Title
Evolutionary Effects of a Eusocial Host Life History on an Endosymbiont

Author
Tolley, Sarah

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Evolutionary Effects of a Eusocial Host Life History on an Endosymbiont

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biology

by

Sarah Jo Auli'i Tolley

2019

ABSTRACT OF THE DISSERTATION

Evolutionary Effects of a Eusocial Host

Life History on an Endosymbiont

by

Sarah Jo Auli'i Tolley

Doctor of Philosophy in Biology

University of California, Los Angles, 2019

Professor Peter Nicholas Nonacs, Chair

The coevolution between hosts and their endosymbionts (organisms that live only in the host's body and cells), has dramatically shaped the history of life. Symbiotic relationships can range from mutualism to parasitism, as observed in the diverse relationships between arthropod hosts and obligately intracellular α-proteobacterium from the genus, *Wolbachia*. *Wolbachia* bacteria may infect as many as 70% of all insect species where it primarily exhibits reproductive parasitism, thereby inducing female-biased sex ratios to favor its own maternal transmission. In addition to altering host physiology, *Wolbachia* has been proposed as a pest control and means of mitigating disease vectoring. Despite its abundance and promising applications, the effects of *Wolbachia* on the development, reproduction, and defensive biology of eusocial insects have remained elusive.

In chapter 1, I explore how horizontal gene migration from endosymbionts to host nuclei may be biased to resolve intergenomic conflicts of interest in favor of the host. This genomic reorganization may help explain why many horizontally transferred *Wolbachia* genes have become non-functional and why genome reduction is greater in mutualistic strains compared to parasitic strains. Chapter 2 investigates an unexplored relationship between *Wolbachia* and leaf-cutter ants and potential mechanisms of *Wolbachia* transfer across host species. The results suggest not all ant-associated *Wolbachia* strains have the same genetic potential for horizontal transmission. In Chapter 3, I use whole genome sequencing and stage-specific RNA-sequencing to delve deeper into the genomic mechanisms of *Wolbachia* interactions across host ant developmental stages. Rather than rely on a single process or pathway, this symbiosis likely relies on a symphony of complex interactions that fluctuate over the development of the host. Genome and transcriptome analyses provide a foundation for future research exploring the functional basis and developmental gene expression dynamics for this widespread endosymbiont infecting a dominant tropical herbivore.

The dissertation of Sarah Jo Auli'i Tolley is approved.

James O. Lloyd-Smith

Robert K. Wayne

Jacob A. Russell

Peter Nicholas Nonacs, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the mentorship, support, collaboration, and encouragement from a great deal of wonderful people. It has been an absolute privilege to learn from so many talented people. First and foremost, I want to thank my advisor, Peter Nonacs, for taking a chance on me and never doubting my ability to succeed. I would also like to thank my thesis committee, Jamie Lloyd-Smith, Bob Wayne, and Jacob Russell for their generous time, guidance, and scholarly feedback.

I am indebted to Panos Sapountzis, who mentored me while I was on a research fellowship at the University of Copenhagen Centre for Social Evolution (CSE), for the immeasurable amount of time he permitted me to bother him with research questions. Without his advice and encouragement, I would have been truly lost. I would also like to thank Koos Boomsma for the incredible opportunity he afforded me to conduct research and collaborate with all the kind, gifted people at his CSE.

I am extremely grateful for the opportunity to collaborate with Fukatsu-san, Koga-san, and the rest of Japan's Advanced Industrial Science and Technology (AIST) community. I learned so much during my time with the Fukatsu Lab and am grateful for all the mentorship I received. Thank you to Bill Wcislo, Paola Galgani, Kristen Saltonstall, and Hermogenez at the Smithsonian Tropical Research Institute for housing me, guiding me, traversing the tropics with me to find ant colonies, and helping me navigate the complex world of Panamanian export permits. At UCLA, I am so appreciative of Alice Mouton's patience in helping me troubleshoot my genomic data analyses.

I would not be where I am today without the love and support of my parents who always did everything they could to help me pursue my passions. Thank you to my big brother, Grant,

for instilling a passion for science in me and for being such an inspiring role model. I am fortunate to have Joe and my loyal hound, Kona, in my life because they give me a reason to smile and laugh every day. Finally, thank you to all my friends for the love, encouragement, and good times throughout the years.

I am incredibly grateful for the funding that made this work possible. I have been fortunate to receive support during my dissertation from a National Science Foundation (NSF) Graduate Research Fellowship (GRFP) under Grant No. DGE-1650604. My research was additionally supported by the supplementary GRFP fellowships, Graduate Research Opportunities Worldwide (GROW) and Graduate Research Internship Program (GRIP). This work was also funded by several UCLA Research and Travel Grants and a Mayhew Graduate Research Award. The contents of this work are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

Chapter 1 is a reprint of a published paper –Nonacs, P. & Tolley, SJ. 2014. Certainty versus stochasticity: Cell replication biases DNA movement from endosymbionts and organelles into nuclei. *Evolutionary Ecology Research*. 16(3): 195-202. Both authors contributed to the conceptualization, writing and editing of the manuscript. The reprint is used here with the permission of co-author, Peter Nonacs, and the journal, Evolutionary Ecology Research. Chapter 2 is a reprint of a published paper – Tolley, SJA, Nonacs, P, Sapountzis, P. 2019. *Wolbachia* Horizontal Transmission Events in Ants: What Do We Know and What Can We Learn? *Frontiers in Microbiology.* 10:296. ST performed the experiments and conducted formal analysis of the data with guidance and supervision from PS and PN. ST wrote the original draft of the manuscript. PS and PN reviewed and edited the manuscript. The paper is used here with the permission of all co-authors and Frontiers in Microbiology.

VITA

SARAH JO AULI'I TOLLEY

EDUCATION

BS, *cum laude,* Biology, University of Colorado, Boulder, 2013

FELLOWSHIPS AND AWARDS

National Science Foundation Graduate Research Fellow (GRFP), 2014-2019, Total $138,000

UCLA Graduate Research and Travel Award, 2014 & 2018, Total $9,190

NSF Graduate Research Opportunities Worldwide (GROW), 2015 & 2016, Total $30,000

NSF Graduate Research Internship Program (GRIP), 2016, Total $5,000

Mayhew Graduate Research Award, 2014, Total $452

PUBLICATIONS

Tolley, S.J.A., Nonacs, P., Sapountzis, P. (2019) *Wolbachia* Horizontal Transmission Events in Ants: What Do We Know and What Can We Learn? *Frontiers in Microbiology.* 10:296. doi: 10.3389/fmicb.2019.00296.

Nonacs, P., Tolley, S.J. (2014) Certainty versus stochasticity: Cell replication biases DNA movement from endosymbionts and organelles into nuclei. *Evolutionary Ecology Research*. 16(3): 195-202.

CHAPTER 1:

CERTAINTY VERSUS STOCHASTICITY:

CELL REPLICATION BIASES DNA MOVEMENT

FROM ENDOSYMBIONTS AND ORGANELLES

INTO NUCLEI

Evolutionary Ecology Research, 2014, **16**: 195–202

Certainty versus stochasticity: cell replication biases DNA movement from endosymbionts and organelles into nuclei

Peter Nonacs and Sarah J. Tolley

Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, California, USA

ABSTRACT

Background: Endosymbiotic bacteria such as *Wolbachia* spend their entire life histories within other organisms' cells. This close proximity of endosymbiont and host genomes allows for transfers of DNA between them. Such events are observed to be strongly biased, however, with overall DNA migration from cytoplasmic elements to host nuclei.

Question: Are DNA transfers from cytoplasmic to nuclear genomes more likely to be retained than those in the opposite direction based on how mitotic and meiotic cell division disperses nuclear and cytoplasmic DNA to daughter cells?

Mathematical model: Simulations track the survival of individual DNA intergenomic transfers in populations across 100 non-overlapping generations. Reproduction is separately modelled as either asexual in a haploid species or sexual in a diploid species.

Key assumptions: Transfers can either have no effect or increase chances of host reproduction by up to 20%. The distribution of genomes into offspring is stochastic (i.e. a given modified genome is as likely to be transmitted as an unmodified one).

Conclusions: Even when DNA transfers are equally bidirectional, transfers into host nuclei are retained more often than ones into cytoplasmic genomes. Consequently, biased migration has potential consequences for life-history evolution, whereby genes that exchange locations also switch 'sides' for intergenomic conflict. Thus, biased migration of genes is a long-term evolutionary process favouring host interests over that of their endosymbionts and organelles.

Keywords: endosymbiont, horizontal gene transfer, intergenomic conflict, mitochondria, *Wolbachia*.

INTRODUCTION

Obligate endosymbionts are organisms whose entire life history is played out within the cells of other species. Hence the fate of an endosymbiont is inextricably entwined with that of its host. If the host fails to survive or reproduce, the endosymbiont suffers the same

Correspondence: P. Nonacs, Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, CA 90095, USA. e-mail: pnonacs@biology.ucla.edu

catastrophic loss in fitness. This connection limits the degree to which any endosymbiont can exploit its host without causing its own extinction. Nevertheless, endosymbionts and hosts can experience a range of conflicting evolutionary interests (Burt and Trivers, 2006). Conflict can arise because host genomes are always transmitted during cell division, while endosymbiont genomes may not be. For example, sperm cells contain replicated host germ line DNA, but endosymbionts are usually absent. Hence males could be valued very differently from the perspective of DNA in nuclei versus DNA in the cytoplasm. One way such a potential conflict could be minimized or eliminated is through horizontal gene transfer (HGT) across genomes. If the same genes are found in both locations, there would be no intergenomic conflict between the two sets of DNA.

Although conflict resolution should equally favour HGT in both directions, the overwhelming majority of reported DNA movement is from endosymbiont to host. Prokaryotic DNA has been found in numerous eukaryotes (reviewed in Dunning Hotopp, 2011), but with only a couple of known instances of the reverse (e.g. Woolfit *et al.*, 2009; Duplouy *et al.*, 2013). This biased migration of genetic material occurs not only in parasitic bacteria such as *Wolbachia* (Saridaki and Bourtzis, 2010), but also in the evolution of mitochondrial and chloroplast organelles. Mitochondria have lost genes to their host's nuclei with little to no migration in the opposite direction (Adams and Palmer, 2003; Brandvain and Wade, 2009). For yeast, Berg and Kurland (2000) estimate approximately one transfer per 10^5 generations from mitochondria to nuclei and less than one transfer per 10^{10} generations in the opposite direction.

Biased migration could result from the greater certainty of vertical transmission for nuclear DNA over cytoplasmic DNA during cell division. Consider a haploid cell with an endosymbiont population. Any DNA transfer into the nuclear chromosome would thereafter be represented in all daughter cells. In contrast, a transfer into a single cytoplasmic element would not be represented in all daughters because cell division tends to distribute cytoplasmic elements randomly (Burt and Trivers, 2006). In a stable population (i.e. only one daughter cell, on average, survives), HGTs into endosymbionts risk being stochastically eliminated in every generation. In diploid species, cytoplasmic HGTs face an added 50% mortality factor as they are almost always only maternally transmitted. For endosymbionts or organelles, any transfer located in a male body has no evolutionary future. This means that even if the likelihood of DNA transfers and fitness benefits are all equal across gene location, there could still be an apparent evolutionarily bias towards DNA moving from cytoplasm to nucleus due to the dynamics of cell division. We estimate by simulation these combined effects of stochasticity in transmission and sexual reproduction in terms of producing biased migration.

MODEL

Simulated populations are 100 'hosts', imagined as either single-celled organisms (the asexual 'haploid' condition) or gamete-producing cells within multicellular organisms (the sexual 'diploid' condition). Each host is infected with 100 endosymbionts with equal fitness effects. The endosymbiont's effect on its host's fitness could be neutral, positive (as in the case of *Buchnera* in aphids), or negative (as in many infections by *Wolbachia*). The model does not differentiate between these alternative evolutionary relationships. Instead, we concentrate on the relative change in host fitness due to genetic lateral transfer. This can be viewed as negative infections becoming less harmful because transferred genes are less effective at manipulating hosts, or that a positive mutualism is proportionally enhanced.

Simulations start with one lateral transfer of genetic material within one host, into either one nuclear chromosome or one genome of one endosymbiont in the within-cell population. Simulations can run up to 100 non-overlapping generations (i.e. the population at $t + 1$ is the offspring of the population at time t), or until the transfer is lost from the population, with no additional transfers occurring during this period. Every combination is replicated 20,000 times.

Transfer events can be fitness-neutral in terms of the endosymbiont's effect on its host's relative likelihood of producing offspring. This is analogous to transferring non-coding DNA, non-functional genes, or genes whose fitness consequences are unaffected by location. Alternatively, transfer events can be fitness-positive, generating a 10–20% greater likelihood of reproducing relative to unmodified individuals. This is analogous to functional genes being better expressed in their new location. Fitness-positive effects are not dosage-dependent: one modified endosymbiont produces equal benefit as do several, and homozygotes and heterozygotes are equally fit in diploids.

In the haploid condition, 100 randomly chosen hosts divide into two cells with one cell randomly selected to survive to the next generation. Haploid chromosomes containing transfers are passed to both daughter cells. For endosymbionts, populations within hosts first replicate. Each extant endosymbiont makes one identical copy. These doubled populations are randomly divided across daughter cells. Therefore, stochastic distribution may result in daughter cells not having the same number of modified endosymbionts as parents. The model assumes that host reproductive success is not influenced by the number of endosymbionts, as seen in a commensal relationship. Clearly, if endosymbionts had positive or negative effects, then number per cell would be an important variable, but endosymbiont load is a separate evolutionary process from lateral transfer effects. We do not, therefore, consider it in this model.

In the diploid condition, 100 hosts are randomly chosen as mothers and randomly mated to another host from the population. One randomly chosen chromosome from each parent is assigned to a single offspring. Thus, offspring could gain modified nuclear chromosomes from either or both parents. For the endosymbiont population, however, offspring receive endosymbionts only from their mother. To simulate meiosis, two cell divisions produce the single, functional egg cell and three evolutionarily dead-end polar bodies. The endosymbiont population replicates as in the haploid case prior to each cell division. Thus, modified endosymbionts can pass into future generations only if they are segregated into eggs.

By chance, some hosts will be selected to produce more than one offspring (particularly if they contain a fitness-positive transfer), and others will not be selected. Therefore, transfers can be lost from populations in several ways: (1) all modified hosts fail to reproduce; (2) with diploidy, modified chromosomes are lost through stochasticity in meiosis; (3) all modified endosymbionts are lost through stochasticity inherent to mitotic or meiotic cell division.

RESULTS

Fitness-positive transfers (10% added benefit) are more likely to be retained across 100 generations than fitness-neutral ones, and transfers to nuclei are more likely to survive than transfers to endosymbionts (Fig. 1). Sexual reproduction increases this asymmetry in survival probabilities of transfers, because unlike nuclear inheritance, any transfer into an

endosymbiont is transmissible only if it occurred in a female. With a fitness increase of 10% (Fig. 1a) or a neutral effect (Fig. 1b), both asexual (haploid) and sexual (diploid) modified chromosomes are considerably more likely than modified endosymbionts to survive. The same qualitative difference across nuclear and cytoplasmic locations remains when the positive effect is doubled to 20%. Modified chromosomes survive in 31.1% and 31.0% cases for asexual and sexual reproduction, respectively. Transfers into endosymbionts survive 6.7% of the time with asexual reproduction and 4.3% with sexual reproduction.

Under positive selection host ploidy has no effect on transfer survival, but under neutral conditions haploid transfers are maintained slightly more often because genetic drift is stronger in the smaller chromosome population.

Transfers are most often lost when rare and susceptible to stochasticity in reproduction (i.e. early in simulations: Fig. 1). If not quickly lost, fitness-positive transfers spread and most individuals in populations are carriers after 100 generations (Fig. 2a: for a 10% positive benefit). Conversely, fitness-neutral transfers must spread through drift, becoming common in relatively few populations (Fig. 2b).

DISCUSSION

Simulations of DNA transfers between endosymbiont and host genomes show that the dynamics of cell replication and division can create a bias for retaining modifications in nuclei over cytoplasmic elements. Horizontal gene transfers (HGTs) into haploid chromosomes of asexually reproducing hosts transmit with certainty into all daughter cells, but a modified cytoplasmic element would be found in only one daughter. The difference is further exaggerated with diploidy and sexual reproduction because both sexes contribute genetic material to their offspring's nuclei, but only mothers pass on endosymbionts. This would create a migration bias of overall movement of DNA flowing from endosymbionts to hosts. This bias is present both when transfers are fitness-enhancing (e.g. functional genes) or without significant fitness consequences (e.g. mobile genetic elements or pseudogenes). The results may even underestimate the bias in migration for positive transfer events by assuming dosage independence for endosymbionts. One modified endosymbiont in a population of 100 may have considerably less effect than one modified nuclear chromosome in a population of one or two.

Gene migration extrapolated over evolutionary time could make endosymbiont 'bodies' redundant and subject to elimination [possibly evidenced in a currently uninfected mosquito and filarial nematodes species, but with *Wolbachia* genes in their genomes (Klasson *et al.*, 2009; McNulty *et al.*, 2010)]. Ioannidis *et al.* (2013) estimate that over 10% of a *Wolbachia* genome has moved into its nematode host's DNA, and that the entire *Wolbachia* genome is potentially transferable. In contrast to the movement of DNA into nuclei, simulated HGTs into endosymbionts had less than a 0.1% chance of being retained unless the transfer increased overall host reproductive fitness.

