSTTDKMM AFRMKKYSLAVMQQHLDQGYSTLPLVVPVLFYHGETSPYPYSLDWFDCFDDKEAARRLDSEPFTFGRCDGYSRRRHSEAWDDCLA*TG PETGPSS*HDGHRALSYSSGQAFPVKR*PI*KFVVLSLSRGRNGLSPVIF*GPFFDHSTGECDDDSRRAEKGRLRRRFKTGCSAGDTAGCPA G*RRGPETGYLGHRQKADGRWRKSRKNPKIYGAIEGRHRSVISL
>17		>17	>BJP42_10000_Polymyxin_resistance_protein_ArnT,_undecaprenyl_phosphate-alpha-L-Ara4N_transferase_Melittin_resistance_protein_PqaB

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

	GEGTPT	GEGTPT	MLKLFKNHMVILVTIFSLLYLIPVDQRLLWQPDETRYAEISREMLERGNWVVPYFLDIRYFEKPIAGYWINNIGQWLFGHHNFSGRFGSIF STALTACLIFCLSIRLFHNRTTATIATLIYLTFFLVLSVGSYAVLDPMLSLWLAAAMVSFYGTLDAKSTQQKFIAYVLLGIACGMGFMTKG FLALAIPVISVLPVMVQQKRLKELFCFGPIALLSAFLLSLPWVLAIAHHEPDFWHYFFWVEHVQRFAEDNAQHKAPFWYYLPIFCIAGLPW LGLMPGALLKGFRERKTRPQLFFLLSWILMPFLFFSLAKGKLLTYILPCMAPLSILLAAYAKDCAAQTRMKAFKINGVINIVFGIFFILVFLV GMKWMPTFLLYQSQETPKIILGSVCALAWASLGAWSLKNNAQNWQYAAACPVLFALLAGYLIPKKLIDTEQPQHFTQNNKALLSGSRYI LVNSVGVGAGVAWELNRSNIMMFHQKGELSYGLSYPDSKNKYVSAADFPRWLAQARRKGDVSLMLLLSEGEGTPTHLPKVDKIESSHR LILLWYKKTS
>18		>18	>BJP42_01180_Glucosamine--fructose-6-phosphate_aminotransferase
	TVDEN	GLGEN	MCGIVGAVAQRNIAKILIEGLRRLEYRGYDSAGLVVLDADGQMTCLREVGKVKALSDAAEKKILQGQTGIAHTRWATHGEPSKVNAHP QISQNICVVHNGIIENYHPLRELLIRQGYLFASETDTEVIAHLVHSEQQQGGSLLQVVQRVIPHLMGSYGAVIMDSHDLNLLIAVRSGSPM VIGLGENENFVASDQLALLPVTRRFIFLEEGDIAQIRRSDVEIFNQKGEKIQRPIWVSESKYDAGDKGIYRHYMQKEIYEQPMVIKNTLQGR LLHGRIDLSELGTRAKGILAQVEHIQIIGCGTSYHSGLVARYWFESLAGIACDVEIASEFRYRRSVVRSNSLLISLSQSGETADTLAALRLSQ SLGYLGSLAICNVEGSSLVRESDFALMTKAGTEIGVASTKAFTTQLTVLLMLLAYIRRLKNQPEALEHQIVKALQILPARIEQIFSLTKDIET LAGDFSDKNHALYLGRGDLYPISMEGALKLKEISYIHAEAHAAGELKHGPLALIDSEIPIVILGPSNGLLEKLKSNIEEVRARGGRLYVFCD KQGEFVANQRMKLIVLPKIEEIIAPIVYSVPLQILSYYVALIKGRNIDQPRNLAKSVTVE
>19		>19	>BJP42_00300_Iron-sulfur_cluster_assembly_protein_SufD
	TDDIN	LNGEG	MAGLPTRVKRELPEKPDVSKNPFSQLFEQKTTAPSHHKEAHWQALLKLGFPHAKHEDWRYTSLNRLLDHTFTLEKKSTVTPEKCGDLAL FMVGQSHKLVFIDGHFIASLSEIDEIPYELLKRKQAVPAPVKPEFFLHLTESLSVENLVIRLAAGQKSKKPLYLLHISSGQETSPLKMAHYR HHLLLEPGSEAEVVEHFVSLNTQRHFTGARLTASIGEEARLHHYKLAFENNQSDHFAHNDFVLASRASVHSAHFLLGAALTRHHTSVQL NGEGASASINSLLLPKEMELFDSRTHVEHNQIGCQSRQLHKTMVSKQGTAVFNGMIKVAPKAIQTDAHMINNNLLLNASAQVNTKPQLEI DADDVKCSHGATVGCIDNQQMFYLQSRGITHSQARKMILFAFASELTEGIKLDLVRESVLSRFSERLSEELS
			>BJP42_06415_Pyruvate_dehydrogenase_E1_component
			MSDPLSNNDVDPIETQDWRKSIESVIREEGAERAQYLIDQVLSAAQKMGAPLSQISLQKNYINSISTEEEPPYPGDLKLERRIRAAIRWNAI MMVLKASKKDLDLGGHIASFQSFATFYEVCFHHFFRAPNQHDGGDLVFFQGHISPGIYARAFLEGRLTEEQLDHFRQEVEGKGLSSYPHP KLMPEFWQFPTVSMGLGPLNAIYQAKFLKYLEHRVLKDTSKQRVYAFLGDGEMDEPESKGAITIAAREKLDNLVFIINCNLQRLDGPVTG NGKIIDELEGIFSGAGWQVIKVIWGGRWDDLLSRDTSGKLVQLMNETVDGDYQTFKSKNGAYVREHFFGRYPETAALVKDMSDDEIWA LNRGGHDPVKIFAALKKAQETKGKPTLILAHTIKGYGMGEAGEGKNIAHQLKKINMEGIRYLRDRFNIAVKDEDIEKLPYVTFEKESEEEK YLHHQRQSLGGYLPSRLKNFTQKLELPSLQDFSALFQEQTKHISTTIAFVRVLNVMLKNKFIKDRIVPIIADEARTFGMEGLFRQIGIYNPG GQQYVPQDREQVAYYKEDAKGQILQEGINELGAASSWLAAATSYSTNDLPMIPFYIYYSIFGFQRIGDLCWAAGDQQARGFLIGGTSGRT TLNGEGLQHEDGHSHIQSLTIPNCISYDPAYAYEVAVIMHQGLERMYGEAQENVYYYLTTLNENYHMPAMPKGAEEGIGKGIYLLESLP GKKSRLQLMGSGAILRHVREAAQILSQQYDIASDVYSVTSFTELARDGQDCERWNMLHPAETPKVPYVAQILHQAPVIASTDYMKLFAE QIRPFIPTRHFRVLGTDGYGRSDSRENLRHHFEVNAAYVVIAALTELTKIGVFNPEEVQEAMIKFNIDPNKPNPRLA
>20		>20	>BJP42_07685_NADH-ubiquinone_oxidoreductase_chain_L
	DSLEN	LSAGS	MNLLFLTLFFPLLGFLLLAFFGRNWSENKAALIGVGAIGLSALTTLYVGVNFITQKNALSSPDRPFFFEQTLWQWISTDYFHIPITLMLDGL SLTMLSVVIGVGFLIHLYASWYMRAKHSSHDEKEQAKVGYARFFAYTNLFIASMVVLVLADNLLLMYLGWEGVGLCSYLLIGFYYTHP ANGAAAMKAFVITRIGDVFLALGLFILYQQLGTLNIRELMILAPTHLAVGSDAITWATLMLLGGAVGKSAQLPLQTWLADAMAGPTPVS ALIHAATMVTAGVYLIARTHGLFLMTPEILHWAGIIGAITLLLAGFAALVQTDIKRILAYSTMSQIGYMFLALGVQAWDAAIFHLMTHAFF KALLFLSAGSVILACHHEQNIFNMRGLRKPLPFLYLCFLVGGSALAALPVITAGFYSKDEILWGAWLHNNIELMIAGLVGAFLTAVYTFR MILIIFHGAPPEHPPYKAAPREISHHVPLWVLLILSTFIGAFIMPPLAGVLPENHALKIGKTILAFCSGLIVITGILLASFLYSGDRHWAQRCQ KTSIGTFLSALWFNAWGFDRLYDRVFVKPYLWIARCLQSDPLNTSMNFLSLLCLWGNRLFSLSENGQIRWYLASMGAGALFILTLLLMN
>21		>21	>BJP42_05815_HP