The mechanics of cell division are an addition to the proposed factors for biases in accumulation. For instance, endosymbionts and organelles can replicate and recycle numerous times before cell division, meaning newly incorporated DNA can be lost due to stochastic processes (Berg and Kurland, 2000), and that pools of 'escaped genes' available for transfer are likely dominated by non-nuclear sources (Adams and Palmer, 2003). Alternatively, bias in DNA distribution could reflect selection after transfer. Endosymbionts and organelles still need to compete for within-cell resources for replication, leading to strong selection to

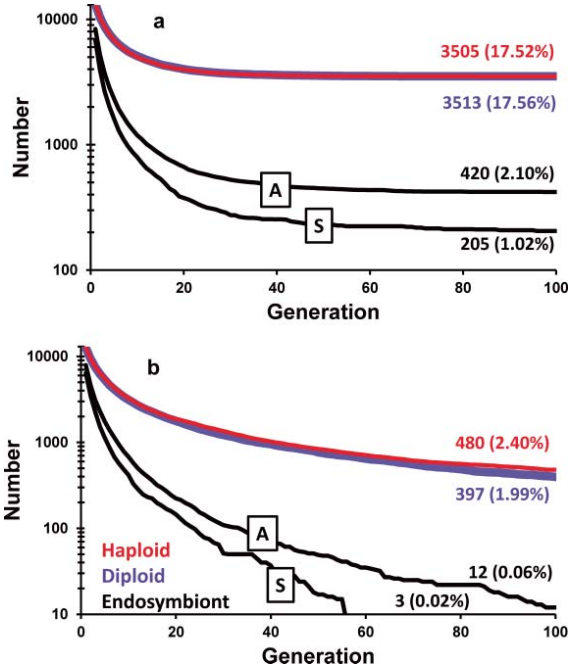


Fig. 1. Number of populations having at least one individual with a modified genome over time. All simulations begin with one transfer event in one individual and numbers denote how many populations (from 20,000) have such individuals after 100 generations. (a) Transfer events increase fitness of affected individuals by 10%. (b) Transfers are selectively neutral. Hosts are diploid, reproducing sexually and meiotically (S) or haploid, reproducing asexually and mitotically (A).

excise non-beneficial transfers to streamline genomes (Kurland, 1992). Also, if mutation rates differ between nuclear and non-nuclear DNA, replicated genes could mutate to non-viable or deleterious versions more often in organelles and endosymbionts (Berg and Kurland, 2000; Brandvain and Wade, 2009). Purifying selection, genetic drift, and replicative efficiency would then favour their loss from cytoplasmic genomes. This is a horizontal ratchet where reacquisition of lost genes is unlikely, and DNA therefore amasses in nuclei (Doolittle, 1998).

Biases created through cell division do additionally predict evolutionary outcomes that specifically correlate to transmission mode effects on HGT survival. First, HGTs are more prevalent in asexual plants than outbred species (Brandvain *et al.*, 2007). Second, the movement of DNA from endosymbionts into hosts accelerates when the endosymbiont switches from a facultative to an obligate life history. This increased movement occurs even though obligate organisms have 4–5 times fewer mobile DNA elements in their genomes (Toft and Andersson,

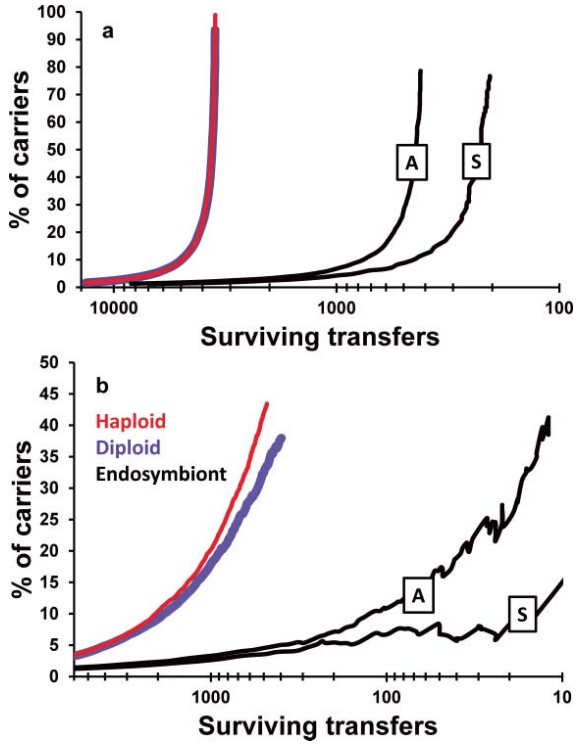


Fig. 2. Mean percentages of individuals having modified genomes in populations where modified genomes are still present. (a) Transfer events increase fitness of affected individuals by 10%. (b) Transfers are selectively neutral. Hosts are diploid, reproducing sexually and meiotically (S) or haploid, reproducing asexually and mitotically (A).

2010). In both cases, a key difference in the compared groups is the increased certainty with which an HGT would be inherited by offspring when in the nuclear genome.

The biased migration model is consistent with the aforementioned examples, but unexplained patterns of gene movement remain. For example, HGTs into hosts of functioning genes from primary or obligate mutualists such as *Buchnera* and *Tremblaya* appear to happen much less often than similar HGTs from secondary or facultative endosymbionts, such as *Wolbachia*, which often negatively affect host reproduction (Nikoh *et al.*, 2010; Husnik *et al.*, 2013; Ioannidis *et al.*, 2013). Perhaps because in mutualistic relationships endosymbionts are often segregated into special cells (e.g. bacteriocytes), this creates both isolation from germ cells and reduced transmission stochasticity (McCutcheon and Moran, 2012).

Although endosymbionts and their hosts have a mutual interest for successful host reproduction, they can be in severe conflict over the details (Burt and Trivers, 2006). If we analogize them as potentially competing 'teams', then gene migration is like trading a player. Consider a *Wolbachia* gene that creates a female-biased sex ratio because, as a maternally inherited element, males are evolutionary dead-ends. If that gene is transferred horizontally, its existing effect would become instantaneously deleterious, because from its new vantage point being in males would be of great selective advantage. Interestingly, such changes in fitness objectives may help explain why many *Wolbachia* HGTs are genes that were or have become non-functional (Ioannidis *et al.*, 2013) and why genome reduction is greater in mutualistic *Wolbachia* strains than in those that manipulate host sex ratios (Toft and Andersson, 2010). Overall, the long-term evolutionary diffusion of DNA from cytoplasmic to nuclear genomes could alter the balance of power in intergenomic conflict and resolve conflicts of interests in favour of the hosts. Many species of ants, for example, are infected with *Wolbachia*, but exhibit no colony-level deleterious effects or sex ratio irregularities (Russell, 2012). Furthermore, in many species of filarial nematodes, *Wolbachia* is an obligate mutualist, such that uninfected nematodes cannot survive (McNulty *et al.*, 2010). The degree to which the evolution of mutualistic and beneficial endosymbiotic relationships is a consequence of genomic reorganization is ripe for further exploration.

REFERENCES

Adams, K.L. and Palmer, J.D. 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol. Phylogenet. Evol.*, **29**: 380–395.

Berg, O.G. and Kurland, C.G. 2000. Why mitochondrial genes are most often found in nuclei. *Mol. Biol. Evol.*, **17**: 951–961.

Brandvain, Y. and Wade, M.J. 2009. The functional transfer of genes from the mitochondria to the nucleus: the effects of selection, mutation, population size and rate of self-fertilization. *Genetics*, **183**: 1129–1139.

Brandvain, Y., Barker, M.S. and Wade, M.J. 2007. Gene co-inheritance and gene transfer. *Science*, **315** (5819): 1685.

Burt, A. and Trivers, R.L. 2006. *Genes in Conflict: The Biology of Selfish Genetic Elements.* Cambridge, MA: Harvard University Press.

Doolittle, W.F. 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.*, **8**: 307–311.

Dunning Hotopp, J.C. 2011. Horizontal gene transfer between bacteria and animals. *Trends Genet.*, **27**: 157–163.

Duplouy, A., Iturbe-Ormaetxe, I., Beatson, S.A., Szubert, J.M., Brownlie, J.C., McMeniman, C.J. *et al.* 2013. Draft genome sequence of the male-killing *Wolbachia* strain wBol1 reveals recent horizontal gene transfers from diverse sources. *BMC Genom.*, **14**: 20.

Husnik, F., Nikoh, N., Koga, R., Ross, L., Duncan, R.P., Fujie, M. *et al.* 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell*, **153**: 1567–1578.

Ioannidis, P., Johnston, K.L., Riley, D.R., Kumar, N., White, J.R., Olarte, K.T. *et al.* 2013. Extensively duplicated and transcriptionally active recent lateral gene transfer from a bacterial *Wolbachia* endosymbiont to its host filarial nematode *Brugia malayi. BMC Genom.*, **14**: 639.

Klasson, L., Kambris, Z., Cook, P.E., Walker, T. and Sinkins, S.P. 2009. Horizontal gene transfer between *Wolbachia* and the mosquito *Aedes aegypti. BMC Genom.*, **10**: 33.

Kurland, C.G. 1992. Evolution of mitochondrial genomes and the genetic code. *Bioessays*, **14**: 709–714.

McCutcheon, J.P. and Moran, N.A. 2012. Extreme genome reduction in symbiotic bacteria. *Nature Rev. Microbiol.*, **10**: 13–26.

McNulty, S.M., Foster, J.M., Mitreva, M., Dunning Hotopp, J.C., Martin, J., Fischer, K. *et al.* 2010. Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. *PLoS ONE*, **5** (6): e11029.

Nikoh, N., McCutcheon, J.P., Kudo, T., Miyagishima, S.-Y., Moran, N.A. and Nakabachi, A. 2010. Bacterial genes in the aphid genome: absence of functional gene transfer from *Buchnera* to its host. *PLoS Genet.*, **6** (2): e1000827.

Russell, J.A. 2012. The ants (Hymenoptera: Formicidae) are unique and enigmatic hosts of prevalent *Wolbachia* (Alphaproteobacteria) symbionts. *Myrmecol. News*, **16**: 7–23.

Saridaki, A. and Bourtzis, K. 2010. *Wolbachia*: more than just a bug in insects [*sic*] genitals. *Curr. Opin. Microbiol.*, **13**: 67–72.

Toft, C. and Andersson, S.I.V. 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. *Nature Rev. Genet.*, **11**: 465–475.

Woolfit, M., Iturbe-Ormaetxe, I., McGraw, E.A. and O'Neill, S.L. 2009. An ancient horizontal gene transfer between mosquito and the endosymbiotic bacterium *Wolbachia pipientis. Mol. Biol. Evol.*, **26**: 367–374.

CHAPTER 2:

WOLBACHIA HORIZONTAL TRANSMISSION EVENTS IN ANTS: WHAT DO WE KNOW AND WHAT CAN WE LEARN?

frontiers in Microbiology

PERSPECTIVE
published: 06 March 2019
doi: 10.3389/fmicb.2019.00296

Wolbachia Horizontal Transmission Events in Ants: What Do We Know and What Can We Learn?

Sarah J. A. Tolley[1], Peter Nonacs[1] and Panagiotis Sapountzis[2]*

[1] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA, United States,
[2] Centre for Social Evolution, University of Copenhagen, Copenhagen, Denmark

While strict vertical transmission insures the durability of intracellular symbioses, phylogenetic incongruences between hosts and endosymbionts suggest horizontal transmission must also occur. These horizontal acquisitions can have important implications for the biology of the host. *Wolbachia* is one of the most ecologically successful prokaryotes in arthropods, infecting an estimated 50–70% of all insect species. Much of this success is likely due to the fact that, in arthropods, *Wolbachia* is notorious for manipulating host reproduction to favor transmission through the female germline. However, its natural potential for horizontal transmission remains poorly understood. Here we evaluate the fundamental prerequisites for successful horizontal transfer, including necessary environmental conditions, genetic potential of bacterial strains, and means of mediating transfers. Furthermore, we revisit the relatedness of *Wolbachia* strains infecting the Panamanian leaf-cutting ant, *Acromyrmex echinatior*, and its inquiline social parasite, *Acromyrmex insinuator*, and compare our results to a study published more than 15 years ago by Van Borm et al. (2003). The results of this pilot study prompt us to reevaluate previous notions that obligate social parasitism reliably facilitates horizontal transfer and suggest that not all *Wolbachia* strains associated with ants have the same genetic potential for horizontal transmission.

Keywords: symbiosis, fungus-growing ants, horizontal transmission, social interactions, *Wolbachia*

OPEN ACCESS

Edited by:
George Tsiamis,
University of Patras, Greece

Reviewed by:
Sandra Breum Andersen,
NYU Langone Health, United States
Aurélien Vigneron,
Yale University, United States
Ana Ješovnik,
Smithsonian Institution, United States

****Correspondence:***
Sarah J. A. Tolley
stolley@ucla.edu

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *09 October 2018*
Accepted: *04 February 2019*
Published: *06 March 2019*

Citation:
Tolley SJA, Nonacs P and Sapountzis P (2019) Wolbachia Horizontal Transmission Events in Ants: What Do We Know and What Can We Learn? Front. Microbiol. 10:296. doi: 10.3389/fmicb.2019.00296

INTRODUCTION

Wolbachia pipientis is a maternally inherited α-proteobacterium widely found in arthropods (Werren et al., 2008). *Wolbachia* exhibits reproductive parasitism in most arthropod species by manipulating the reproductive physiology of hosts and inducing female-biased sex ratios via one of four mechanisms: cytoplasmic incompatibility, feminization of genetic males, male-killing, or enforcing parthenogenesis (for reviews see Stouthamer et al., 1999; Werren et al., 2008). Although transmission within species is strictly vertical, *Wolbachia* phylogenies rarely correspond to host phylogenies, suggesting horizontal transmission (HT) also occurs (Zhou et al., 1998; Vavre et al., 1999; Raychoudhury et al., 2009; Stahlhut et al., 2010; Ahmed et al., 2013).

Horizontal transmission of intracellular bacterial symbionts require intimate tissue-level interaction between current and future hosts; predator–prey and host–parasitoid relationships have therefore been proposed to explain observed HT events (e.g., Heath et al., 1999; Noda et al., 2001; Yang et al., 2013; Carvalho et al., 2014; Ahmed et al., 2015; Mascarenhas et al., 2016). Ants are the

only lineage of social Hymenoptera where permanent social parasites, closely related to their host, commonly invade mature colonies (Boomsma et al., 2014). As ants are perennial, intimate inquiline cohabitation where social parasites live with hosts across generations offers ample opportunities for HT. This idea was first explored in *Acromyrmex echinatior* colonies, which are considered closed systems for endosymbionts since workers are highly aggressive toward non-nestmates (Larsen et al., 2014). Colonies can, however, be infiltrated by socially parasitic *Acromyrmex insinuator* queens, which invade and adopt the host colony odor (Lambardi et al., 2007; Nehring et al., 2015). Van Borm et al. (2003) first suggested that HT events occur between *Wolbachia* endosymbionts of *A. echinatior* and *A. insinuator* based on shared bacterial genotypes between cohabiting ant species. Later research suggested the same for *Solenopsis daguerrei*, a social parasite of *S. saevissima* (Dedeine et al., 2005; Martins et al., 2012) and for another fungus-growing ant, *Sericomyrmex amabilis*, and its social parasite *Megalomyrmex symmetochus* (Adams et al., 2013; Liberti et al., 2015).

In this perspective, we summarize research that has investigated *Wolbachia* HT events in ants, examine limitations of methods and study systems used, and propose future research. We also partially repeat one of the first field studies (Van Borm et al., 2003) characterizing *Wolbachia* endosymbionts of two fungus-growing ant species: the leaf-cutting ant, *A. echinatior*, and its social parasite, *A. insinuator*. Our pilot results, originating from a single *A. echinatior* colony parasitized by three *A. insinuator* queens, only partially confirmed these earlier findings. This highlights the importance of re-evaluating past and current methods and redirecting future efforts to include whole genome sequencing (WGS) data, which could increase the resolution of phylogenetic relationships and reveal pioneering insights into the genes and mechanisms that allow *Wolbachia* to jump to new hosts.

METHODS

A single *A. echinatior* queen and three parasitic *A. insinuator* queens cohabiting a mature colony (Ae724; collected in Gamboa, Panama, May 2015) were isolated in separate sterile petri dishes (similar to Stürup et al., 2014). After a 36-h period, ca. 40 eggs were collected from each queen and stored at −20°C. DNA was extracted using the DNeasy Tissue Kit (Qiagen) and a 603 bp region of the *Wolbachia* surface protein (*wsp*) was amplified using 81F/691R primers (Braig et al., 1998) and PCR conditions as described in Baldo et al. (2006b). PCR products were purified using the Invitek PCR purification kit, cloned using the TOPO TA cloning kit (Invitrogen, United States), and 24 colonies from each cloning were sent for Sanger sequencing (MWG, Germany). We checked chromatographs and removed primer sequences using Geneious (v. 9.0.4). Trimmed sequences (MG547478-MG547559) were queried against the non-redundant NCBI database to compile the top 100 hits. All sequences were aligned with ClustalW, sites with gaps were removed and sequences that could not align to the entire 426 bp reduced alignment were removed. Maximum Likelihood phylogenetic trees with 1,000 bootstrap iterations and the TVM+G model (jmodeltest v2.1.7) were run in Garli version 2.01.1067 (Zwickl, 2006). A consensus tree was configured in Geneious v9 (Kearse et al., 2012), and one representative *wsp* sequence from the same host species (>99%) was picked. The tree was further modified in FigTree v1.4.3 (Rambaut, 2016). As described in Baldo et al. (2006b), the strain profiles for each *wsp* sequence from this study was identified based on four conserved hypervariable regions (HVR) (**Supplementary Tables S1**, **S2**). Since eggs were pooled for sampling, coinfecting strains present in each species may not occupy the same individuals.

RESULTS

Van Borm et al. (2003) originally characterized nine *Wolbachia* infections: two strains in *A. echinatior*, four in *Acromyrmex octospinosus*, and three in their social parasite, *A. insinuator*. Some strains were specific to *Acromyrmex* ant species (A1 and B2), while others were present across multiple species (B1 and Bcons). Considering this earlier study was conducted >15 years ago, we reconstructed the phylogenetic relationships of previously identified *wsp* sequences (Van Borm et al., 2003), *wsp* sequences generated in our pilot study (from one host and three cohabiting parasitic queens), and closely related *wsp* sequences available on NCBI from other arthropod hosts (**Figure 1**). Using similar methods as Van Borm et al. (2003) (with the exception of extracting DNA from eggs rather than gynes), we identified three *wsp* genotypes named HVR1-3 (**Supplementary Tables S1**, **S2**). While HVR-1 was the dominant *A. echinatior* strain in our study (**Supplementary Figure S1**), it was not identified in the previous study. HVR-2 was identical (>99%) to strain B1 (AF472563; Van Borm study). We identified HVR-2 in both *A. insinuator* and *A. echinatior* while Van Borm et al. (2003) found HVR-2 only in *A. insinuator* and a closely related but not identical strain (Bcons) in *A. echinatior*. HVR-2 has also been found in *A. octospinosus* (Van Borm et al., 2003; Andersen et al., 2012). HVR-3 was identical (>99%) to strain B2 (AF472560; Van Borm study) and, as before, was only found in *A. insinuator*. Two strains identified before (AF472558-9) were not found in the colony we analyzed. The Van Borm study suggested multiple HT events occurred for *Acromyrmex* ants to acquire their *Wolbachia*, as evidenced by their findings showing distantly related *Wolbachia* strains shared by closely related *Acromyrmex* hosts and the reverse, closely related *Wolbachia* present in distantly related host species. Our results were consistent with the Van Borm study where *wsp* sequences from *Acromyrmex* hosts were most similar to those from distantly related *Solenopsis* fire ant hosts. Our new phylogeny also revealed additional ant hosts harboring closely related *wsp* sequences. HVR-2 seemed the most cosmopolitan strain in ant hosts as it is present in at least nine ant genera (**Figure 1**).

DISCUSSION

HT events are believed to have largely contributed to the *Wolbachia* pandemic, where an estimated 50–70% of all insect

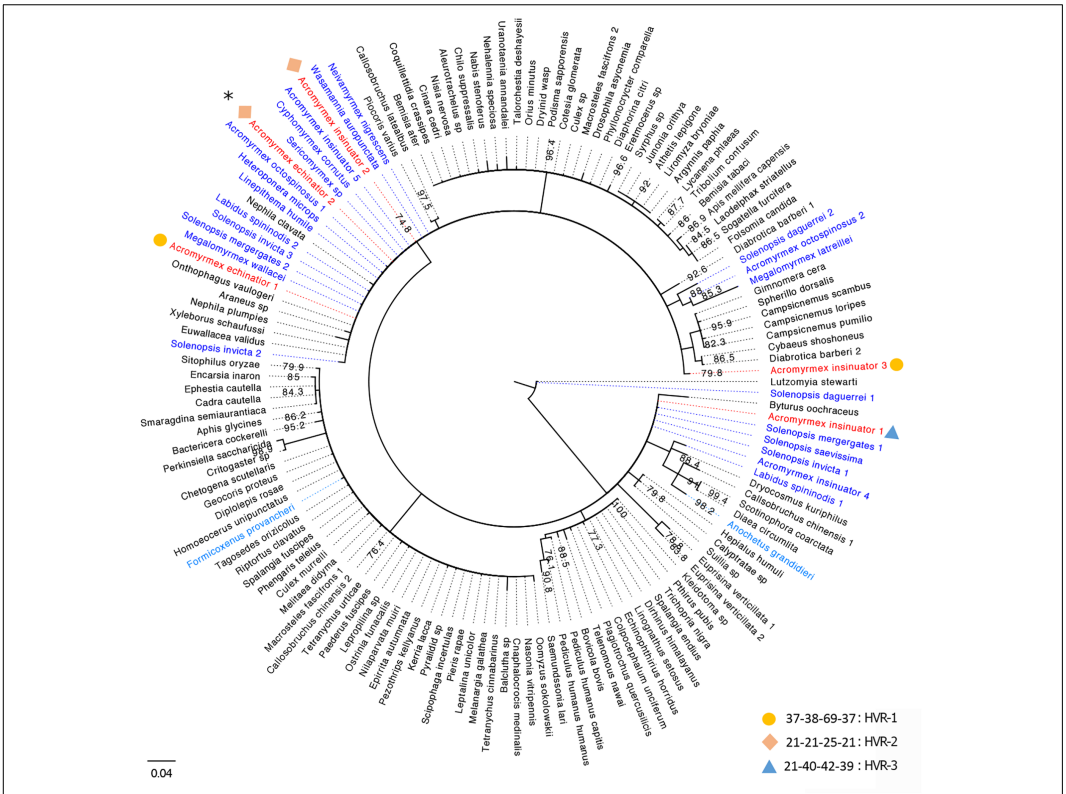


FIGURE 1 | Consensus tree for *Wolbachia* strains based on the *wsp* gene. Strains are represented by the infected arthropod host species with which they are associated. Host names based on sequences generated in the current study are colored red and those from previous studies on ant species are colored blue (Neotropical species in dark blue, two others in light blue). Sequences from *Wolbachia* strains associated with non-ant hosts are presented in black font and bootstrap support is shown at each node. In total, 24 unique *Wolbachia* sequences from ant hosts and 112 sequences from non-ant arthropod hosts were downloaded from GenBank and used in this phylogenetic analysis. Numbers next to species names represent genetically distinct strains harbored in the same species. Information about the HVR barcoding is given next to each of the *Wolbachia* strains identified in our study using circles, diamonds, and triangles (legend bottom right), for details see **Supplementary Tables S1**, **S2**. Since HVR barcoding was only analyzed for the 83 *wsp* sequences from the four *Acromyrmex* queens used in this study, only the representative strains from our study in red font are given a corresponding HVR type. Strains matching the Van Borm et al. (2003) isolates, B2 and B1, are labeled here as *Acromyrmex insinuator* 4 and 5. The majority of BLAST hits clustering with the *A. echinatior* and *A. insinuator* sequences generated in this study are *Wolbachia* strains from Neotropical New World ant species. The only BLAST hits from ant hosts that are not Neotropical New World ant species were phylogenetically isolated (light blue species; *Formicoxenus provancheri*, occurring in North America, and *Anochetus grandidieri*, a species endemic to Madagascar). The asterisk at the top left of the figure marks the HVR-2 strain that is widespread among mainly ants, but also other insect hosts in the Americas.

species are infected (Werren et al., 2008; Saridaki and Bourtzis, 2010; Weinert et al., 2015). High frequency of phylogenetic incongruences between hosts and *Wolbachia* strains (as seen in **Figure 2**) suggests HT events are relatively common on an evolutionary time scale despite the fact that they are difficult to predict and observe in nature. The results of our pilot experiment support the hypothesis that HT has occurred between *A. echinatior* and its social parasite, *A. insinuator*, originally proposed by Van Borm et al. (2003). As in the Van Borm study, we found distantly related *Wolbachia* strains occupying the same host (HVR-2 and -3 in *A. insinuator*) as well as identical strains occupying distantly related hosts (HVR-2; **Figures 1**, **2**). Although social parasitism should provide ample opportunity for HT, our results suggest some strains, like HVR-2, may be better equipped to "jump" between hosts. Although much about HT remains unknown, minimum conditions must be fulfilled for HT to occur: (1) there must be suitable environmental conditions (in the new host as well as the medium/environment the bacteria transitions through), (2) the bacterial strain must have the genetic potential for transfer, and (3) there must be a mechanism that will mediate the HT event.

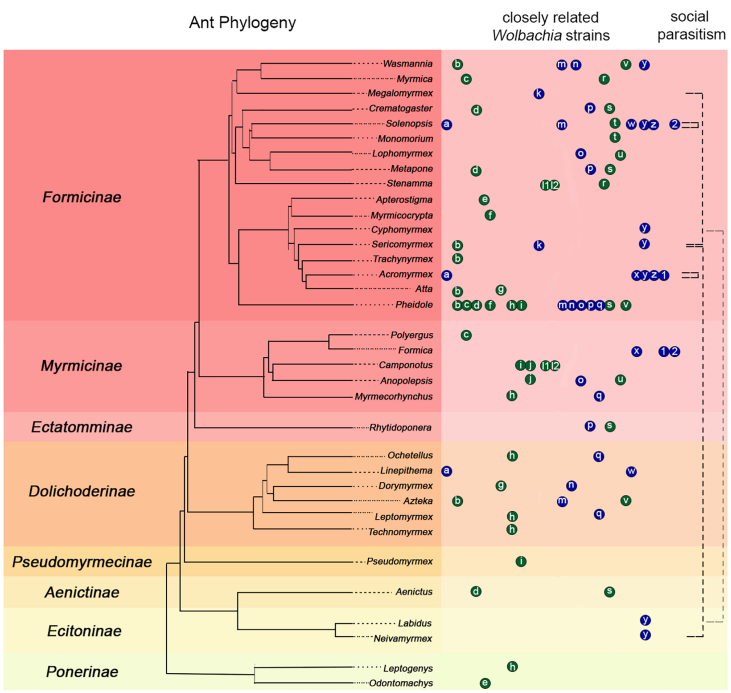


FIGURE 2 | Compilation of previous literature including ant phylogenies, *Wolbachia* infections, strain typing, and known connections between social parasitism and shared *Wolbachia* strains. The ant phylogeny on the left was constructed using data from Moreau et al. (2006), Schultz and Brady (2008), and Branstetter et al. (2017). Ant subfamilies are presented on the left of the phylogeny with colored backgrounds separating them. Colored circles with small case letters on the right of the ant phylogeny connect ant hosts suspected to carry similar *Wolbachia* strains based on previous phylogenies using the *wsp* gene (blue) or MLST typing (green). Letters in the circles indicate the published source where authors built phylogenies that showed potentially common *Wolbachia* strains in different host ant species: a: (Dedeine et al., 2005); b–j: (Frost et al., 2010); k: (Liberti et al., 2015); l1: (Ramalho et al., 2017); m–q: (Rey et al., 2013); h, l2, r–v: (Russell et al., 2009); w: (Fernando de Souza et al., 2009); x: (Tsutsui et al., 2003; Reuter et al., 2005); y: (this study); z: (Van Borm et al., 2003); 1, 2: (Viljakainen et al., 2008). Each letter circle occupying the same column represents a set of highly similar *Wolbachia* strains defined as belonging to the same clade (or a potentially common identical strain) according to the corresponding source publication. Circle order is alphabetical and distances between circles are not indicative of the genetic similarity of strains. Therefore, some heavily sampled genera, such as *Pheidole*, may have the same strain represented in more than one publication in which different ant genera were analyzed. The black dashed lines in the rightmost column connect known instances of social parasitism between ants based on previous literature (see text for details). Black dashed lines that start and end in the same genus highlight that *Solenopsis* and *Acromyrmex* have social parasites within their genera and they share identical *Wolbachia* strains with them (Van Borm et al., 2003; Dedeine et al., 2005; Martins et al., 2012; our study). The light gray dashed line connecting *Labidus* and *Cyphomyrmex* highlight an almost identical shared *Wolbachia* strain (differing only by 1 bp; **Figure 1**). However, there is no data suggesting *Labidus* predates *Cyphomyrmex* colonies.

Part 1: *Wolbachia* Genetic Potential

Wolbachia can be artificially transferred across insect genera in the lab (e.g., Zabalou et al., 2004; Hoffmann et al., 2011) and following transfers, adaptations to new hosts may rapidly occur (McMeniman et al., 2008). This ability to invade new hosts is consistent with the identification of genetically similar strains in taxonomically unrelated hosts (e.g., Heath et al., 1999; Raychoudhury et al., 2009). *Wolbachia* is obligately intracellular yet is capable of surviving extracellularly for several months before reinvading new cells and establishing a stable infection (Rasgon et al., 2006). Although mechanisms of natural HT remain elusive, *Wolbachia* has demonstrated the ability to successfully "jump" across cells, cross somatic tissues, and reach reproductive organs (Frydman et al., 2006; White et al., 2017). Successful transfers may be attributed to the bacterium's ability to adapt to new environments. This could be accomplished by recombination, likely mediated by inactive bacteriophages introducing "exotic genes," resulting in gene gains and diversification of the bacterium's genome (Wu et al., 2004; Klasson et al., 2009; Vos and Didelot, 2009; Ellegaard et al., 2013). Indeed, the *Wolbachia* genome has a high number of repetitive elements and ankyrins, mostly introduced by bacteriophages (Ishmael et al., 2009; Kent and Bordenstein, 2010; Leclercq et al., 2011; Siozios et al., 2013). While the function of these gene

gains has not been fully deciphered, genomic comparisons with a mutualistic strain infecting nematode hosts, *w*Bm (Foster et al., 2005), suggest they play a role in the bacterium's ability to induce reproductive phenotypes in arthropods.

Considering the significant genomic differences and tissue tropisms between *Wolbachia* strains, we expect not all strains have the same potential for transmission. For example, while *Wolbachia* is typically localized in the reproductive tract (e.g., *w*Mel, *w*Sty), there are some B-group strains that colonize somatic (non-reproductive) tissues (e.g., *w*No, *w*Ma; Veneti et al., 2004). As expected, not all strains can survive a transfer or induce reproductive phenotypes necessary to facilitate its spread in new host populations (Zabalou et al., 2008; Veneti et al., 2012). Phylogenetic comparisons using *wsp* sequences (Van Borm et al., 2003; **Figure 1**) also suggest that one of the strains in *Acromyrmex* (HVR-2) may have a greater propensity for HT than HVR-1 and HVR-3. HVR-2 is not only common across the Panamanian *Acromyrmex* species (*A. echinatior*, *A. insinuator*, *A. octospinosus*), where it has been identified as *w*SinvictaB (Andersen et al., 2012), but also in ant hosts across four subfamilies (**Figure 1**). In contrast, HVR-1 and HVR-3 appear specific to their respective host species and are far more dominant in those hosts than the shared HVR-2 (**Supplementary Figure S1**). This distribution suggests that HVR-1 and HVR-3 are better adapted to their respective host species while HVR-2 is a generalist capable of infecting hosts with diverse life histories. Interestingly, HVR-2 (*w*SinvictaB) appears to be dominant in *A. octospinosus* (Andersen et al., 2012), but occurs as either a single or double infection with the rare and sparse *w*SinvictaA (Andersen et al., 2012).

Part 2: Potential Transmission Routes in Ants

Ant sociality offers ample opportunities for *Wolbachia* transfer across hosts and may be especially favorable for species prone to interspecific social interactions or with less restrictive tissue tropisms. For example, fungus-growing ants are a host where *Wolbachia* has uncommon tissue tropism; it is present extracellularly in the gut lumen and may reach high titers in the hemolymph (Andersen et al., 2012; Frost et al., 2014; Sapountzis et al., 2015). A common resource, such as a fungal garden, may thus facilitate HT of *Wolbachia* strains between cohabiting *A. echinatior* and *A. insinuator*, as the ants deposit their feces in the fungus, feed on it, and cover their brood with it (which also feeds on the fungus). Similarly, an identical *Wolbachia* strain has been found between a workerless social parasite, *S. daguerrei*, and its host ant species *S. invicta* (Dedeine et al., 2005). However, a shared *Wolbachia* strain was not found between *M. symmetochus* social mercenaries and its host, *S. amabilis*, suggesting cohabitation does not always result in HT (Liberti et al., 2015).

Inquiline mites may also have the capacity to vector *Wolbachia* between attine species cohabiting the same nest or foraging on the same plants. However, mites in *Acromyrmex* nests appear to be saprophytic, not parasitic (Peralta and Martínez, 2013), making this alternative transmission route unlikely. Parasitic phorid flies could also serve as a common vector between all three ant species (Brown and Feener, 1998; Fernández-Marín et al., 2006; Pérez-Ortega et al., 2010; Guillade and Folgarait, 2015), however, so far there is no data suggesting they have contributed to HT events (Dedeine et al., 2005).

Independent of being intra- or extra-cellular symbionts, HT may also be mediated by predators such as *Neivamyrmex*, a genus of army ant known to raid nests of fungus-growing ants and consume their brood (Lapolla et al., 2002; Powell and Clark, 2004). Army ant taxa (subfamilies Aenictinae, Dorylinae, and Ecitoninae) are often infected with *Wolbachia* and thus offer exciting opportunities for studying potential HT (**Figure 2**). HVR-2 is distributed across species from the subfamilies Myrmicinae (*Acromyrmex* and *Sericomyrmex*) and Ecitoninae (*Neivamyrmex*; **Figures 1**, **2**). Similarly, an identical *Wolbachia* strain is shared between *Cyphomyrmex* and army ants of the genus *Labidus* (subfamily Ecitoninae; **Figure 1**), however, there is no known data confirming whether these army ants attack fungus-growing ants (**Figure 2**).

Part 3: Genomic Data and Sampling Power Limitations

Wolbachia strain typing has relied on several different genes, one of them being the 16S rDNA gene used when performing targeted sequencing (e.g., Kautz et al., 2013; Ramalho et al., 2017). This method is not appropriate to build phylogenies as the 16s gene is highly conserved and cannot distinguish closely related *Wolbachia* strains (Andersen et al., 2012). The *wsp* gene has also been used extensively for *Wolbachia* characterization because its rapid sequence evolution enables differentiation between closely related strains and it contains four HVRs useful in solidifying strain identification (Baldo et al., 2006b). However, the relatively short sequence length (<600 bp), high recombination rate (Baldo et al., 2005) and, in some arthropod hosts, strong positive selection (Jiggins et al., 2002), make *wsp* suboptimal for constructing phylogenies. Nevertheless, the *wsp* gene remains a useful "quick and dirty" approach to distinguish phylogenetic relationships of *Wolbachia* strains and is, in most cases, the only sequence available to build phylogenies. Due to these limitations, multilocus sequence typing (MLST) was introduced, which uses concatenated alignments of five housekeeping genes (Baldo et al., 2006a; Bordenstein et al., 2009). However, due to frequent recombination, WGS is the only accurate method to infer phylogenetic relationships (Bleidorn and Gerth, 2018).