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

	LNGEG	LSDEN	MHLIISREDFLHHHWQKHPVLLKKAISDFINPISPEELEKLVKDNSLESQLIQKRHGKCQLGYQAFNGCASLGKRNWSLKVEAINHWHRA AEEFLSLFRVFPDWVKEALTVSFSVSGGGLGPQTRPSDVLVIQGMGRSRWRVGDRGASPIHDGQNDFAESIINEELSAGDMLYIPKGFPHE SMSSEAAMSYCLNFWTDNSLRMIRNWTESLSDENHRGIEYSPSPDLLLREDPTEILPQDLTALQDMMSQFLLQKRGHLKTWFVQEMSQT SYELPKAPEAEVYSPSQVQTLLQQGHLNSERRSSLLL
			>BJP42_06165_HP
			MIRNWTESLSDENHRGIEYSPSPDLLLREDPTEILPQDLTALQDMMSQFLLQKRGHLKTWFVQEMSQTSYELPKAPEAEVYSPSQVQTLL QQGHLNSERRSSLLL
>22		>22	>BJP42_12115_Putative_toxin_subunit
	EPTNT	QESPT	MQNSDTLQLNAPSLPAGGGTINGLKGDMAGAGPDGAASLSVPLPISSGRGYAPALTLNYHSRAGNGPFGIGWDIGLAAIRRRTRKGVPA YKETDEFTGPDGEVLVPALTSDGQAVIRTADTLINAKLSQPYTVRTYRSRLERDFSCFEYWQSPASGHADFWVIYSPDGQVHLLGRNPQA RLSYPQQPSQTAGWLQESSVSATGEQIYYQYRAEDEMGCEAAEKTAHPAATAQRYLVAVYYGNRHAGRSLPALTEIPTENAWLFILTLD YGERAQALATAPIWHSPGAGAWACRQDVFSDYGYGFELRTRRLCHQVLMFHRLATLAGSAKGGDTPALISCLQLTYQESPTISTLISVQQ RAYAPDGTACTLPPITFGWQTFTPPTTPAWQSQNDWGHFTPLQPYQWVDLNGEGLAGLLYRDNDAWWYRSPIRVHKANDPNAVSWGG ATCLPTLPSLQKGGTLTDLNGDGRLQWLVTTPTVAGHYDRTSEGDWLHFTPLSALPTEYFHPRAQLADIIGAGLSDMVLIGPRSVRLYVG TEAGWKTGQTVMQLDGITLPVAGADARTLVAFSDLPGTGQQHLVAVTADGVRYWPNLGHGHFGQPITVPGFHQPAPAFHPDRVFLADI DGSGTTDLIYVHPDHLAIYLNQSGNRFAPPFSVPLPEQVRYDSTSGLQVADIQGLGIASLLLTTAHPTPRHWVCHLAQDKPWLLNSMNNH LGARHTLHYRSSAQFWLDEKAAVLAKGNPTPPSYLPFALHTLWRSEIEDDITGNRLVSQVRYQHGVWDGQEREFRSFGFVEVHDTDTAS SRGTAAEVNSPTITRRWFATGLPAVDNRLPDEYWQGDTAAFARFSARFTTGSREHETAFTPDTPTAFWLNRATKGLLLRSELYGADGSQ QAAIPYTVTENRPQVRLIETRGAMPVVWPVMVESRTYVYERVSNDPQCSQVVLLASDENGLPLRQVTLNYPRRAKPAVNPYPDTLPTTL FASSDDDQQHPLRLTLQQSRLHTLKDTVKGVWLLGVPDATRNDGFVHPASAAPAAGLTLEALQKTNGLMTDKAAVFAGQQQAWYLD AQDKPTLKPTDLPPKVAFAETAVLDEALVNALAQDITTDALTRAGYQQTGYLFPRGPEKTAKLWVARQGYATYASAEHFWLPVTYRDT LLTAPVTVTRDPHDCVVTQYQDAAGLTTRAEYDWRFLTPVKVTDANDNVHTVTLDALGRVITARFHGTEAGKPVGYRNKPVTLPTDVN TALALTAPLPVAQLMLYVPDSRMKNAGDAERLPPHILTLTTDRYDDDPAQQIRQQVIFSDGLGRLLQTAIRQAPGDAWQRDANGALVKG PDGKPLSAFTTFRWLVTGRTEYDNKGQAIRTYQPYFLNDWRYVSDDSARTDLIADTHDYDPTGRVWQVKTTKGYLKRVYFTPWWVAR EDENDTAATL
>23		>23	>BJP42_10910_TonB-dependent_hemin_,_ferrichrome_receptor
	AGDIN	SIDDK	MSFTLPKLLRLSSLTLAISLPSLAYAESGMPDKIETEEKKPSFNKQSLKTETITVTATGNPRDTFSTPMMITVIDRDNAKSDTASSAADILLS TPGITIEGVGRKNGQDVNMRGYGGKGILTLIDGVRQNLDTGHMSGHFIDPSFIKQVEIVRGPSALLYGSGALGGVISYQTVDTVDLLQPGR NSGYRVFSMGATGDRSLGIGATVFGKTDDLDGLITFSVRDVGNIKESDGINAPNDEGIRNLMAKGTWKIDDSQSLSANVRYYRNNAEEIL NPQEVNPSLRNPMTDRKTVQKDVQFSYALNPKTSDWLDAKTTLYYSDVSIDDKVKKRADQGRTQKTLGTKLENRSRLLTHSPAAHLLT YGTEFYQQKQKPEGHAKAFPDAKTNFASGWVQDEITFRDLPVSLLIGTRFDRYKAKNPKNADISADNWSSRGAVTLTPTDWLMLFASYG QAFRAPTMGEMYNDAQHFRGNRFKPNPNLKPEKNATQEYGFGVKFDDLVLTEDSVEFKASYFSTQATDYITTEVNFRKGTTQYVNVPD TNIWGWDAVLNYKTPWFGWNLAYNRTHGKNEKTGLYISSINPDTLTSSLDIPIPKTDLSAGWVMTLAESPDFMKSPGTTHSQEIKAQPGY AVHDFYLSYKGQNQFQGVTTNLVLANAFDKEYYSPQGVPQLGRTAKLLISYQW
>24		>24	>BJP42_02745_Translation_elongation_factor_P
	DQLGG	SVEGA	MATYSSNEFRSGLKIMLDGEPCSIIESEFVKPGKGQAFVRTRIRKLISGKLLEKTFKSTDSVEGADVVDMNLIYLYNDGDFWHFMNNDTFE QLAADVKATGESAKWLVEQAECILTLWNGQPISVTPPNFVELQIVDTDPGLKGDTAGTGGKPATLTTGAVVKVPLFVQIGEIIKVDTRSG EYVSRVK
>25		>25	>BJP42_09690_Mobile_element_protein
	VESEN	TDAPS	MESGSFVDEDMKAYQNDILYSLKTKKGKGYLYILIESVRRESRFDYCLNSVEGARRTVSPVLAY
>26		>26	>BJP42_07990_hypothetical_protein_12_APSE