A particular challenge to studying the evolutionary relationships of *Wolbachia* in arthropods is that hosts are frequently infected with multiple strains (Hiroki et al., 2004; Mouton et al., 2004; Frost et al., 2010; Andersen et al., 2012; Zhao et al., 2013), making even MLST and WGS approaches exceedingly challenging. *Acromyrmex* ants are one such example as they almost always contain multiple strains (Van Borm et al., 2003; Andersen et al., 2012) and we do not yet have *Wolbachia* genome data. *Wsp* typing has confirmed distinct, species-specific *Wolbachia* strains for *A. echinatior* (HVR-1) and *A. insinuator* (HVR-3) as well as a shared strain between the two species and *A. octospinosus* (HVR-2; Van Borm et al., 2003; Andersen et al.,

2012). Differences from this study and Van Borm et al. (2003) could mean strains are transient or that diversity is greater than what is currently known. On the other hand, differences may be related to limited ant colony sampling. Many ant species have wide geographic distributions (e.g., *Linepithema*, *Monomorium*, *Solenopsis*, *Atta*, and *Acromyrmex* genera) and show significant differences in infections among colonies and geographic locations (e.g., Reuter et al., 2005; Frost et al., 2010; Martins et al., 2012; Zhukova et al., 2017). Thus, despite previous efforts to illustrate *Wolbachia* HT events, success has been limited because we have only characterized small subpopulations and because *Wolbachia* may be evolving and spreading to new hosts faster than we currently study it.

Part 4: Implications for Future Research

Although limited, existing data suggests *Wolbachia* associated with ants are uniquely shaped by the ant microenvironment and have occasionally taken advantage of opportunities offered by the hosts' wide range of social interactions to "jump" to other ant species or genera. Comparisons between the widespread HVR-2 and less common strains, HVR-1 and -3, offer an exciting opportunity for future research because these strains (i) have different specificity to ant hosts (frequencies, infection levels), and (ii) have strikingly different distributions across phylogenetically distant ant hosts (although this may be driven by under-sampling). This suggests HVR-2 may have acquired (or lost) a set of genes that have facilitated its "ecological success." Future genomic comparisons may allow us to answer important questions about *Wolbachia* evolution and HT including, why strains like HVR-2 have greater ecological success (spread), and what genes and mechanisms are associated with the ability to spread successfully across distantly related host species.

The most reliable *Wolbachia* phylogenies have been built using WGS data (Klasson et al., 2009; Ellegaard et al., 2013; Gerth et al., 2014; Gerth and Bleidorn, 2016). These phylogenies have resolved important gaps in our knowledge of *Wolbachia* origin and supergroup diversification as they are typically built using conserved orthologs unaffected by recombination, which would render topologies invalid (Gerth et al., 2014; Gerth and Bleidorn, 2016). Further mapping of *Wolbachia* diversity on host ant trees and more genomic data, particularly involving ants not hailing from the Americas, will be required to assess biogeography patterns, such as whether there are specialized *Wolbachia* lineages infecting New World ants (Russell et al., 2009; Frost et al., 2010). The existence of major consortia like the GAGA project[1], which aims to sequence and perform comparative bacterial genomics for 200 ant genomes, shows tremendous promise for furthering knowledge of *Wolbachia* associations with a broader taxonomic host range. Comparative genomics (e.g., identification of selection signatures in genes) can shed light onto genetic prerequisites for HT. Besides advancing phylogenomic and comparative genomic approaches, WGS can provide insight into HT mechanisms for future functional studies (similar to Frydman et al., 2006; White et al., 2017) allowing us to pinpoint specific *Wolbachia* genes to relevant phenotypes.

[1]http://antgenomics.dk/

AUTHOR CONTRIBUTIONS

ST performed the experiments and conducted formal analysis of the data with guidance and supervision from PS and PN. ST wrote the original draft of the manuscript. PS and PN reviewed and edited the manuscript.

FUNDING

Funding was provided by an NSF Graduate Research Fellowship to ST (DGE-1650604).

ACKNOWLEDGMENTS

We thank Mariya Zhukova for helping with the collection of eggs and Jacobus Boomsma for comments on an earlier version of the manuscript. The Smithsonian Tropical Research Institute in Panama made facilities available and the Autoridad Nacional del Ambiente (ANAM) of Panama issued collection and export permits.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.00296/full#supplementary-material

FIGURE S1 | Relative proportion of HVR types 1–3 across *Acromyrmex* queen eggs. Pie chart showing the relative proportion of eggs sequenced from queens in this study with each HVR type, as described in **Supplementary Table S1**. The legend lists the HVR reference number used in this study followed by parentheses including the name of identical strains and the host species harboring those strains as described by Van Borm et al. (2003). Although closely related, the asterisk indicates that the strain Bcons was not a perfect match to HVR-2 from our study, unlike strain B1 (present in *A. insinuator* and *A. octospinosus*) which was identical to HVR-2.

TABLE S1 | Top BLAST matches and corresponding HVR type for *A. echinatior* and *A. insinuator wsp* sequences generated in this study. Summary of HVR typing results from **Supplementary Table S2** and additional *wsp* characterizations from the *A. echinatior* host queen (HQ) and the three parasitic *A. insinuator* (PQ) queens. From left to right: queen ant used to collect egg DNA samples for this study, host species and accession numbers for the top BLAST hits matching respective *Wolbachia* sequences, the geographic origin of NCBI samples of other ants, the average % match identity in BLAST, the number of sequences that returned these hits, HVR (hypervariable region) typing according to Baldo et al. (2006b) used in this study, and the percentage of sequences from each queen with respective HVR types. HVR typing is based on the four hypervariable regions of the *wsp* gene which are comprised of relatively conserved amino acid motifs used to identify recombination points, allowing discrimination between closely related *Wolbachia* strains, comparable to the use of antigens for serotyping pathogenic bacteria (Baldo et al., 2006b). The four HVRs occupy consecutive conserved regions of the *wsp* gene and the combination of all four HVR haplotypes make up its WSP profile. All *A. echinatior* and *A. insinuator* sequences were classified by their four HVRs using the PubMLST database[2] (best match to existing sequences in the database), which revealed three distinct *wsp* genotypes, here referred to as HVR 1-3. Color coding for the HVR types matches those

[2]https://pubmlst.org/wolbachia/

shown in **Figure 1**, except for one case where sequences were chimeras (type 21-38-69-37/21-21-25-37) of HVR-1 and HVR-2. As identified in other strains, recombination was localized in the HVRs, which suggests the two *A. echinatior* sequences are true chimeras rather than sequencing errors (Andersen et al., 2012).

TABLE S2 | Individual sequence information and alignment to HVR reference sequences. From left to right: colony of origin, accession numbers for sequences from this study, host species and accession numbers for the top BLAST hits matching respective *Wolbachia* sequences, % identity to BLAST match, closest match to sequences in pubMLST database (Baldo et al., 2006b) for *wsp* (nucleotide query), HVR1-4 (amino acid query) with sequence differences listed below. Colors coordinate with HVR types 1–3 used in **Figure 1**. Asterisks indicate sequences that differed in HVR reference number relative to other strains in the same HVR type (1–3) classified in this study. The three sequences with asterisks only exhibited a 1–3 bp difference from other sequences in their respective HVR characterization (types 1–3) and were thus included in analyses.

REFERENCES

Adams, R. M. M., Liberti, J., Illum, A. A., Jones, T. H., Nash, D. R., and Boomsma, J. J. (2013). Chemically armed mercenary ants protect fungus-farming societies. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15752–15757. doi: 10.1073/pnas.1311654110

Ahmed, M. Z., de Barro, P. J., Ren, S.-X., Greeff, J. M., and Qiu, B.-L. (2013). Evidence for horizontal transmission of secondary endosymbionts in the *Bemisia tabaci* cryptic species complex. *PLoS One* 8:e53084. doi: 10.1371/journal.pone.0053084

Ahmed, M. Z., Li, S.-J., Xue, X., Yin, X.-J., Ren, S.-X., Jiggins, F. M., et al. (2015). The intracellular bacterium *Wolbachia* uses parasitoid wasps as phoretic vectors for efficient horizontal transmission. *PLoS Pathog.* 10:e1004672. doi: 10.1371/journal.ppat.1004672

Andersen, S. B., Boye, M., Nash, D. R., and Boomsma, J. J. (2012). Dynamic *Wolbachia* prevalence in *Acromyrmex* leaf-cutting ants: potential for a nutritional symbiosis. *J. Evol. Biol.* 25, 1340–1350. doi: 10.1111/J.1420-9101.2012.02521.X

Baldo, L., Bordenstein, S., Wernegreen, J. J., and Werren, J. H. (2006a). Widespread recombination throughout *Wolbachia* genomes. *Mol. Biol. Evol.* 23, 437–449. doi: 10.1093/molbev/msj049

Baldo, L., Dunning Hotopp, J. C., Jolley, K. A., Bordenstein, S. R., Biber, S. A., Choudhury, R. R., et al. (2006b). Multilocus sequence typing system for the endosymbiont *Wolbachia* pipientis. *Appl. Environ. Microbiol.* 72, 7098–7110. doi: 10.1128/AEM.00731-06

Baldo, L., Lo, N., and Werren, J. H. (2005). Mosaic nature of the *Wolbachia* surface protein. *J. Bacteriol.* 187, 5406–5418. doi: 10.1128/JB.187.15.5406-5418.2005

Bleidorn, C., and Gerth, M. (2018). A critical re-evaluation of multilocus sequence typing (MLST) efforts in *Wolbachia*. *FEMS Microbiol. Ecol.* 94. doi: 10.1093/femsec/fix163

Boomsma, J. J., Huszár, D. B., and Pedersen, J. S. (2014). The evolution of multiqueen breeding in eusocial lineages with permanent physically differentiated castes. *Anim. Behav.* 92, 241–252. doi: 10.1016/j.anbehav.2014.03.005

Bordenstein, S. R., Paraskevopoulos, C., Dunning Hotopp, J. C., Sapountzis, P., Lo, N., Bandi, C., et al. (2009). Parasitism and mutualism in *Wolbachia*: what the phylogenomic trees can and cannot say. *Mol. Biol. Evol.* 26, 231–241. doi: 10.1093/molbev/msn243

Braig, H. R., Zhou, W., Dobson, S. L., and O'Neill, S. L. (1998). Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *J. Bacteriol.* 180, 2373–2378.

Branstetter, M. G., Ješovnik, A., Sosa-Calvo, J., Lloyd, M. W., Faircloth, B. C., Brady, S. G., et al. (2017). Dry habitats were crucibles of domestication in the evolution of agriculture in ants. *Proc. Biol. Sci.* 284:20170095. doi: 10.1098/rspb.2017.0095

Brown, B. V., and Feener, D. H. (1998). Parasitic phorid flies (Diptera: Phoridae) associated with army ants (Hymenoptera: Formicidae: Ecitoninae, Dorylinae) and their conservation biology. *Biotropica* 30, 482–487. doi: 10.1111/j.1744-7429.1998.tb00084.x

Carvalho, G. A., Corrêa, A. S., de Oliveira, L. O., and Guedes, R. N. C. (2014). Evidence of horizontal transmission of primary and secondary endosymbionts between maize and rice weevils (*Sitophilus zeamais* and *Sitophilus oryzae*) and the parasitoid *Theocolax elegans*. *J. Stored Prod. Res.* 59, 61–65. doi: 10.1016/j.jspr.2014.05.004

Dedeine, F., Ahrens, M., Calcaterra, L., and Shoemaker, D. D. (2005). Social parasitism in fire ants (*Solenopsis spp.*): a potential mechanism for interspecies transfer of *Wolbachia*. *Mol. Ecol.* 14, 1543–1548. doi: 10.1111/j.1365-294X.2005.02499.x

Ellegaard, K. M., Klasson, L., Näslund, K., Bourtzis, K., and Andersson, S. G. E. (2013). Comparative genomics of *Wolbachia* and the bacterial species concept. *PLoS Genet.* 9:e1003381. doi: 10.1371/journal.pgen.1003381

Fernández-Marín, H., Zimmerman, J. K., and Wcislo, W. T. (2006). *Acanthopria* and *Mimopriella* parasitoid wasps (Diapriidae) attack *Cyphomyrmex* fungus-growing ants (Formicidae, Attini). *Naturwissenschaften* 93, 17–21. doi: 10.1007/s00114-005-0048-z

Fernando de Souza, R., Daivison Silva Ramalho, J., Santina de Castro Morini, M., Wolff, J. L. C., Araújo, R. C., and Mascara, D. (2009). Identification and characterization of *Wolbachia* in Solenopsis saevissima fire ants (Hymenoptera: Formicidae) in southeastern Brazil. *Curr. Microbiol.* 58, 189–194. doi: 10.1007/s00284-008-9301-y

Foster, J., Ganatra, M., Kamal, I., Ware, J., Makarova, K., Ivanova, N., et al. (2005). The *Wolbachia* Genome of Brugia malayi: endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* 3:e121. doi: 10.1371/journal.pbio.0030121

Frost, C. L., Fernandez-Marin, H., Smith, J. E., and Hughes, W. O. (2010). Multiple gains and losses of *Wolbachia* symbionts across a tribe of fungus-growing ants. *Mol. Ecol.* 19, 4077–4085. doi: 10.1111/j.1365-294X.2010.04764.x

Frost, C. L., Pollock, S. W., Smith, J. E., and Hughes, W. O. H. (2014). *Wolbachia* in the flesh: symbiont intensities in germ-line and somatic tissues challenge the conventional view of *Wolbachia* transmission routes. *PloS One* 9:e95122. doi: 10.1371/journal.pone.0095122

Frydman, H. M., Li, J. M., Robson, D. N., and Wieschaus, E. (2006). Somatic stem cell niche tropism in *Wolbachia*. *Nature* 441, 509–512. doi: 10.1038/nature04756

Gerth, M., and Bleidorn, C. (2016). Comparative genomics provides a timeframe for *Wolbachia* evolution and exposes a recent biotin synthesis operon transfer. *Nat. Microbiol.* 2:16241. doi: 10.1038/nmicrobiol.2016.241

Gerth, M., Gansauge, M.-T., Weigert, A., and Bleidorn, C. (2014). Phylogenomic analyses uncover origin and spread of the *Wolbachia* pandemic. *Nat. Commun.* 5:5117. doi: 10.1038/ncomms6117

Guillade, A. C., and Folgarait, P. J. (2015). Effect of phorid fly density on the foraging of *Atta vollenweideri* leafcutter ants in the field. *Entomol. Exp. Appl.* 154, 53–61. doi: 10.1111/eea.12255

Heath, B. D., Butcher, R. D., Whitfield, W. G., and Hubbard, S. F. (1999). Horizontal transfer of *Wolbachia* between phylogenetically distant insect species by a naturally occurring mechanism. *Curr. Biol.* 9, 313–316. doi: 10.1016/S0960-9822(99)80139-0

Hiroki, M., Tagami, Y., Miura, K., and Kato, Y. (2004). Multiple infection with *Wolbachia* inducing different reproductive manipulations in the butterfly *Eurema hecabe*. *Proc. Biol. Sci.* 271, 1751–1755. doi: 10.1098/rspb.2004.2769

Hoffmann, A. A., Montgomery, B. L., Popovici, J., Iturbe-Ormaetxe, I., Johnson, P. H., Muzzi, F., et al. (2011). Successful establishment of *Wolbachia* in Aedes populations to suppress dengue transmission. *Nature* 476, 454–457. doi: 10.1038/nature10356

Ishmael, N., Dunning Hotopp, J. C., Ioannidis, P., Biber, S., Sakamoto, J., Siozios, S., et al. (2009). Extensive genomic diversity of closely related *Wolbachia* strains. *Microbiol. Read. Engl.* 155, 2211–2222. doi: 10.1099/mic.0.027581-0

Jiggins, F. M., Hurst, G. D. D., and Yang, Z. (2002). Host-symbiont conflicts: positive selection on an outer membrane protein of parasitic but not Mutualistic Rickettsiaceae. *Mol. Biol. Evol.* 19, 1341–1349. doi: 10.1093/oxfordjournals.molbev.a004195

Kautz, S., Rubin, B. E. R., and Moreau, C. S. (2013). Bacterial Infections across the Ants: frequency and prevalence of *Wolbachia*, *Spiroplasma*, and *Asaia*. *Psyche J. Entomol.* 2013:e936341. doi: 10.1155/2013/936341

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (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

Kent, B. N., and Bordenstein, S. R. (2010). Phage WO of *Wolbachia*: lambda of the endosymbiont world. *Trends Microbiol.* 18, 173–181. doi: 10.1016/j.tim.2009.12.011

Klasson, L., Westberg, J., Sapountzis, P., Näslund, K., Lutnaes, Y., Darby, A. C., et al. (2009). The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5725–5730. doi: 10.1073/pnas.0810753106

Lambardi, D., Dani, F. R., Turillazzi, S., and Boomsma, J. J. (2007). Chemical mimicry in an incipient leaf-cutting ant social parasite. *Behav. Ecol. Sociobiol.* 61, 843–851. doi: 10.1007/s00265-006-0313-y

Lapolla, J. S., Mueller, U. G., Seid, M. A., and Cover, S. P. (2002). Predation by the army ant *Neivamyrmex rugulosus* on the fungus-growing ant *Trachymyrmex arizonensis* - semantic scholar. *Insect. Soc.* 49, 251–256. doi: 10.1007/s00040-002-8310-2