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

	TDAPS	TDDIN	MTEATQSDNPNDIDSLSTLKIQAVSGNLSPSIYANGRNQLPIEITAKAVNKDGNPLKFSNNTWISILNLCFAESDKKLNRNGNSGWCFTEV ENQYSKEIQLGSHRSKRFLSGEDGTAIIILYVYTDDINTKRIAVSVDTPNGKHFTTADIHSGAEKMSVTVAAIQQIIYRKNMLTITPIYELEK GVNKVIYSDGYTREFDCHYDNYYITIPLPIINAIAHGYGNGTTYPKWICAYDKQSNNQHMIVAHPDNTKKGTETFGFEGRATFLGLDSGA RFDLTKKVTFSEHDNAICFTQIAFQTGGNAWRLNDGQCLSNFVPYAFDTWFELFDIYGNYGKFNVQFTDDKLYIKITDRYK
>27		>27	>BJP42_12795_IncF_plasmid_conjugative_transfer_protein_TraG_pseudo
	ESASP	TDPAS	MAGTDPASEAGQLATNSPLVQNEHTADTSALRNQAQQTERDVSAPELKKAQDNLMNSLPEMSWSASSWGAWDNSTDWMDRRAEQSG GALVAGSQAGAEAFSKAMDQLRTMTPEQRDQFIAATQRGDQALKEEFGWAGDAMVGVAKLGRNVIGAAASGYDAAKEWLTGKSDLS KAAKGMSIEERGAFYAAAFASASEAGGGAAQQFMNQYGDEFKETMQSIAQSRYGLTQAQSAVYAESFDTNEGRMKQAVQNLKMEYA ERNPDGSTIMQGAKPVLSQQNEAFTDKLVNVLQNSTGAGDRSGSYLTSVRGYNIANRRF
>28		>28	>BJP42_00090_Thioredoxin_reductase
	GEGTPT	TDVEN	MNISRHHQCIILGSGPAGYTAAIYAARANIKPILITGMEKGGQLTTTTDVENWPGSPESLTGPTLMERMHQHAEKFNTEILFDHINKVDLQ RRPFYLFGNDQEYSCDALIIATGASARYLGLKSEEDFKGKGVSACATCDGFFYRNQKVAVVGGGNTAVEEALYLANIAREVHLIHRRQE FRSEKILRDRLMKKVASGHIILHTDSILDEILGDLSGVTAVRLKPSHTENGSVFQILDVTGVFIAIGHIPNTELFKGQLELEKGYIKVKGGLT GNATQTSVEGVFAAGDVMDPIYRQAITSAGSGCMAALDAERYLDIMDNRFK
>29		>29	>BJP42_02435_T1SS_secreted_agglutinin_RTX_pseudo

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

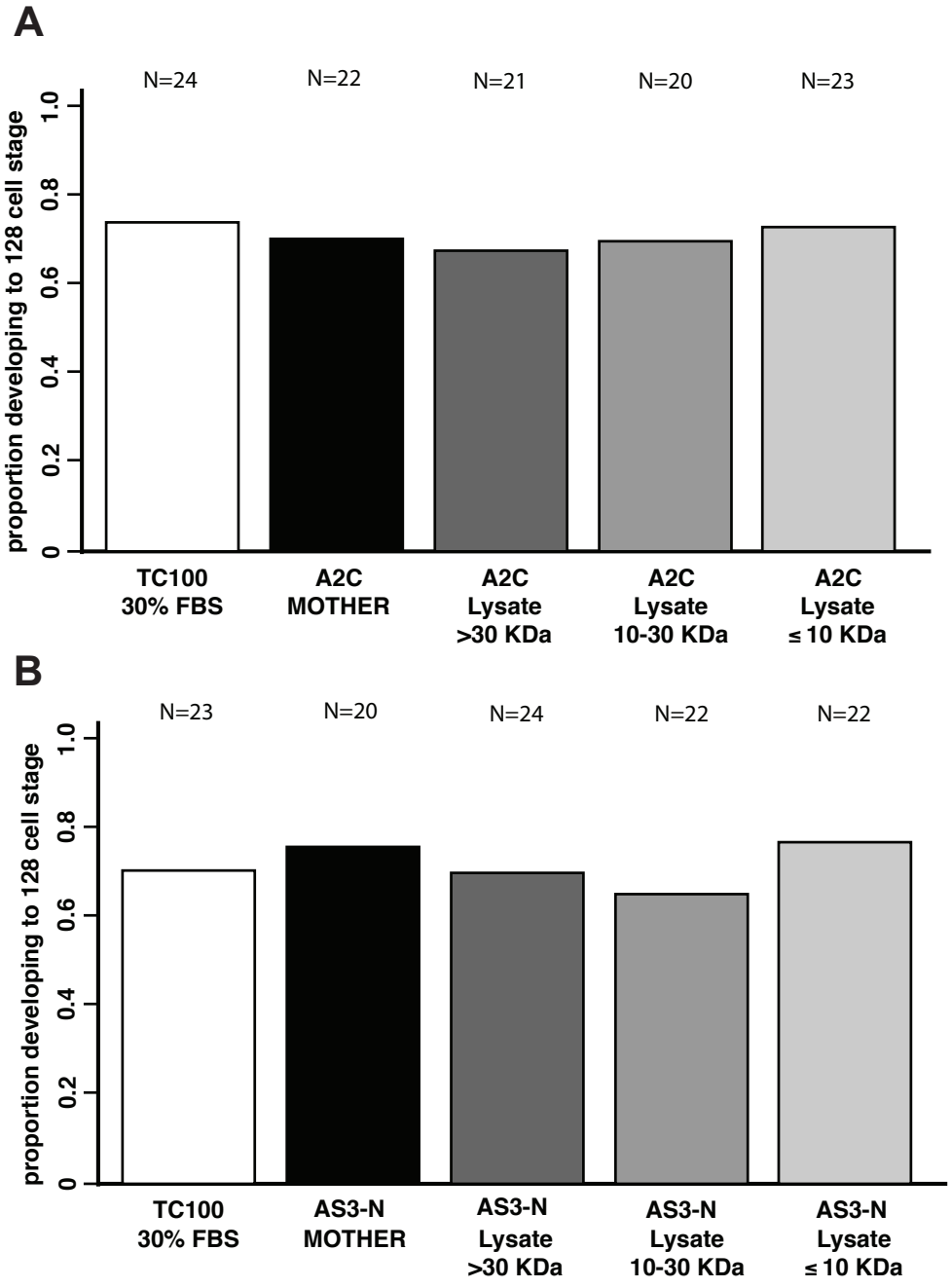
	TDVEN	TTPST	MKLKKPVVDGVGEEETEIKTITKKEGKFNLSDLNGIDLSLNANGHYTVTVEASNNDIKSSSALTVRFYGTVLTPEILLMNQSVENGQLVLS ESQKPSFEIKNVDPHAREIHVTATPIKPTGKQAQKFKVVLNQNGSIKTNFDNGFFEKIEGNSDYKFTPTEGWEEGRYSIDVSVINEANKESL NSSPLELIISTSELQKVDIELDDKDKTGDKKSNFTKNKNPTLIFRTQEENIDRIEIFKKVEERIKKIGTLKSSEGFEKNKDQHWPIPDNIFDKE GKYTITARPYKASKEGSPTEIDIDFYQHAPEVEKIQLNAKSKSTFDDSGIYTNKPTPGFLITVPEEDRHRVWKMQVLVIQKSDKKRLETIEIS DNLFQSSNPEPVFTKETWSKGSYDIAVTLTDKADNKSKQETQIIHFDNTPPQPVIAFAPEVSPEKSGEKNIIKNQNRKFVIKGTEEKGIEESD EIEVTLTQSNKNSSKETLYNSKKKKNEKLFSYDKDNTEQTFSLSSQEALDKGDFKLNVKVTDRYGNTGNSQEPLSFTIQDSVEKPKIHLLD DTSYEQQAGVYITKKPQPRFNLTQLDEHATRVEIKTTPPILNSDPSNETFDELATNNDKKEELKNGFQLKESWTTNGDYKLSVNTFVDEEP SPETTLKVLFDNELKQPQIALDEETIKSSSASDSGATVASQTPKFNLLNIEDDVKHILLQVYKDAQEKPVETIDFDIKDLKEDIERKGIKATN NRGYTYQPLEWEDKTKYKITLTVTDRAKNSNEKNPAEFEVKIDTEAAKTPSITLGESDNQAFDKIKKFITKNQTPSFQLDNIDDRAKTITITI YKLGTTDVLASKTFRGETTTFNVNNLIRADNKSQLSSLEDGQYTIQAQVTLKTNISSALGSLDFTIDRADAPTHTIELVGNTGTNDKPVTN KTTPKFKIKDIALDRLDWDNRDSIVVTFQNKKTKESVTQNLTESGIQENNTVEFTANQLEQGEYTVTADVTYQAGQVRSGRLNEDKGMT IYTTPPKSTIEFENPEKVGNDGGKNKVQKQDLKFKILVDENEPMLQKGSSQVKVSLKKDGITQGSEIEAQQNSNVWTVDLSRQSIPKGDY TIEANVTDKHGNTGTSEPMNFYILDGLLQPKIEIDTNDHTPYDPTEYKDNKKITNKKQPTFIIKNIDPEAESVAFSVSKKGEDAPFATSNYK KNEGGWPFSGNQKKIKVEFQDDGDYEIKLDVVGVGGVQQSAKALLLTLNTQVDSPKISLESQKDVATPLTSEKKPTFLIEQIGEHVREIKII REIQGKDHAKKEFVCKINDEGEVQTLNNDSEGKLTKDTKNKGTYRFTPKDDWADGTYKVRVEVTNRAKKSAMSNEDKIITLTIDNNLPE VPKIELHPDDNNTYPVSLKKTETWTQKNSPKFNITVKKNDVNPVTKVHVKVQKPSGNETIEQATQKDGKWTFTPGTPFTKGAYTLSVES ENKVGRKSNFSSNITVNFNDYPPGQPEIRLDSESITGMDRLTDITNHQNPKLSITNIKTIEDAVDPEAVEVTLNDSKKQALATNAIATYDSK NNTFFYVQRHLSEGKYTATVKATNKAGIPSTSDPFTFEIDTTAPNPPEIEILKGSYVTHQKLNKVLKNNKQLGFLVKTAEEIKDLKYIEVSL