Larsen, J., Fouks, B., Bos, N., d'Ettorre, P., and Nehring, V. (2014). Variation in nestmate recognition ability among polymorphic leaf-cutting ant workers. *J. Insect Physiol.* 70, 59–66. doi: 10.1016/j.jinsphys.2014.09.002

Leclercq, S., Giraud, I., and Cordaux, R. (2011). Remarkable abundance and evolution of mobile group II introns in *Wolbachia* bacterial endosymbionts. *Mol. Biol. Evol.* 28, 685–697. doi: 10.1093/molbev/msq238

Liberti, J., Sapountzis, P., Hansen, L. H., Sørensen, S. J., Adams, R. M. M., and Boomsma, J. J. (2015). Bacterial symbiont sharing in *Megalomyrmex* social parasites and their fungus-growing ant hosts. *Mol. Ecol.* 24, 3151–3169. doi: 10.1111/mec.13216

Martins, C., Souza, R. F., and Bueno, O. C. (2012). Presence and distribution of the endosymbiont *Wolbachia* among *Solenopsis* spp. (*Hymenoptera*: *Formicidae*) from Brazil and its evolutionary history. *J. Invertebr. Pathol.* 109, 287–296. doi: 10.1016/j.jip.2012.01.001

Mascarenhas, R. O., Prezotto, L. F., Perondini, A. L. P., Marino, C. L., and Selivon, D. (2016). *Wolbachia* in guilds of *Anastrepha* fruit flies (Tephritidae) and parasitoid wasps (Braconidae). *Genet. Mol. Biol.* 39, 600–610. doi: 10.1590/1678-4685-GMB-2016-0075

McMeniman, C. J., Lane, A. M., Fong, A. W. C., Voronin, D. A., Iturbe-Ormaetxe, I., Yamada, R., et al. (2008). Host adaptation of a *Wolbachia* strain after long-term serial passage in mosquito cell lines. *Appl. Environ. Microbiol.* 74, 6963–6969. doi: 10.1128/AEM.01038-08

Moreau, C. S., Bell, C. D., Vila, R., Archibald, S. B., and Pierce, N. E. (2006). Phylogeny of the ants: diversification in the age of angiosperms. *Science* 312, 101–104. doi: 10.1126/science.1124891

Mouton, L., Dedeine, F., Henri, H., Boulétreau, M., Profizi, N., and Vavre, F. (2004). Virulence, multiple infections and regulation of symbiotic population in the Wolbachia-*Asobara tabida* symbiosis. *Genetics* 168, 181–189. doi: 10.1534/genetics.104.026716

Nehring, V., Dani, F. R., Turillazzi, S., Boomsma, J. J., and d'Ettorre, P. (2015). Integration strategies of a leaf-cutting ant social parasite. *Anim. Behav.* 108, 55–65. doi: 10.1016/j.anbehav.2015.07.009

Noda, H., Miyoshi, T., Zhang, Q., Watanabe, K., Deng, K., and Hoshizaki, S. (2001). *Wolbachia* infection shared among planthoppers (Homoptera: Delphacidae) and their endoparasite (Strepsiptera: Elenchidae): a probable case of interspecies transmission. *Mol. Ecol.* 10, 2101–2106. doi: 10.1046/j.0962-1083.2001.01334.x

Peralta, L., and Martínez, P. A. (2013). Ensambles de ácaros oribátidos en hormigueros de *Acromyrmex* spp. *(Hymenoptera, Formicidae). Ecol. Austral* 23, 209–217.

Pérez-Ortega, B., Fernández-Marín, H., Loiácono, M., Galgani, P., and Wcislo, W. T. (2010). Biological notes on a fungus-growing ant, *Trachymyrmex* cf. zeteki (Hymenoptera, Formicidae, Attini) attacked by a diverse community of parasitoid wasps (Hymenoptera, Diapriidae). *Insectes Sociaux* 57, 317–322. doi: 10.1007/s00040-010-0086-1

Powell, S., and Clark, E. (2004). *Combat Between Large Derived Societies: A Subterranean Army Ant Established As a Predator of Mature Leaf-Cutting Ant Colonies*. Available at: http://repository.si.edu//handle/10088/3728 [Accessed December 7, 2017].

Ramalho, M. O., Martins, C., Silva, L. M. R., Martins, V. G., and Bueno, O. C. (2017). Intracellular symbiotic bacteria of camponotus textor, Forel (Hymenoptera, Formicidae). *Curr. Microbiol.* 74, 589–597. doi: 10.1007/s00284-017-1201-6

Rambaut, A. (2016). *FigTree Version 1.4.3 [Computer Program]*. Available at: http://tree.bio.ed.ac.uk/software/figtree/

Rasgon, J. L., Gamston, C. E., and Ren, X. (2006). Survival of *Wolbachia* pipientis in cell-free medium. *Appl. Environ. Microbiol.* 72, 6934–6937. doi: 10.1128/AEM.01673-06

Raychoudhury, R., Baldo, L., Oliveira, D. C., and Werren, J. H. (2009). Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex. *Evol. Int. J. Org. Evol.* 63, 165–183. doi: 10.1111/j.1558-5646.2008.00533.x

Reuter, M., Pedersen, J. S., and Keller, L. (2005). Loss of *Wolbachia* infection during colonisation in the invasive argentine ant linepithema humile. *Heredity* 94, 364–369. doi: 10.1038/sj.hdy.6800601

Rey, O., Estoup, A., Facon, B., Loiseau, A., Aebi, A., Duron, O., et al. (2013). Distribution of endosymbiotic reproductive manipulators reflects invasion process and not reproductive system polymorphism in the little fire ant *Wasmannia auropunctata*. *PloS One* 8:e58467. doi: 10.1371/journal.pone.0058467

Russell, J. A., Goldman-Huertas, B., Moreau, C. S., Baldo, L., Stahlhut, J. K., Werren, J. H., et al. (2009). Specialization and geographic isolation among *Wolbachia* symbionts from ants and lycaenid butterflies. *Evol. Int. J. Org. Evol.* 63, 624–640. doi: 10.1111/j.1558-5646.2008.00579.x

Sapountzis, P., Zhukova, M., Hansen, L. H., Sørensen, S. J., Schiøtt, M., and Boomsma, J. J. (2015). *Acromyrmex* leaf-cutting ants have simple gut microbiota with nitrogen-fixing potential. *Appl. Environ. Microbiol.* 81, 5527–5537. doi: 10.1128/AEM.00961-15

Saridaki, A., and Bourtzis, K. (2010). *Wolbachia*: more than just a bug in insects genitals. *Curr. Opin. Microbiol.* 13, 67–72. doi: 10.1016/j.mib.2009.11.005

Schultz, T. R., and Brady, S. G. (2008). Major evolutionary transitions in ant agriculture. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5435–5440. doi: 10.1073/pnas.0711024105

Siozios, S., Ioannidis, P., Klasson, L., Andersson, S. G. E., Braig, H. R., and Bourtzis, K. (2013). The diversity and evolution of *Wolbachia* ankyrin repeat domain genes. *PLoS One* 8:e55390. doi: 10.1371/journal.pone.0055390

Stahlhut, J. K., Desjardins, C. A., Clark, M. E., Baldo, L., Russell, J. A., Werren, J. H., et al. (2010). The mushroom habitat as an ecological arena for global exchange of *Wolbachia*. *Mol. Ecol.* 19, 1940–1952. doi: 10.1111/j.1365-294X.2010.04572.x

Stouthamer, R., Breeuwer, J. A., and Hurst, G. D. (1999). *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. *Annu. Rev. Microbiol.* 53, 71–102. doi: 10.1146/annurev.micro.53.1.71

Stürup, M., Nash, D. R., Hughes, W. O. H., and Boomsma, J. J. (2014). Sperm mixing in the polyandrous leaf-cutting ant *Acromyrmex echinatior*. *Ecol. Evol.* 4, 3571–3582. doi: 10.1002/ece3.1176

Tsutsui, N. D., Kauppinen, S. N., Oyafuso, A. F., and Grosberg, R. K. (2003). The distribution and evolutionary history of *Wolbachia* infection in native and introduced populations of the invasive argentine ant (*Linepithema humile*). *Mol. Ecol.* 12, 3057–3068. doi: 10.1046/j.1365-294X.2003.01979.x

Van Borm, S., Wenseleers, T., Billen, J., and Boomsma, J. J. (2003). Cloning and sequencing of wsp encoding gene fragments reveals a diversity of co-infecting *Wolbachia* strains in *Acromyrmex* leafcutter ants. *Mol. Phylogenet. Evol.* 26, 102–109. doi: 10.1016/S1055-7903(02)00298-1

Vavre, F., Fleury, F., Lepetit, D., Fouillet, P., and Boulétreau, M. (1999). Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. *Mol. Biol. Evol.* 16, 1711–1723. doi: 10.1093/oxfordjournals.molbev.a026084

Veneti, Z., Clark, M. E., Karr, T. L., Savakis, C., and Bourtzis, K. (2004). Heads or tails: host-parasite interactions in the Drosophila-*Wolbachia* system. *Appl. Environ. Microbiol.* 70, 5366–5372. doi: 10.1128/AEM.70.9.5366-5372.2004

Veneti, Z., Zabalou, S., Papafotiou, G., Paraskevopoulos, C., Pattas, S., Livadaras, I., et al. (2012). Loss of reproductive parasitism following transfer of male-killing *Wolbachia* to *Drosophila melanogaster* and *Drosophila simulans*. *Heredity* 109, 306–312. doi: 10.1038/hdy.2012.43

Viljakainen, L., Reuter, M., and Pamilo, P. (2008). *Wolbachia* transmission dynamics in formica wood ants. *BMC Evol. Biol.* 8:55. doi: 10.1186/1471-2148-8-55

Vos, M., and Didelot, X. (2009). A comparison of homologous recombination rates in bacteria and archaea. *ISME J.* 3, 199–208. doi: 10.1038/ismej.2008.93

Weinert, L. A., Araujo-Jnr, E. V., Ahmed, M. Z., and Welch, J. J. (2015). The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc. Biol. Sci.* 282:20150249. doi: 10.1098/rspb.2015.0249

Werren, J. H., Baldo, L., and Clark, M. E. (2008). *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* 6, 741–751. doi: 10.1038/nrmicro1969

White, P. M., Pietri, J. E., Debec, A., Russell, S., Patel, B., and Sullivan, W. (2017). Mechanisms of horizontal cell-to-cell transfer of *Wolbachia* spp. in *Drosophila melanogaster*. *Appl. Environ. Microbiol.* 83, e03425-e16. doi: 10.1128/AEM.03425-16

Wu, M., Sun, L. V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J. C., et al. (2004). Phylogenomics of the reproductive parasite *Wolbachia* pipientis wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol.* 2:E69. doi: 10.1371/journal.pbio.0020069

Yang, X.-H., Zhu, D.-H., Liu, Z., Zhao, L., and Su, C.-Y. (2013). High levels of multiple infections, recombination and horizontal transmission of *Wolbachia* in the *Andricus mukaigawae* (Hymenoptera; Cynipidae) communities. *PloS One* 8:e78970. doi: 10.1371/journal.pone.0078970

Zabalou, S., Apostolaki, A., Pattas, S., Veneti, Z., Paraskevopoulos, C., Livadaras, I., et al. (2008). Multiple rescue factors within a *Wolbachia* strain. *Genetics* 178, 2145–2160. doi: 10.1534/genetics.107.086488

Zabalou, S., Riegler, M., Theodorakopoulou, M., Stauffer, C., Savakis, C., and Bourtzis, K. (2004). *Wolbachia*-induced cytoplasmic incompatibility as a means for insect pest population control. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15042–15045. doi: 10.1073/pnas.0403853101

Zhao, D.-X., Chen, D.-S., Ge, C., Gotoh, T., and Hong, X.-Y. (2013). Multiple infections with *Cardinium* and two strains of *Wolbachia* in the spider mite *Tetranychus phaselus* Ehara: revealing new forces driving the spread of *Wolbachia*. *PLoS One* 8:e54964. doi: 10.1371/journal.pone.0054964

Zhou, W., Rousset, F., and O'Neil, S. (1998). Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. Biol. Sci.* 265, 509–515. doi: 10.1098/rspb.1998.0324

Zhukova, M., Sapountzis, P., Schiøtt, M., and Boomsma, J. J. (2017). Diversity and transmission of gut bacteria in *Atta* and *Acromyrmex* leaf-cutting ants during development. *Front. Microbiol.* 8:1942. doi: 10.3389/fmicb.2017.01942

Zwickl, D. J. (2006). *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion*. Ph.D. thesis, University of Texas, Austin, TX.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplemental Results

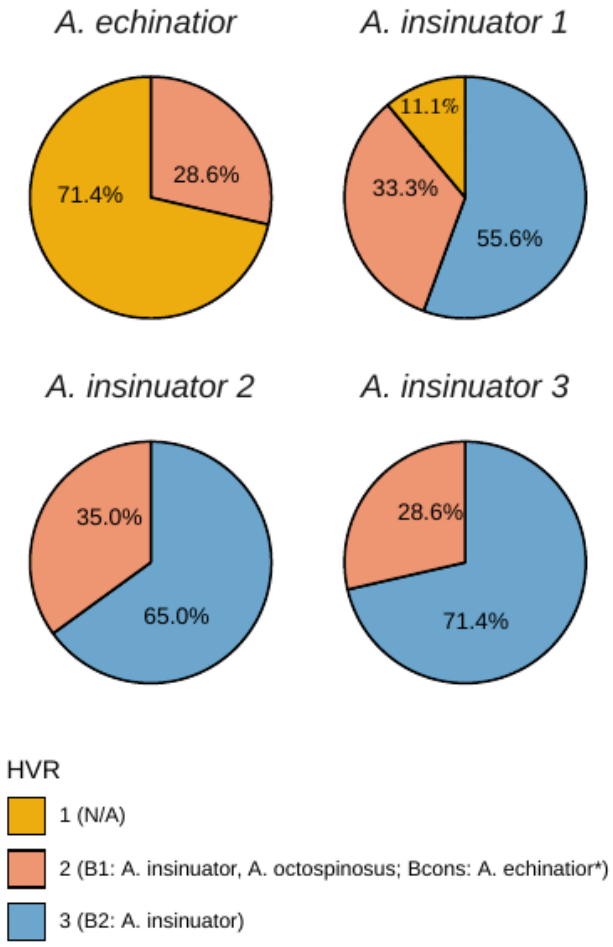


FIGURE S1. Relative proportion of HVR types 1–3 across *Acromyrmex* queen eggs. Pie chart showing the relative proportion of eggs sequenced from queens in this study with each HVR type, as described in Supplementary Table S1. The legend lists the HVR reference number used in this study followed by parentheses including the name of identical strains and the host species harboring those strains as described by Van Borm et al. (2003). Although closely related, the asterisk indicates that the strain Bcons was not a perfect match to HVR-2 from our study, unlike strain B1 (present in *A. insinuator* and *A. octospinosus*) which was identical to HVR-2.

Queen	Most Closely Related Hosts	Accession Numbers	Geographic Origin	Average % Identity	# of Sequences	HVR Type	HVR Reference # (This Study)	% of Sequences with HVR Type
A. echinatior	*Megalomyrmex wallacei*	LC027874	Costa Rica	99.0%	15	37-38-69-37	1	71.43%
	Sericomyrmex sp. & Neivamyrmex nigrescens	LC027866 & KC137187	Panama, Mexico	99.5%	6	21-21-25-21	2	28.57%
	Solenopsis invicta & Linepithema humile	HM747152 & AY446990	Brazil	99.2%	2	21-38-69-37 / 21-21-25-37	N/A	NOT Included
A. insinuator 1	*Megalomyrmex wallacei*	LC027874	Costa Rica	98.5%	2	37-38-69-37	1	11%
	Sericomyrmex sp. & Neivamyrmex nigrescens	LC027866 & KC137187	Panama, Mexico	99.7%	4	21-21-25-21	2	33.33%
	Wasmannia auropunctata & Solenopsis invicta	JX499066 & DQ842483	French Guiana, Brazil	99.5%	2			
	Acromyrmex insinuator & Solenopsis invicta	AF472560 & HM747159	Panama, Brazil	99.7%	10	21-40-42-39	3	55.56%

TABLE S1. Top BLAST matches and corresponding HVR type for *A. echinatior* and *A. insinuator wsp* sequences generated in this study. Summary of HVR typing results from Supplementary Table S2 and additional *wsp* characterizations from the *A. echinatior* host queen (HQ) and the three parasitic *A. insinuator* (PQ) queens. From left to right: queen ant used to collect egg DNA samples for this study, host species and accession numbers for the top BLAST hits matching respective *Wolbachia* sequences, the geographic origin of NCBI samples of other ants, the average % match identity in BLAST, the number of sequences that returned these hits, HVR (hypervariable region) typing according to Baldo et al. (2006b) used in this study, and the percentage of sequences from each queen with respective HVR types. HVR typing is based on the four hypervariable regions of the *wsp* gene which are comprised of relatively conserved amino acid motifs used to identify recombination points, allowing discrimination between closely related *Wolbachia* strains, comparable to the use of antigens for serotyping pathogenic bacteria (Baldo et al., 2006b). The four HVRs occupy consecutive conserved regions of the *wsp* gene and the combination of all four HVR haplotypes make up its WSP profile. All *A. echinatior* and *A. insinuator* sequences were classified by their four HVRs using the PubMLST database[2] (best match to existing sequences in the database), which revealed three distinct *wsp* genotypes, here referred to as HVR 1-3. Color coding for the HVR types matches those shown in Figure 1, except for one case where sequences were chimeras (type 21-38-69-37/21-21-25-37) of HVR-1 and HVR-2. As identified in other strains, recombination was localized in the HVRs, which suggests the two *A. echinatior* sequences are true chimeras rather than sequencing errors (Andersen et al., 2012).

Queen	Accession #	Top 3 BLAST Hits	Accession #	% Identical	pubMLST database	*wsp* typing (nucleotide query)	HVR1 (amino query)	HVR2 (amino query)	HVR3 (amino query)	HVR4 (amino query)
A. echinatior		*Megalomyrmex wallacei*	LC027874	99.4	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99.4						
	MG547498	*Megalomyrmex wallacei*	LC027872	99.4	**Differences:**	317C →T			30K →R	
		Megalomyrmex wallacei	LC027874	99.4	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99.4						
	MG547478	*Megalomyrmex wallacei*	LC027872	99.4	**Differences:**	317C →T			30K →R	
		Megalomyrmex wallacei	LC027874	99	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99		31G→A, 143G→A,				
	MG547481	*Megalomyrmex wallacei*	LC027872	99	**Differences:**	114T→G, 317C→T	11G→S	15S→K	30K →R	
		Megalomyrmex wallacei	LC027874	99.4	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99.4						
	MG547489	*Megalomyrmex wallacei*	LC027872	99.4	**Differences:**	7G→A, 317C→T	3V→I		30K→R	
		Megalomyrmex wallacei	LC027874	99.2	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99.2						
	MG547499	*Megalomyrmex wallacei*	LC027872	99.2	**Differences:**	317C→T			30K →R	
		Megalomyrmex wallacei	LC027874	99.1	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99.1						
	MG547483	*Megalomyrmex wallacei*	LC027872	99.1	**Differences:**	317C→T			30K →R	
		Megalomyrmex wallacei	LC027874	98.9	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	98.9						
	MG547497	*Megalomyrmex wallacei*	LC027872	98.9	**Differences:**	317C→T,378T→C			30K→R	
		Megalomyrmex wallacei	LC027874	99	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99						
	MG547490	*Megalomyrmex wallacei*	LC027872	99	**Differences:**	91T→C,317C→T	31Y→H		30K→R	
		Megalomyrmex wallacei	LC027874	99.1	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99.1						
	MG547492	*Megalomyrmex wallacei*	LC027872	99.1	**Differences:**	317C→T			30K→R	
		Megalomyrmex wallacei	LC027874	98.8	**Closest match:**	58	30*	38	69	21*
		Megalomyrmex wallacei	LC027873	98.8		42T→G, 317C→T,				
	MG547487	*Megalomyrmex wallacei*	LC027872	98.8	**Differences:**	452T→C, 467C→A			30K→R	
		Megalomyrmex wallacei	LC027874	99.2	**Closest match:**	58	37	38	69	21*
		Megalomyrmex wallacei	LC027873	99.2						
	MG547494	*Megalomyrmex wallacei*	LC027872	99.2	**Differences:**	317C→T, 467C→A			30K→R	19A-V
		Megalomyrmex wallacei	LC027874	98.3	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	98.3				2D→G,		
	MG547486	*Megalomyrmex wallacei*	LC027872	98.3	**Differences:**		29F→S	12S→L	30K→R	
		Megalomyrmex wallacei	LC027874	98.7	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	98.7		85T→C,135A→C,				
	MG547480	*Megalomyrmex wallacei*	LC027872	98.7	**Differences:**	317C→T	29F →L		30K →R	
		Megalomyrmex wallacei	LC027874	98.9	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	98.9						
	MG547491	*Megalomyrmex wallacei*	LC027872	98.9	**Differences:**	135A→C,317C→T			30K→R	
		Megalomyrmex wallacei	LC027874	99.2	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	99.2						
	MG547495	*Megalomyrmex wallacei*	LC027872	99.2	**Differences:**	317C→T			30K→R	
		Solenopsis invicta	HM747152	98.9	**Closest match:**	58	21	38	69	37
		InvB (Linepithema humile)	AY446990	98.7		7G→A,				
	MG547493	*InvB (Linepithema humile)*	AY446989	98.7	**Differences:**	20C→A,28A→G,42T→G, 105C→T, 317C→T			30K→R	
		Sericomyrmex sp.	LC027866	100	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	100						
	MG547500	*Neivamyrmex nigrescens*	KC137187	100	**Differences:**					
		Sericomyrmex sp.	LC027866	99.8	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99.8						
	MG547482	*Neivamyrmex nigrescens*	KC137187	99.8	**Differences:**	143G→A,144T→G		15S→K		
		Sericomyrmex sp.	LC027866	99	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99						
	MG547479	*Neivamyrmex nigrescens*	KC137187	99	**Differences:**			12S→L		
		Sericomyrmex sp.	LC027866	99.8	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99.8						
	MG547485	*Neivamyrmex nigrescens*	KC137187	99.8	**Differences:**					
		Sericomyrmex sp.	LC027866	99.2	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99.2		41A→G, 270T→C,				

	MG547488	*Neivamyrmex nigrescens*	KC137187	99.2	**Differences:**	358A→G	14K→R		39K→E	
		Sericomyrmex sp.	LC027866	99.4	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99.4						
	MG547496	*Neivamyrmex nigrescens*	KC137187	99.4	**Differences:**	16A→G,113T→C	6I→V	5V→A		
		Solenopsis invicta	HM747152	99.8	**Closest match:**	28	21	21	25	37
		InvB (Linepithema humile)	AY446990	99.6						
	MG547484	*InvB (Linepithema humile)*	AY446989	99.6	**Differences:**	452C→T,467A→C				
A. insinuator 1		*Acromyrmex insinuator 2*	AF472560	99.6	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747159	99.6						
	MG547503	*Solenopsis invicta*	HM747156	99.6	**Differences:**	134C →T		12S→L		
		Acromyrmex insinuator 2	AF472560	99.8	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747159	99.8						
	MG547507	*Solenopsis invicta*	HM747156	99.8	**Differences:**					
		Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747159	99.6						
	MG547514	*Solenopsis invicta*	HM747156	99.6	**Differences:**	131A→T		11Y→F		
		Solenopsis invicta	HM747159	99.8	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747156	99.8						
	MG547506	*Acromyrmex insinuator 2*	AF472560	99.6	**Differences:**					
		Solenopsis invicta	HM747159	99.8	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747156	99.8						
	MG547508	*Acromyrmex insinuator 2*	AF472560	99.6	**Differences:**					
		Solenopsis invicta	HM747159	100	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747156	100						
	MG547504	*Acromyrmex insinuator 2*	AF472560	100	**Differences:**					
		Solenopsis invicta	HM747159	100	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747156	100						
	MG547518	*Acromyrmex insinuator 2*	AF472560	100	**Differences:**					
		Acromyrmex insinuator 2	AF472560	99.4	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747159	99.3						
	MG547505	*Solenopsis invicta*	HM747156	99.3	**Differences:**	131A→T		11Y→F		
		Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747159	99.6						
	MG547513	*Solenopsis invicta*	HM747156	99.6	**Differences:**	127C→G		10L→V		
		Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747159	99.6						
	MG547515	*Solenopsis invicta*	HM747156	99.6	**Differences:**	413T→C				8F→S
		Sericomyrmex sp.	LC027866	99.6	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99.6						
	MG547517	*Neivamyrmex nigrescens*	KC137187	99.6	**Differences:**	261A→G				
		Sericomyrmex sp.	LC027866	100	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	100						
	MG547511	*Neivamyrmex nigrescens*	KC137187	100	**Differences:**					
		Sericomyrmex sp.	LC027866	99.8	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99.8						
	MG547501	*Neivamyrmex nigrescens*	KC137187	99.8	**Differences:**					
		Wasmannia auropunctata	JX499066	99.8	**Closest match:**	28	21	21	25	21
		Solenopsis invicta	DQ842483	99.8						
	MG547509	*Solenopsis invicta*	HM747152	99.8	**Differences:**	261A→G				
		Wasmannia auropunctata	JX499066	99.2	**Closest match:**	28	21	21	25	21
		Solenopsis invicta	DQ842483	99.2						
	MG547512	*Solenopsis invicta*	HM747152	99.1	**Differences:**	133T→C, 135A→C, 366T→C		12S→P		
		Sericomyrmex sp.	LC027866	99.2	**Closest match:**	28	21	21	25	21
		Sericomyrmex sp.	LC027865	99.2						
	MG547502	*Neivamyrmex nigrescens*	KC137187	99.2	**Differences:**	135A→C, 137A→C, 322A→G		13Q→P	27K→E	
		Megalomyrmex wallacei	LC027874	98.5	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	98.5						
	MG547510	*Megalomyrmex wallacei*	LC027872	98.5	**Differences:**	28A→G, 90T→C, 166A→G, 204T→C, 317C→T	10T→A	23T→A	30K→R	
		Megalomyrmex wallacei	LC027874	98.5	**Closest match:**	58	37	38	69	37
		Megalomyrmex wallacei	LC027873	98.5						
	MG547516	*Megalomyrmex wallacei*	LC027872	98.5	**Differences:**	28A→G, 135A→C, 166A→G, 204T→C, 317C→T	10T→A	23T→A	30K →R	
A. insinuator 2		*Acromyrmex insinuator 2*	AF472560	99.8	**Closest match:**	59	21	40	42	39
		Solenopsis invicta	HM747159	99.8						
	MG547529	*Solenopsis invicta*	HM747156	99.8	**Differences:**					

	Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.6						
MG547531	*Solenopsis invicta*	HM747156	99.6	**Differences:**	214T→C		39Y→H		
	Acromyrmex insinuator 2	AF472560	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.8						
MG547535	*Solenopsis invicta*	HM747156	99.8	**Differences:**	107T→A		3I→N		
	Solenopsis invicta	HM747159	100	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	100						
MG547538	*Acromyrmex insinuator 2*	AF472560	100	**Differences:**					
	Acromyrmex insinuator 2	AF472560	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.8						
MG547519	*Solenopsis invicta*	HM747156	99.8	**Differences:**					
	Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.6						
MG547530	*Solenopsis invicta*	HM747156	99.6	**Differences:**	127C→G		10L→V		
	Solenopsis invicta	HM747159	99.3	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.3						
MG547533	*Acromyrmex insinuator 2*	AF472560	99.2	**Differences:**	78T→C, 127C→G		10L→V		
	Solenopsis invicta	HM747159	100	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	100						
MG547537	*Acromyrmex insinuator 2*	AF472560	100	**Differences:**					
	Solenopsis invicta	HM747159	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.8						
MG547524	*Acromyrmex insinuator 2*	AF472560	99.8	**Differences:**					
	Solenopsis invicta	HM747159	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.6						
MG547532	*Solenopsis invicta*	JQ425789	99.3	**Differences:**	348A→G, 444A→G				
	Solenopsis invicta	HM747159	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.8						
MG547526	*Acromyrmex insinuator 2*	AF472560	99.6	**Differences:**					
	Solenopsis invicta	HM747159	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.6						
MG547527	*Acromyrmex insinuator 2*	AF472560	99.4	**Differences:**	69A→G	23I→M			
	Acromyrmex insinuator 2	AF472560	99.2	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.1		127C→G, 135A→G,				
MG547523	*Solenopsis invicta*	HM747156	99.1	**Differences:**	320A→G		10L→V	29K→R	
	Sericomyrmex sp.	LC027866	99.8	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.8						
MG547520	*Neivamyrmex nigrescens*	KC137187	99.8	**Differences:**					
	Sericomyrmex sp.	LC027866	99.6	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.6						
MG547534	*Neivamyrmex nigrescens*	KC137187	99.6	**Differences:**	135A→C				
	Sericomyrmex sp.	LC027866	99.8	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.8						
MG547521	*Neivamyrmex nigrescens*	KC137187	99.8	**Differences:**					
	Sericomyrmex sp.	LC027866	99.6	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.6						
MG547536	*Neivamyrmex nigrescens*	KC137187	99.6	**Differences:**	69A→G	23I→M			
	Sericomyrmex sp.	LC027866	99.6	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.6						
MG547525	*Neivamyrmex nigrescens*	KC137187	99.6	**Differences:**	60A→G				
	Wasmannia auropunctata	JX499066	99.4	**Closest match:**	28	21	21	25	21
	Solenopsis invicta	DQ842483	99.4						
MG547528	*Solenopsis invicta*	HM747152	99.3	**Differences:**	135A→C, 137A→C		13Q→P		
	Wasmannia auropunctata	JX499066	99.6	**Closest match:**	28	21	21	25	21
	Solenopsis invicta	DQ842483	99.6						
MG547522	*Solenopsis invicta*	HM747152	99.6	**Differences:**	243A→G				
A. insinuator 3	*Acromyrmex insinuator 2*	AF472560	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.6						
MG547539	*Solenopsis invicta*	HM747156	99.6	**Differences:**					
	Solenopsis invicta	HM747159	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.8						
MG547552	*Acromyrmex insinuator 2*	AF472560	99.6	**Differences:**					
	Solenopsis invicta	HM747159	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.8						
MG547550	*Acromyrmex insinuator 2*	AF472560	99.6	**Differences:**					
	Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39

	Solenopsis invicta	HM747159	99.6						
MG547555	*Solenopsis invicta*	HM747156	99.6	**Differences:**	401G→A				4G→D
	Acromyrmex insinuator 2	AF472560	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.8						
MG547541	*Solenopsis invicta*	HM747156	99.8	**Differences:**					
	Acromyrmex insinuator 2	AF472560	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.8						
MG547556	*Solenopsis invicta*	HM747156	99.8	**Differences:**					
	Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.6						
MG547548	*Solenopsis invicta*	HM747156	99.6	**Differences:**	243T→C				
	Solenopsis invicta	HM747159	100	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	100						
MG547559	*Acromyrmex insinuator 2*	AF472560	100	**Differences:**					
	Solenopsis invicta	HM747159	100	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	100						
MG547549	*Acromyrmex insinuator 2*	AF472560	100	**Differences:**					
	Solenopsis invicta	HM747159	99.8	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.8						
MG547546	*Acromyrmex insinuator 2*	AF472560	99.6	**Differences:**					
	Solenopsis invicta	HM747159	99.3	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747156	99.3						
MG547547	*Acromyrmex insinuator 2*	AF472560	99.2	**Differences:**	36A→G, 270T→C				
	Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.6						
MG547545	*Solenopsis invicta*	HM747156	99.6	**Differences:**	186A→G				
	Acromyrmex insinuator 2	AF472560	99.6	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.6						
MG547540	*Solenopsis invicta*	HM747156	99.6	**Differences:**	30A→T				
	Acromyrmex insinuator 2	AF472560	99.4	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.3						
MG547551	*Solenopsis invicta*	HM747156	99.3	**Differences:**					
	Acromyrmex insinuator 2	AF472560	99.2	**Closest match:**	59	21	40	42	39
	Solenopsis invicta	HM747159	99.1		15T→C, 127C→G,				
MG547553	*Solenopsis invicta*	HM747156	99.1	**Differences:**	241C→T		10L→V	3P→S	
	Sericomyrmex sp.	LC027866	99.6	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.6						
MG547543	*Neivamyrmex nigrescens*	KC137187	99.6	**Differences:**	135A→C				
	Sericomyrmex sp.	LC027866	99.8	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.8						
MG547554	*Neivamyrmex nigrescens*	KC137187	99.8	**Differences:**					
	Sericomyrmex sp.	LC027866	100	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	100						
MG547557	*Neivamyrmex nigrescens*	KC137187	100	**Differences:**					
	Sericomyrmex sp.	LC027866	99.8	**Closest match:**	28	21	21	25	21
	Sericomyrmex sp.	LC027865	99.8						
MG547542	*Neivamyrmex nigrescens*	KC137187	99.8	**Differences:**					
	Wasmannia auropunctata	JX499066	99.6	**Closest match:**	28	21	21	25	21
	Solenopsis invicta	DQ842483	99.6						
MG547544	*Solenopsis invicta*	HM747152	99.6	**Differences:**	441T→441C				
	Sericomyrmex sp.	LC027866	99.6	**Closest match:**	28	21	21	261*	21
	Sericomyrmex sp.	LC027865	99.6					29Q-L,	
MG547558	*Neivamyrmex nigrescens*	KC137187	99.6	**Differences:**	329A-T, 377A-G			30K-R	

TABLE S2. Individual sequence information and alignment to HVR reference sequences. From left to right: colony of origin, accession numbers for sequences from this study, host species and accession numbers for the top BLAST hits matching respective *Wolbachia* sequences, % identity to BLAST match, closest match to sequences in pubMLST database (Baldo et al., 2006b) for *wsp*(nucleotide query), HVR1-4 (amino acid query) with sequence differences listed below. Colors coordinate with HVR types 1–3 used in Figure 1. Asterisks indicate sequences that differed in HVR reference number relative to other strains in the same HVR type (1–3) classified in this study. The three sequences with asterisks only exhibited a 1–3 bp difference from other sequences in their respective HVR characterization (types 1–3) and were thus included in analyses.

CHAPTER 3:

DEFINING *WOLBACHIA* SYMBIOSIS IN LEAF-CUTTER ANTS BY COMPARATIVE GENOMICS AND DYNAMIC STAGE-SPECIFIC GENE EXPRESSION

Introduction

Microbial symbionts have a profound effect on the ecology and evolution of their hosts with interactions spanning the continuum between mutualism, commensalism, and parasitism (Bordenstein et al., 2009; Engelstädter et al., 2009; Werren et al., 2008). Such variation in symbiotic relationships can be observed in interactions between eukaryotic hosts and intracellular α-proteobacteria from the genus *Wolbachia*, which include obligate mutualism in nematodes (Comandatore et al., 2015) and reproductive parasitism in many arthropods. In addition to altering host biology, these symbiotic relationships have potential applications in pest control (Bourtzis, 2008) and reduced disease transmission (e.g. introduction of CI-inducing *Wolbachia* in populations of *A. aegypti* to reduce vectoring of the dengue virus; Walker et al., 2011). *Wolbachia* is the most widespread endosymbiotic bacterium on Earth (Werren et al., 2008) yet despite its abundance and promising applications, much of *Wolbachia*'s evolutionary history remains to be described (Gerth & Bleidorn 2016). In particular, although the phenotypic effects of *Wolbachia* infection has been clarified across several groups of solitary arthropods (LePage et al., 2017), much less progress has been made in understanding how *Wolbachia* influences the development, reproduction, and defensive biology of eusocial insects, which consist of over 12,000 species including several invasive pests (Andersen et al., 2012; Russell, 2012).