DGNKLNKNLLDPTNKGGLFSGTWLYNKNVWEFTYNHLFSLGTYNLLVTVTDKAGNTSNAQDQFMVLGELHPPKIEMPDNTNAAGPTST PHMINKVNMDQMDKRNLTISLDDNRLFPYVEMKAELIPNQNGSSKTTTINLRTEKEVWMGEIPTHQPDGEYTLKVTATDVLGIKKETNV EFVVDTTPPETPTIKLRDSVTKHAESKATQETTQEKPNFVIGSIPNDIYKINIKITGDDYDYDYEVGNLNNVENKLKEWSGIQNALKDNQT YSITVTFTDRAGNTSSNADTPFKIKKIVPYTFIDPTLILDPETDSGTQGDFKTNHNQPVLLFKNMSDAEVTEVNLKLENKTDKKTYTYSLK DFEPIHGKNSKEYKIELIKESSLTKEAPFPEGEYKVTVELTYASNEQYTKATTFEKPLTINRKKPQAVYDVKYTFKDGDQDQGKTINIEGK KPKNTDIFIQFHEQAEINIKKKNGDSLFSDDKFITTQPWGNRKALPSSLKTISATNPRILRSTYHSHLPSLDIVL*KTV*LRYLSKKKT*N*TIR LSLKIKLIMKYTNIQ*MKLI*KIKVFNLQLTKHSQTNLRLMSPGQSTKAKKLKKLYPKLLKMLNIPSQEAVVTSWWLKEN*KKDKH*I*AII MMIFLQKVTKLLEIKF**R*SHLIRI*GNLTCLLQMQMGEKVRLLKKKYPMSLFIGILSQNNIRLDLV*TILLNTFQTNSHAKKSVSILNMYW RCETVPHLLIMCTTRD*QALKNLIK*LSLITR*IMS*QNISI*IKKKIRTRFMEIIMAIQQKLGGVSDTMILIVFQKN*IKTIVKLNLSIKNRL*NF YLDLRVSITI*VY*R*RQ*KKRIKT*WKPDGICGLKRFLLKEKTTLMEKKRI*SE*QSMKNCKKDI*N*L*QMLITKNKRLSLTQAKIMKHLK *R*QEASKGKGLQ*TTGL*EFF*KIKIQVLSLILER*RLVE*HLMILQRLKMARKTLPQ*KFFQVGKNISEKTILKAEANVA*KKTL*KNQTT TPLKRRCIKSTRQQQIKTTPACHSQNSAPHSQSKRKTL*TLRNSQST*LNPKKRISLLSWWRNRRRPPHKILLSSTTPSTQRKTPPWISPPAS HSTPRKKRLMPSKRWSPWIIAPNIRLS*LITKGYSLSRCP*QKARMNSR*NLSIRMGIGSAWIKRSPSMRLLKEY*AAWKPQTDIKATFQQG QKRSRQKKMQIF***RLFYICLSMRRNLCIMD
>30		>30	>BJP42_02275_Cell_division_trigger_factor
	QESPT	TVDEN	MQVLVETTEDLKKRATITLFSDDIEKAIKDKVVKTAKTAKMDGFRKGKVPINLIEQRYGESIQQDILIDFMQDNFSKTILQEKLNLAGRPQ YLPGKYIKNENYTYSVEFEVYPEIKLKDLELIEVEKPIVSLNDDDIDKMLEKLRHQKKTWQEIPDLPAEMGDRVTFDFTGSVDGEKFEGGT ASDFVMNLGEGNMIPGFEEGIVGHKMSEEFNITLTFPDEYHSESLKGKESQFLIHLKKIEKFQLPEINEEFIKNFDLPEESSIEALRAEIRKNM QRELTKAIRKRIKEQTITSLCDLNKTKVPLSLIEEEIDALRTKAASYSDKKDIIKELPRSVFAEKAESRVLVGLLLTEVIDQNKLTVDENKVQ VFLEEIAEAYENPKETIEFYKNNKKLMNNIYNLALEEEAIQVLLKNAKITEKPISFSEFMKQIHETDTAND
>31		>31	>BJP42_02435_T1SS_secreted_agglutinin_RTX_pseudo

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

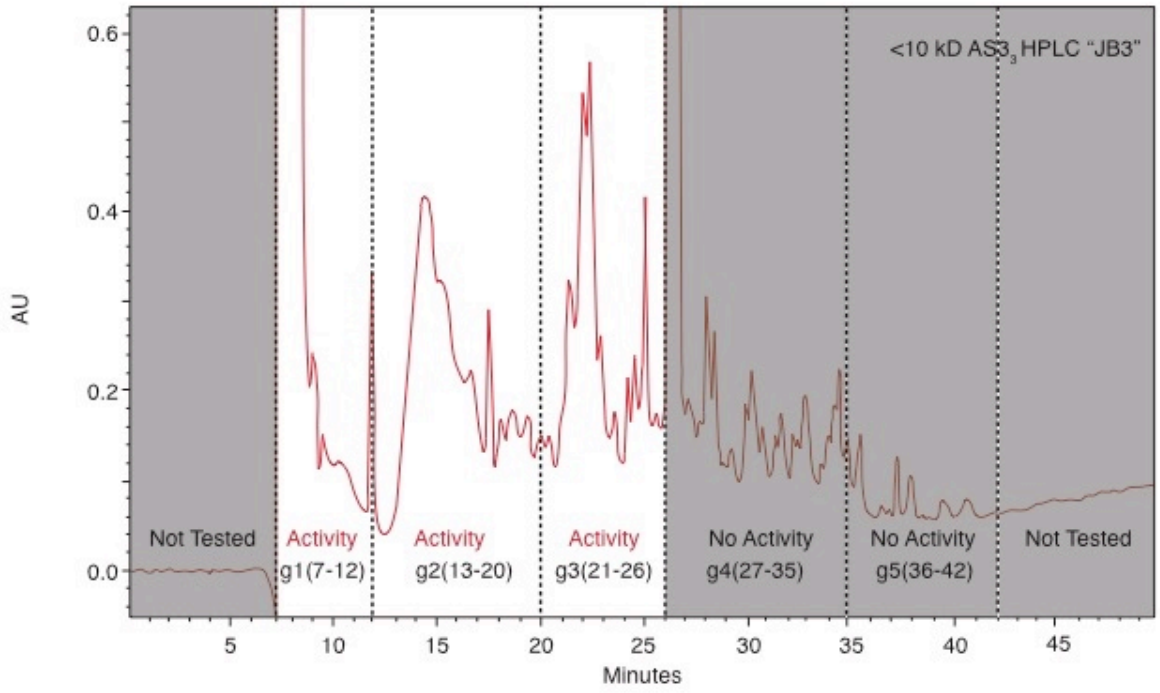
	DTVEN	VESEN	MKLKKPVVDGVGEEETEIKTITKKEGKFNLSDLNGIDLSLNANGHYTVTVEASNNDIKSSSALTVRFYGTVLTPEILLMNQSVENGQLVLS ESQKPSFEIKNVDPHAREIHVTATPIKPTGKQAQKFKVVLNQNGSIKTNFDNGFFEKIEGNSDYKFTPTEGWEEGRYSIDVSVINEANKESL NSSPLELIISTSELQKVDIELDDKDKTGDKKSNFTKNKNPTLIFRTQEENIDRIEIFKKVEERIKKIGTLKSSEGFEKNKDQHWPIPDNIFDKE GKYTITARPYKASKEGSPTEIDIDFYQHAPEVEKIQLNAKSKSTFDDSGIYTNKPTPGFLITVPEEDRHRVWKMQVLVIQKSDKKRLETIEIS DNLFQSSNPEPVFTKETWSKGSYDIAVTLTDKADNKSKQETQIIHFDNTPPQPVIAFAPEVSPEKSGEKNIIKNQNRKFVIKGTEEKGIEESD EIEVTLTQSNKNSSKETLYNSKKKKNEKLFSYDKDNTEQTFSLSSQEALDKGDFKLNVKVTDRYGNTGNSQEPLSFTIQDSVEKPKIHLLD DTSYEQQAGVYITKKPQPRFNLTQLDEHATRVEIKTTPPILNSDPSNETFDELATNNDKKEELKNGFQLKESWTTNGDYKLSVNTFVDEEP SPETTLKVLFDNELKQPQIALDEETIKSSSASDSGATVASQTPKFNLLNIEDDVKHILLQVYKDAQEKPVETIDFDIKDLKEDIERKGIKATN NRGYTYQPLEWEDKTKYKITLTVTDRAKNSNEKNPAEFEVKIDTEAAKTPSITLGESDNQAFDKIKKFITKNQTPSFQLDNIDDRAKTITITI YKLGTTDVLASKTFRGETTTFNVNNLIRADNKSQLSSLEDGQYTIQAQVTLKTNISSALGSLDFTIDRADAPTHTIELVGNTGTNDKPVTN KTTPKFKIKDIALDRLDWDNRDSIVVTFQNKKTKESVTQNLTESGIQENNTVEFTANQLEQGEYTVTADVTYQAGQVRSGRLNEDKGMT IYTTPPKSTIEFENPEKVGNDGGKNKVQKQDLKFKILVDENEPMLQKGSSQVKVSLKKDGITQGSEIEAQQNSNVWTVDLSRQSIPKGDY TIEANVTDKHGNTGTSEPMNFYILDGLLQPKIEIDTNDHTPYDPTEYKDNKKITNKKQPTFIIKNIDPEAESVAFSVSKKGEDAPFATSNYK KNEGGWPFSGNQKKIKVEFQDDGDYEIKLDVVGVGGVQQSAKALLLTLNTQVDSPKISLESQKDVATPLTSEKKPTFLIEQIGEHVREIKII REIQGKDHAKKEFVCKINDEGEVQTLNNDSEGKLTKDTKNKGTYRFTPKDDWADGTYKVRVEVTNRAKKSAMSNEDKIITLTIDNNLPE VPKIELHPDDNNTYPVSLKKTETWTQKNSPKFNITVKKNDVNPVTKVHVKVQKPSGNETIEQATQKDGKWTFTPGTPFTKGAYTLSVES ENKVGRKSNFSSNITVNFNDYPPGQPEIRLDSESITGMDRLTDITNHQNPKLSITNIKTIEDAVDPEAVEVTLNDSKKQALATNAIATYDSK NNTFFYVQRHLSEGKYTATVKATNKAGIPSTSDPFTFEIDTTAPNPPEIEILKGSYVTHQKLNKVLKNNKQLGFLVKTAEEIKDLKYIEVSL DGNKLNKNLLDPTNKGGLFSGTWLYNKNVWEFTYNHLFSLGTYNLLVTVTDKAGNTSNAQDQFMVLGELHPPKIEMPDNTNAAGPTST PHMINKVNMDQMDKRNLTISLDDNRLFPYVEMKAELIPNQNGSSKTTTINLRTEKEVWMGEIPTHQPDGEYTLKVTATDVLGIKKETNV