Efforts to pinpoint *Wolbachia* genes that play essential roles in mediating host interaction and maintaining their infection status have been impeded by the unculturable nature of this bacterium in a free-living environment. As a result, comparative genomic methods and genome-wide expression profiling are used to identify candidate microbial genes involved in host-symbiont interactions. For example, despite significant genome reductions, arthropod *Wolbachia*

strains have a surprisingly high number of ANK genes accounting for up to 43 genes or 3% of the total genes in strains *w*Mel, *w*Pip, and *w*Ri (Klasson et al., 2008; Klasson et al., 2009; Sinha, 2019; Wu et al., 2004). This contrasts with the genomes of closely related α-Proteobacteria like *Rickettsia* and *Ehrlichia*, which typically only harbor one to three ANK genes (Andersson et al., 1998). Interestingly, these genes have been demonstrated to suppress host innate immunity, regulate transcription, inhibit host cell apoptosis, and reduce reactive oxygen species (Liu et al., 2012; Rikihisa and Lin, 2010). Similarly, genomic comparisons with other *Wolbachia* strains have resulted in the notion that the nature of *Wolbachia*-host relationships largely depends on how *Wolbachia* and host cells interact across iron-dependent pathways (Brownlie et al., 2009; Gill et al., 2014; Kremer et al., 2009). For instance, *Wolbachia*-infected *Drosophila melanogaster* was shown to have increased fecundity relative to uninfected females when reared on either iron-enriched or -reduced diets (Brownlie et al., 2009). Additionally, studies of the wasp host, *Asobara tabida*, reveal *Wolbachia* infection may regulate iron homeostasis to limit oxidative stress and cell death thereby ensuring its own persistence in the cell (Kremer et al., 2009).

Here, we present the draft genomes of two *Wolbachia* strains isolated from fungus-growing ants (Hymenoptera: Formicidae), a common host family that until now has been excluded from phylogenetic analyses. We also report the dynamics of gene expression across different life cycle stages of the leaf-cutter ant host, *Acromyrmex echinatior*, in an effort to identify microbial genes potentially involved in host-symbiont interactions. Our results provide insight into the mechanisms of host interaction across developmental stages and provide a foundation for future research exploring the functional basis of ant-associated *Wolbachia* symbiosis.

Materials and Methods

Rearing and handling of ant colonies

Apterostigma dentigerum and *Acromyrmex echinatior* colonies were collected in Gamboa, Panama and maintained in a rearing room at ca. 25˚C and 70% relative humidity under a 12:12 h photoperiod at the Centre for Social Evolution (University of Copenhagen, Denmark). Lab colonies used in this study had been kept in laboratory conditions for at least 2 years on a diet of privet leaves and rice.

DNA extraction, bacterial isolation and multiple displacement amplification (MDA)

Sequencing libraries were created by extracting DNA from *A. echinatior* eggs and *A. dentigerum* testes and fat bodies. For *A. echinatior*, a single queen from colony Ae280 was isolated in a petri dish with a small piece of fungus (cleared of all previously laid eggs to ensure they were not produced by workers) for 48 hours, after which approximately 50 eggs were collected. *A. echinatior* is infected with two *Wolbachia* strains, however in eggs, only the wSinvictaB strain can be detected (Andersen et al., 2012). As such, eggs were used for sequencing to isolate this established dominant strain and the existence of only one strain was verified after genome sequencing. In contrast, *A. dentigerum* has a single *Wolbachia* infection, which was isolated by dissecting the fat body tissues of 25 workers and testes from 30 males from colony RMMA150520-03. Fat bodies were selected because they have been identified as having high titers of *Wolbachia* (Frost et al., 2014).

Prior to the experiments, all *A. echinatior* eggs were placed in a small sieve and surface sterilized by submerging in 70% ethanol for 1 minute, rinsing twice in autoclaved Milli Q water,

submerging in 50% bleach for 2 minutes, and again rinsing twice more in autoclaved Milli Q water. Similarly, *A. dentigerum* workers and males were anesthetized and surfaced sterilized following the same protocol. These ants were then dissected under a stereomicroscope and the fat body cells and testes were removed. *A. echinatior* eggs and *A. dentigerum* fat bodies and testes were treated as separate samples following the same bacterial isolation protocol.

For the bacterial isolations we used a previously described protocol with a few modifications (Iturbe-Ormaetxe et al., 2011). After surface sterilization and dissection, samples were immediately transferred under a laminar flow hood into a sterile 15mL glass homogenizer (Wheaton) on ice along with 1000µL of cold SPG buffer (218 mM sucrose, 3.8 mM KH_2PO_4, 7.2 mM K_2HPO_4, 4.9 mM l-glutamate, pH 7.2). Using a glass pestle, we disrupted the tissue on ice and immediately transferred tissue into a new 1.5ml Eppendorf tube. Both samples were centrifuged at (4˚C) for 15 minutes at 3,200 rfc. The supernatant was transferred to a new 1.5 mL microcentrifuge tube and centrifuged again with the same settings. The supernatant was subsequently filtered through a 5 µm (Acrodisc) and then a 2.7 µm (Whatman) syringe filter and transferred to a new 1.5 mL microcentrifuge tube and centrifuged for 20 minutes at 18,000 rfc and 4˚C. The supernatant was discarded and the pellets were re-eluted in 5ul of SPG buffer. Approximately 1ul was used for Multiple Displacement Amplification (MDA) to amplify whole genomic DNA using the Qiagen REPLI-g Midi Kit following the manufacturer's instructions.

The template and a dilution of the amplified DNA were amplified using the universal 16S rRNA primers 515F and 806R (Caporaso et al., 2011) as previously described (Sapountzis et al., 2015), purified using an Invitek PCR purification kit (PCRapace) and sent to MWG (Germany) for Sanger sequencing. As a precaution for environmental contamination, when we performed the MDA, we also used a blank reaction with sterile MilliQ water as template which produced no

16S rRNA PCR-product. After confirming that the 16S rRNA amplified was coming from *Wolbachia* and the chromatograms had no apparent signs of other bacterial 16S rDNA sequences, the MDA-amplified DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's instructions. For both samples, the extracted DNA was quantified using a spectrophotometer (Nanodrop) and subsequently was sent to seqIT (Germany) where libraries were generated using the Nextera XT DNA Library Preparation Kit (Illumina, USA) from 100-200 ng of DNA. Miseq sequencing was performed at 2 × 250-bp read length and generated approximately 3.000.000 reads per sample. For the PacBio sequencing approximately 5 ug of purified *Wolbachia* DNA was used as input for SMRTbell library construction using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA). Mean library size was approximately 12 kbp. Each library was sequenced on 1 SMRT cell using the PacBio RSII platform (Pacific Biosciences, Menlo Park, CA, USA) at the University of Washington PacBio Sequencing Services (Seattle, WA, USA). Movies were collected using the P6-C4 chemistry with a movie length of 240 minutes.

Assembly, annotation and quality control

Quality-control of the reads were performed with FastQC (Andrews, 2010). For the Illumina reads the Nextera adaptors were removed from the fastq files using Trim Galore (Babraham Institute), and the quality of reads was checked again with FastQC. We used the SPAdes Genome Assembler (version 3.5.0) to generate a de novo hybrid (illumina and pacBio reads) assembly using the "--careful" option which reduces the number of mismatches and short indels and runs MismatchCorrector with kmer sizes of 21, 33, 55 and 77. Following the assembly we used the Burrows-Wheeler Aligner (BWA) to map the reads to the assembled contigs. The

SAM file generated was further analyzed using SAMTOOLS and converted to a BAM file which was viewed using the Bamviewer v1.2.11 (Carver et al., 2010; Carver et al., 2012). Errors in the assembled contigs were further checked using the Reapr v1.0.18 software (Hunt et al., 2013). Assembled contigs that had less than 2x coverage or smaller than 250bp were removed. We performed a blastx search in all 6 codon frames using the above contigs as queries against the Uniref50 database and examined the best blast hits (>1e-10, >30%). Contigs that gave hits to ant genomes were removed.

For each assembly, genes for each contig were predicted with MAKER using the prokaryotic default settings with a few modifications. In short, Genemark S was trained using the assembled contigs and we also run ab-initio prediction programs on the unmasked sequences. We assigned functional categories by performing BLAST v2.2.28+ searches against the SwissProt database and using a standalone version of InterproScan v-5.17-56.0 with Pfam categories.

To assign genes Clusters of Orthologous Gene annotations we downloaded the bactNOG database v4.5 (page accessed April 2016), which we used to compare the predicted proteins with HMMER v3.1.b1. To identify and compare metabolic pathways we used the KAAS tool (Moriya et al. 2007) provided by the KEGG database (Kanehisa and Goto 2000; Kanehisa et al. 2010) using the BLAST algorithm and single best hit (SBH) with default settings.

Phylogenetic analysis

For phylogenetic reconstruction we used the available genome sequenced strains present in the Ensembl database and used in two previous phylogenetic analyses (Comandatore et al., 2013; Gerth et al., 2014; Table S1). Using the orthofinder software with default settings we

identified the single-copy orthologs in the following *Wolbachia* strains: *WolAcro1, wAdent, wUni, wGlos, wMel, wMelpop, wRi, wHa, wNferru, wNeluco, wNpanzer, wNflava, wAlbB, wNo, wTricho, wVitB, wPel, wJhb, wBm, wCimex, wOncho*. For each of the 269 genes identified we created a nucleotide alignment using muscle, which we examined with the PHITEST. Alignments that showed signs of recombination either with the PHITEST or the MaxChi method were excluded. The remaining 92 alignments were examined using the Xia method for saturation and 1 more alignment was excluded. For the remaining 92 genes we created a concatenated alignment using amas v0.98 (Borowiec, 2015) which was filtered using the gblocks server. An amino acid alignment using the same genes was also created. The nucleotide alignment was tested using jmodeltest and the protein alignment was tested using ProtTest v3.4 (Abascal et al., 2005). The GTR+I+G or the HIVw model was used for phylogenetic analyses (Jones et al., 1992; Le and Gascuel, 2008). Phylogenetic relationships were constructed using the PhyML software using 10,000 bootstrap replicates (Guindon et al., 2010). The final tree was further processed using FigTree v1.4.2 (Rambaut, 2016). For the Bayesian phylogenetic analyses, the same alignments were used and for the nucleotide alignment the GTR+I+G model was used with 4 heated and 1 cold chain and the program was run for 1.000.000 generations using a sampling frequency of 1000. Afterwards, the standard deviation was checked (it approached 0) as well as the effective sample size (ESS) using the sump command (where we verified that ESS was >100). We followed a similar approach for the amino acid alignment. This analysis differed in that we did not specify a model but instead using ‘aamodelpr=mixed’ which averages over models.

Recombination detection and positive selection analyses on orthologs of Wolbachia supergroup A

We performed all comparisons based on the phylogenetic reconstruction and selected a set of 16 most closely related *Wolbachia* strains belonging to supergroups A (*wMel, wMelpop, wRi, wHa, wUni, wGlos, wNpanzer, wNflava, wNleuco, wNferru*) and B (*wPel, wJhb, wVitB, wGmm, wNo, wAlbB*) to compare with *WolAcro1* and *wAdent* strains for signatures of positive selection. To identify single-copy orthologs among the above strains we used the Orthofinder software (Emms and Kelly 2015).

These searches identified a dataset of 747 CDS which we extracted. For each gene we constructed an amino acid sequence using muscle v3.6. The nucleotide sequences and the amino acid alignment were used to construct a codon alignment using the pal2nal software (Suyama et al., 2006). Each codon alignment was checked for stop codons and misalignments manually. We examined all alignments for recombination using the RDP4 software (maxchisq) which revealed approximately 2/3 (3/4) of all CDS were subject to recombination, a percentage similar to what has been reported before (Klasson et al., 2009).

For the positive selection analyses, we used the branch and branch-site REL tests (Yang and Nielsen 2002; Zhang et al. 2005) implemented in the ete toolkit (Huerta-Cepas et al., 2016). In brief, each codon alignment as well as the tree constructed using the concatenated alignment in the current study (Figure 1), were used as inputs. The following models were run for each set of genes: site models bsA and bsA1, b_free and b_neut, b_free and M0 for each of foreground strains (*WolAcro1* and *wAdent*). A likelihood ratio test was used to compare each pair of models and the significance cutoff was set at 1/3,470 (the reciprocal of the number of genes tested).

Reverse transcription-PCR (RT-PCR) and quantitative (qPCR) PCR for selected genes

Four *A. echinatior* colonies were selected for RNA extractions: Ae480, Ae332, Ae331, and Ae280. Colonies were maintained at the Centre for Social Evolution as described above. From each colony, three individuals from each developmental stage were pooled per sample for a total of 3-5 replicates per colony from worker, brown pupae, white pupae, larvae, and male sample types and for 2 replicates per colony for alate samples. In total, 90 samples were analyzed across all four colonies. Individuals were picked from the colony, immediately frozen with liquid N_2, crushed with a plastic mortar and pestle in Qiazol and placed in a tissue lyser with glass beads for 5 minutes at 50Hz. Approximately 15% of the lysate was set aside for later DNA extraction using the Qiagen DNeasy extraction protocol. Chloroform was added to the remaining 85% of the lysate followed by differential centrifugation (12,000 x g for 15 min at 4˚C) in which only the top aqueous layer was recovered and used for subsequent RNA extraction using a RNeasy mini kit (Qiagen, Germany). Approximately 500 ng of total RNA was treated with RQ1 RNase-free DNase I (Promega Corporation, Madison, WI) and reverse transcribed using an iScript RT kit (Bio-Rad) to obtain first-strand cDNA. As a negative control, DNase-treated RNA was examined alongside cDNA samples by PCR (denaturation for 30 s at 95˚C followed by 35 cycles of 15 s at 95˚C and 25 s at 65˚C and a 5-min final extension at 65˚C) and PCR products were visualized on a 1.5% agarose gel. Primers targeting riboflavin genes were designed in Geneious (v. 9.0.4; Table S2). PCR amplification was performed for each primer pair using the same template cDNA and the above settings and the product was purified with the MinElute PCR Purification Kit (Qiagen). The purified DNA concentration was measured with a Qubit 4 Fluorometer (ThermoFisher Scientific) and used to calculate the serial dilutions to create standard curves for each primer pair. RT-PCR reactions were made using SYBR Premix Ex Taq

(Takara) and run in triplicate on a Roche LightCycler 480 system, following the manufacturer's instructions.

Sample collection, RNA isolation, library preparation, and sequencing

Three *A. echinatior* colonies were collected in Gamboa, Panama. Samples were sorted and flash frozen in RNAlater at the Smithsonian Tropical Research Institute. RNA was extracted using the RNAiso Plus extraction protocol (Takara) followed by RNA cleanup using the RNeasy MinElute Cleanup Kit (Qiagen) at the Institute of Advanced Industrial Science and Technology in Tsukuba, Japan. For extraction, samples included three females from each developmental stage (larva, white pupa, brown pupa, worker) with three biological replicates each for a total of 12 samples. Ribosomal RNA (rRNA) depletion was performed using the Ribo-Zero Gold rRNA Removal Kit (Illumina). Libraries were prepared following the TruSeq RNA Sample Preparation Guide (Illumina). RNA quality was assessed using the Agilent 2100 Bioanalyzer and concentrations were quantified using the qPCR library quantification protocol (KAPA biosystems). Libraries were multiplexed and sequenced on the Illumina HiSeq2500 platform (Chemical Dojin, Japan) with 150bp paired-end reads.

RNA sequencing alignment and differential expression analysis

Read quality was assessed using FastQC (Andrews, 2010) and adaptors were removed using Trim Galore! v0.3.1 (Krueger et al., 2015). RNA-seq samples were prepared using the whole ant and as such, reads were mapped separately to the *WolAcro1 Wolbachia* genome as well as the host *A. echinatior* ant genome (GenBank ID: 279248). The annotated genomes were indexed, and reads were subsequently mapped with Burrows-Wheeler Aligner (BWA). The

resulting BAM files were sorted with Cufflinks (v2.2.1; Trapnell et al; 2012) and read counts for each gene were generated with HTSeq (v0.6.1p2; Anders et al., 2015). HTSeq count tables were used as input for DESeq2 (v1.20.0; Anders and Huber, 2010) to identify differentially expressed genes ($p<0.05$) between host developmental stages in both the bacteria and ant.

Weighted gene co-expression network construction and module identification

Weighted gene co-expression network analysis (WGCNA) was conducted with the R package, WGCNA (Langfelder and Horvath, 2008), using rlog normalized expression data from all twelve RNA-seq samples mapped to the *Wolbachia* genome as input. Weighted gene clustering requires the use of a soft thresholding power, β, which preserves the continuous nature of the underlying correlation information. In WGCNA, genes are assigned to networks using an adjacency matrix, calculated by raising the similarity score of gene co-expression across samples to the power of β, to quantify how strongly genes are connected to one another. We selected our β power using the biologically motivated criterion of approximate scale-free topology (see Zhang and Horvath, 2005). Genes were assigned to networks, defined by gene co-expression similarity across samples and by how strongly genes are connected to one another, using a soft thresholding power of $\beta = 8$ to ensure a scale-free network (scale free $R^2 = 0.85$; Figure S1 B-C). We selected the power $\beta = 8$ because it is the lowest power for which the scale-free topology fit curve flattens out upon reaching a high scale independence (approximately $R^2 = 0.90$). There is a trade-off between maximizing how well a network satisfies a scale-free topology model fit (quantified by R^2) and maintaining a high mean number of connections. Node connectivity or centrality is a key concept of network analysis and is highest for nodes overlapping with the greatest number of other nodes.