EFVVDTTPPETPTIKLRDSVTKHAESKATQETTQEKPNFVIGSIPNDIYKINIKITGDDYDYDYEVGNLNNVENKLKEWSGIQNALKDNQT YSITVTFTDRAGNTSSNADTPFKIKKIVPYTFIDPTLILDPETDSGTQGDFKTNHNQPVLLFKNMSDAEVTEVNLKLENKTDKKTYTYSLK DFEPIHGKNSKEYKIELIKESSLTKEAPFPEGEYKVTVELTYASNEQYTKATTFEKPLTINRKKPQAVYDVKYTFKDGDQDQGKTINIEGK KPKNTDIFIQFHEQAEINIKKKNGDSLFSDDKFITTQPWGNRKALPSSLKTISATNPRILRSTYHSHLPSLDIVL*KTV*LRYLSKKKT*N*TIR LSLKIKLIMKYTNIQ*MKLI*KIKVFNLQLTKHSQTNLRLMSPGQSTKAKKLKKLYPKLLKMLNIPSQEAVVTSWWLKEN*KKDKH*I*AII MMIFLQKVTKLLEIKF**R*SHLIRI*GNLTCLLQMQMGEKVRLLKKKYPMSLFIGILSQNNIRLDLV*TILLNTFQTNSHAKKSVSILNMYW RCETVPHLLIMCTTRD*QALKNLIK*LSLITR*IMS*QNISI*IKKKIRTRFMEIIMAIQQKLGGVSDTMILIVFQKN*IKTIVKLNLSIKNRL*NF YLDLRVSITI*VY*R*RQ*KKRIKT*WKPDGICGLKRFLLKEKTTLMEKKRI*SE*QSMKNCKKDI*N*L*QMLITKNKRLSLTQAKIMKHLK *R*QEASKGKGLQ*TTGL*EFF*KIKIQVLSLILER*RLVE*HLMILQRLKMARKTLPQ*KFFQVGKNISEKTILKAEANVA*KKTL*KNQTT TPLKRRCIKSTRQQQIKTTPACHSQNSAPHSQSKRKTL*TLRNSQST*LNPKKRISLLSWWRNRRRPPHKILLSSTTPSTQRKTPPWISPPAS HSTPRKKRLMPSKRWSPWIIAPNIRLS*LITKGYSLSRCP*QKARMNSR*NLSIRMGIGSAWIKRSPSMRLLKEY*AAWKPQTDIKATFQQG QKRSRQKKMQIF***RLFYICLSMRRNLCIMD
>32		>32	>BJP42_08825_Transcription_accessory_protein_(S1_RNA-binding_domain)
	DTAVT	VSEAG	MIKIMMNSLNHLLAEEIKADPKQVSCAVQLLDEGNTVPFIARYRKEITGGLDDTQLRLLESRLGYLRELEDRRHSILKSIEEQGKLTDALA NAINSTLNKAELEDLYLPYKPKRRTKGQIAIEAGLESLAETLWQDPHQDPEILAQQYINAEKNITDTKSALEGARYILMERFAEDAALLSK VRQYLWKNAHFVSQVLPEKEEVGAKFRDYFDYQEPIAKIPSHRALALFRGRSEGVLQLALKVLSDEDPSEAILVEHFNIRLNNAPSDHWR KTVLQWAWRVKILPHLETELMGNVREMAQEQAIEVFALNMKDLLMAAPAGMRPTMGMDPGLRTGVKIAVVDATGKLLATDTVYPHT GQAAKAAIRVAHLCLEHQVEWVAIGNGTASRETERFYAEVQKQYPEVNSQKVIVSEAGASVYSASELAAEEFPDLDVSIRGAVSIARRLQ DPLAELVKIDPKSIGVGQYQHDVSQTQLAKKLTAVVEDCVNSVGVDLNTASVPLLTRVAGLSTSMAKNIVNWRDLNGRFIHRQQLLNIN QVGPKAFEQSAGFLRINNGDNPLDASTVHPEAYPVVENILKSTKQHLSQLMGNISILRSLNPENFVNDRFGLPTVRDILKELEKPGRDPRPE FKTPIFSDEIETLKDLIPGMILQGSVTNITHFGAFIDIGVHQDGLVHISSMANQFVTDPHQLIKTGDIVQVKVLSIDMERKRIALSMRLDEQIA DKKDKTASENASKEKRLTKKRPLKNAPISHFKNSALSDAFASAIKNKRS
			>BJP42_08930_Inositol-1-monophosphatase

Supplemental Table 4.2. Unique peptide sequences from active fractions mapped back to APSE3 and *H. defensa* strain AS3 genes with complete sequences.

			MHPMLNIAIRATRKAGNFIAKCYETPSYNQENNEFVSFVQQQAAQRIIETIHQSYPTHIVTTKYDDAIPEKESDMKWIVDPINGINNFLKRL PHFSISVALRVKGRTEIAIIYDPIRNELFTTTRGQGAQLNGYRLRMNNKTKDLTGSTISTLLSIEQKKHSPHELSQLIMLLINKKMNFHCSGS SALDFAYVAAGRLDSFFAIAIKKQDFWGGEMLVSEAGGLVTDLSGGHDHEISENIIAAHPMMIKALLKAIGEIKMDKE



Supplemental Figure 4.1. Molecular weight cutoff fractions from uninfected *H. defensa* strains A2C and AS3 have no effect on *A. ervi* development in vitro. Proportion of *A. ervi* developing to the critical 128-cell stage. Mother samples consisted of remixed molecular weight cutoff fractions and had no activity in either the A2C (A, X^2=0; df=4; p=1) or AS3 (B, X^2=0; df=4; p=1) strains. Similarly, no single fractions had activity against parasitoid development.



6lo

Supplemental Figure 4.2. HPLC fractions showing activity against *A. ervi* development in vitro. Schematic showing HPLC readouts with individual active fraction groups highlighted (A). Fractions were first pooled into groups of four or more based on spectrum readouts so that peaks were fully encompassed in a single pool. These pools were then bioassayed to test for *A. ervi* killing activity (B). Proportion of *A. ervi* developing to 128-cell stage in the presence of pooled HPLC fractions. Pooled fractions 14-17 were found to have activity even after one 1/10 dilution and was therefore selected for single fraction bioassays (X^2=12.12; df=5; p≤0.001). Single fraction bioassays revealed that fractions 15 and 16 significantly reduced the number of *A. ervi* developing to the 128-cell stage while fractions 14 and 17 did not (C, X^2=12.12; df=4; p≤0.001). Asterisks indicate treatment was significantly different than control of untreated TC100 + 30% FBS.