Gene expression similarity is determined using a pair-wise weighted correlation metric, and clustered according to a topological overlap metric into modules. Networks of genes were clustered using a topological overlap measure (TOM) of normalized expression data, to determine network interconnectedness or proximity. This proximity was used as input for average linkage hierarchical clustering to group interconnected genes with high correlation based on their expression patterns into modules. Module co-expression similarity is quantified by calculating and clustering each module's eigengene, which are considered the best summary of the standardized module expression data. We employed a dynamic branch cutting approach (cut height = 0.3) to the resulting cluster tree to merge highly similar modules, which resulted in twelve remaining modules. The significance of hub GO terms was calculated using the R package, topGO (Alexa and Rahnenfuhrer, 2018) and visualized with REVIGO (Supek et al., 2011). The same WGCNA analysis was also conducted using RNA-seq reads mapped to the *A. echinatior* genome.

Results

General characteristics of the genome

Given its intracellular life style and inability to be cultured in media, *Wolbachia* DNA was purified from the host sample using a combination of filtering techniques and multiple-displacement amplification to obtain sufficient quantities of genomic material for sequencing. The assembled *Wolbachia* genome isolated from *Acromyrmex echinatior* (*WolAcro1*) is estimated to be 1.4 Mb in length with a 35.3% GC content and 1,231 predicted functional protein-coding genes. The *wAdent Wolbachia* genome isolated from *Apterostigma dentigerum* is estimated to be 1.2 Mb in length with a 34.5% GC content and 961 predicted functional protein-

coding sequences. Illumina Miseq coupled with PacBio sequencing generated long reads with a 700x depth of coverage to resolve complex repeat regions. We were unable to completely close the genomes resulting in 17 final scaffolds for *WolAcro1* and 60 scaffolds for *wAdent*. This can likely be attributed to a high amount of recombination, sequence polymorphisms, and repetitive elements, as has been described in other *Wolbachia* genomes (Klasson et al. 2008; Newton et al., 2016).

Wolbachia Phylogenomics

Phylogenetic relationships of 21 *Wolbachia* strains were investigated using single copy orthologs as previously described (Comandatore et al., 2013; Gerth et al., 2014; Table S1). As expected, the A, B, C, D, and F *Wolbachia* supergroups formed distinct branching clades with 100% bootstrap support for all except supergroups C and F, which had 81% bootstrap support (Figure 1). Strain and supergroup placements are consistent with previously described phylogenies (Comandatore et al., 2013; Gerth et al., 2014; Lindsey et al., 2016). Together, *WolAcro1* and *wAdent* formed a separate sister clade to the supergroup A *Wolbachia* strains. These results did not change in any appreciable way when we used amino acid or nucleotide sequences and using Bayesian or maximum likelihood methods (Figures S2-S4).

Genes under positive selection

Positive selection analyses were conducted using all single-copy orthologs from *WolAcro1* and *wAdent* as well as 16 A and B *Wolbachia* supergroup strains. We classified these strains into five groups based on their phylogenetic relationships (Figure 1) and found distinct genes carrying signatures of positive selection among the following groups: (1) The two new ant-

associated *Wolbachia* strains presented in this study; (2) Recently sequenced *Nomada*-associated *Wolbachia* strains (Gerth et al., 2016); (3) *Wolbachia* strains belonging to supergroup B; (4) The more distinctly placed supergroup A strains, *wGlos* and *wUni*; and (5) The remaining supergroup A *Wolbachia* strains. Several genes related to metabolism, recombination and virulence were shown to carry positive selection signatures in these groups. These included genes acting on the purine and terpenoid backbone biosynthesis pathway in supergroup A strains as well as virulence-related genes in supergroup B. Interestingly, we found two genes carrying signatures of positive selection in more than one group: the riboflavin synthase gene, *ribE*, in both fungus-growing ant strains and mosquito-infecting supergroup B strains (Figure 2), and a transcriptional regulator common among several *Wolbachia* strains infecting Nomada species and other supergroup A members.

Wolbachia qPCR results

We chose to further examine the riboflavin production genes with qPCR because the riboflavin operon had shown signatures of positive selection for the for the *Wolbachia* strains infecting fungus-growing ants (Figure 2). Using *WolAcro1* specific primers, each targeting one of the five genes of the riboflavin operon (Table S2), we measured the expression of the genes compared to the *wsp* expression in the same developmental stages we performed the transcriptome analysis plus males and reproductive females (gynes). The results showed there was generally increased expression of many of the genes in gynes and a reduced expression in males, however none of these contrasts were significant (p-value > 0.05; Figure S5).

The Wolbachia transcriptome and age-dependent dynamically expressed genes

To study the nature of *Wolbachia*'s symbiosis with *A. echinatior* and the role that *Wolbachia* may play during the ant's development, we analyzed differentially expressed *Wolbachia* genes at each ant developmental stage. Over 400 million RNA-seq reads were generated in total with an average of 30,864,393 reads per sample. Approximately 12.5% of reads from each sample mapped to the WolAcro1 genome and the three biological replicates per developmental stage cluster well together (Figure S6). The most highly expressed gene across all life stages was *aspB*, an aspartate aminotransferase.

We found 10.3% of genes (127/1231) to be differentially expressed with an adjusted p-value of <0.05 between at least one of the four replicated conditions (larvae, white pupae, brown pupae, workers). Differentially expressed genes generally fall into one of two categories. The first category (Group 1) is a subset of genes (45.7%; 58/127 genes) that are upregulated in the earlier developmental stages (larvae and white pupae; genes marked with a circle in Figure 3). Eight genes from Group 1 were annotated to the KEGG metabolism pathway (*purL*, *purE*, and *ndk*, purine metabolism; *ftsl*, peptidoglycan biosynthesis; *araM*, glycerophospholipid metabolism; *fabD*, fatty acid biosynthesis; *ppsR*; carboxypeptidase Taq) and three ABC transporters (*znuA*, *pstA*, and *ABCB-BAC*). The purL gene, responsible for the conversion of 5-phosphoribosyl N-formylglycinamide (FGAR) to 5-phosphoribosyl N-formylglycinamidine (FGAM) in the purine biosynthesis pathway, was among the top 5% of most highly expressed genes in larvae and white pupae. Additionally, three out of 26 total Ankyrin repeat domain (ANK) genes found in the *WolAcro1* genome were upregulated in the younger developmental stages.

The second category of differentially expressed genes (Group 2) are those upregulated in later developmental stages (brown pupae and workers; 54.3%; 69/127 genes; genes marked with

a shaded triangle in Figure 3). Eighteen of these genes annotated with the KEGG database are involved in genetic information processing (*rplA, rplB, rplQ, RP-L13, fusA, SCO2, argS, rpoBC rpsD, rpsO, trmD, hscA, rspF, dnaE1, secY, groEL, rho, DPO3D1*). These include the chaperone gene, *groEL*, which was the fifteenth most highly expressed gene in brown pupae according to rlog normalized counts and plays an essential role in oxidative stress response (Susin et al., 2006). Thirteen genes from Group 2 were identified by the KEGG database as involved in metabolism (*dapE*, lysine biosynthesis; *ubiD*, ubiquinone biosynthesis; *coxA*, *nuoF*, and *atpB*, oxidative phosphorylation; *pqqL*, zinc protease; *ispDF* and *dxr*, terpenoid backbone biosynthesis; *glyA*, glyoxylate metabolism; *plsY*, glycerophospholipid metabolism; *hflK*, aminoglycoside resistance; *glmU*, amino sugar metabolism; E2.7.7.49). Finally, three genes in Group 2 were identified as ANK genes however they were always more highly expressed in the brown pupa stages relative to the worker stage. The largest number of differentially expressed *Wolbachia* genes was in the contrast between larva and brown pupa (96 genes) and the smallest between brown pupa and workers (13 genes; Table 1). Gene ontology (GO) analysis of the 96 differentially expressed genes between larva and brown pupa reveal the genes upregulated in the brown pupa stage are mostly involved in ion transmembrane transporter activity, oxidoreductase activity (acting on a heme group of donors), GTP binding, iron-sulfur cluster binding, and the structural constituent of ribosome (Figure 8).

Wolbachia weighted gene co-expression network construction and identification of hub genes in the brown module

WGCNA is a useful data mining technique for describing the correlation patterns among genes in biological networks across genomic samples. This analysis identified 12 groups of

interconnected, highly co-expressed networks of *Wolbachia* genes referred to here as modules. The brown module exhibited the highest gene significance and correlation amongst developmental stages with genes from this module negatively correlated with the larva and white pupa stages and positively correlated with the brown pupa and worker stages (p-value = 6e-4, R^2 = 0.84; Figure 5 B-C). Intramodular hub genes are genes with the highest connectivity in their modules and are thus likely to play important roles in the module's respective biological processes. We identified 51 hub genes in the brown module with an absolute value Pearson's correlation of module membership > 0.8 and absolute value of trait significance > 0.2 (Figure 5 C). Of these, 37 hub genes were positively correlated with expression in the older life stages and 14 hub genes were negatively correlated. Amongst the positively correlated hub genes, iron ion binding and single-stranded DNA binding were tied as the two most significantly enriched GO terms (Figure 6). Some notable genes within the positively correlated brown module included the CI-inducing gene, *cifA* and *bfr* (similar to the eukaryotic gene, ferritin, involved in iron uptake) while the negatively correlated genes contained several purine metabolism genes including *purC*, *purE*, and *purL*.

Analysis of cytoplasmic incompatibility-inducing Wolbachia genes

Co-expression of the genes, *cifA* and *cifB*, enables CI-like embryonic lethality (Beckmann et al., 2017; LePage et al., 2017). Similar to the CI-inducing strain, *wMel* (Lindsey et al., 2018), the *cifA* gene was expressed at much higher levels than the *cifB* gene in *WolAcro1* (approximately 3.5-fold greater based on rlog transformed values). In concordance with Lindsey et al. (2018), the *cifA* gene had relatively higher expression during late developmental stages (brown pupae and workers) while the *cifB* gene had the highest average expression in larva,

however differences in expression were marginal. WGCNA clustering results show the *WolAcro1 cifA* gene is positively correlated with genes in the brown module, which had the greatest significance amongst module relationships with older vs. younger developmental stages (Figure 5 B). Genes in the brown module sharing similar expression patterns across samples were mostly involved in iron ion binding and single-stranded DNA binding (Figure 6). The *WolAcro1 cifA* expression profile across samples was most highly correlated with the DNA gyrase gene, *gyrB*, while the *cifB* gene shared the most similar expression profile with the regulatory protein, *repA*. WGCNA analysis clustered the *cifB* gene with the green module which only had a marginally significant negative trait relationship with the white pupae stage (p-value = 0.05; Figure 5 B). GO annotation clustering revealed genes in the green module are most highly involved in nucleoside metabolism, stress response, phosphotransferase activity, and double-stranded RNA binding. Comparisons with *cifA* and *cifB* homologs in other *Wolbachia* strains reveal the *cif* genes in *WolAcro1* and *wAdent* are similar to those in several CI-inducing strains including *wHa* and *wAlbB* (Table 2).

Differentially expressed A. echinatior genes across developmental stages

On average, 70% of reads from each sample mapped to the *A. echinatior* genome. A total of 5,782/ 12,367 unique genes were differentially expressed at an adjusted p-value < 0.05 across at least one pair-wise developmental stage comparison. The comparison between white pupa and workers produced the largest number of differentially expressed genes while the comparison between larva and white pupa produced the lowest (Table 1). By grouping the two younger (larva and white pupa) and two older (brown pupa and workers) stages together we narrowed down our list to 1,190 differentially expressed genes (adjusted p-value < 0.05, log fold change >

2). GO analysis reveals genes upregulated in older stages are largely involved in iron homeostasis related pathways such as heme binding, iron ion binding, and oxidoreductase activity (Figure 4).

A. echinatior weighted gene co-expression network construction and module identification

We explored the patterns of *A. echinatior* gene expression following the same methods as described for *Wolbachia* WGCNA. The turquoise module exhibited the greatest gene significance and correlation when comparing younger and older stages (p-value = 0.0005; Figure S7). Genes in this module were positively correlated with brown pupae and workers and were negatively correlated with larvae and white pupae. The significance of hub GO terms was calculated using topGO and gene ontology enrichments were clustered using DAVID with default settings (Huang et al., 2009). Positively correlated genes in this module were associated with heme, iron, oxidoreductase activity, and transmembrane transport (Figure S8).

Discussion

Wolbachia is ubiquitous across all *A. echinatior* samples, abundantly present in both reproductive and non-reproductive tissues (Andersen et al., 2012; Sapountzis et al., 2015). However, the nature of this symbiosis has remained elusive until now. Using whole genome sequencing and RNA-seq gene expression assays, we are beginning to elucidate the mechanisms of *Wolbachia*'s symbiotic interactions with leaf-cutter ant hosts. We present two draft *Wolbachia* genomes and provide insights into their phylogenetic relationship with previously sequenced strains from supergroups A and B. Transcriptional profiling suggests there are stage-

specific requirements from *Wolbachia* over the course of female *A. echinatior* ant development. Pair-wise comparisons of differentially expressed *Wolbachia* genes across ant developmental stages revealed genes predicted to be involved in stress response and host-microbe interactions and generally fell into the functional categories of lipid metabolism, membrane transport, DNA replication, energy production, and nucleotide metabolism.

Genomic comparisons of WolAcro1 and wAdent to previously sequenced Wolbachia

Phylogenetic analyses suggest *WolAcro1* and *wAdent* form a sister clade to the supergroup A strains. This is consistent with previous studies using PCR methods, which also found the majority of ant-associated *Wolbachia* strains belonged to supergroup A (Frost et al., 2010; Werren and Windsor, 2000). Supergroups A and B differ from other supergroups because of their ability to adapt to new hosts and exhibit a huge diversity of symbiotic interactions ranging from parasitism to facultative and proximate mutualism (Gerth et al., 2014).

Positive selection analysis across 18 supergroup A and B *Wolbachia* strains identified positive selection acting on different classes of genes, potentially reflecting their adaptation to different host niches. However, the *ribE* riboflavin synthase gene was found to have signatures of positive selection in both the leaf-cutter ant-associated *Wolbachia* as well as mosquito-associated *Wolbachia* strains. *Wolbachia* has been shown to provision riboflavin in hosts with specialized diets (e.g. the *w*Cle strain associated with the bedbug, *Cimex lectularius*; Moriyama et al., 2015) however has not been documented in mosquito-infecting *Wolbachia* strains. Leaf-cutter ants have a specialized diet consisting solely of fungal gongylidia. As such, we explored the likelihood that the riboflavin pathway is important for *Wolbachia* maintenance in ants (similar to Fallon et al., 2014) or that *WolAcro1* supplements host riboflavin acquisition during

development by measuring the expression differences across all riboflavin synthase genes using RNA-seq data as well as targeted qPCR assays. Surprisingly, we found no significant difference in expression levels across developmental stages amongst any of the riboflavin genes (Figure S5; Figure 7). As discussed in previous positive selection analyses, the strength of positive selection forces acting on genes may not necessarily correlate with their expression patterns and in fact, genes with reduced expression may be in a evolutionarily flexible state where they are more likely to come under positive selection (Drummond et al., 2005; Kosiol et al., 2008; Larracuente et al., 2008). Thus, low *ribE* expression does not negate the finding that this gene carries signatures of evolution under positive selection and the symbiotic relevance of riboflavin biosynthesis genes in supergroup A and B strains deserves future exploration.

Cytoplasmic incompatibility-inducing Wolbachia genes are expressed in A. echinatior

The most common parasitic *Wolbachia* phenotype is cytoplasmic incompatibility (CI), where eggs from *Wolbachia*-uninfected females fertilized by infected males fail to develop (Yen and Barr, 1971). Although CI and male-killing are theoretically feasible reproductive consequences of *Wolbachia* infection in ants, neither phenomena is evident in social Hymenoptera (Bouwma and Shoemaker, 2011; Keller et al., 2001; Russell, 2012; Wenseleers et al., 2002), apart from *Wolbachia*-infected pharaoh ant colonies exhibiting a modest increase in their female sex ratio (Pontieri et al., 2017). Interestingly, both *cifA* and *cifB* genes are present in the *WolAcro1* and *wAdent* genomes and expressed across all *A. echinatior* developmental stages (Figure 7) with similar expression profiles to other CI-inducing strains (Lindsey et al., 2018). These results appear to support the notion that *WolAcro1* is a CI-inducing strain however more studies will be needed to clarify the role of these genes in ants.

Wolbachia genes involved in iron homeostasis are highly co-expressed and upregulated in brown pupa

Iron is a fundamental part of both macro- and microorganism pathways including heme and ATP biosynthesis, apoptosis, aging, and recycling of intracellular resources (Gill et al., 2014; Schaible and Kaufmann, 2004). Thus, limiting bacterial access to iron can decrease infectivity. Iron is an essential micronutrient however an excess of iron in the cell is harmful. Briefly, iron ABC transporters move iron from the outer membrane into the bacterial cytosol where it is used to produce iron-sulfur clusters and heme, including NADH dehydrogenase I and cytochrome C oxidase. Electron leakage reacts with iron via the Fenton reaction to produce highly reactive hydroxyl radicals that signal an immune response and apoptosis. Cells have various mechanisms of mitigating the accumulation of hydroxyl radicals and *Wolbachia* appears to be one of these contributing forces (Brownlie et al., 2009; Gill et al., 2014; Kremer et al., 2009). Although *Acromyrmex* has an intact heme pathway, this would not preclude additional provision by *Wolbachia* especially during times of high oxidative stress such as molting where *Wolbachia* may supplement host supplies or aid in iron homeostasis (Brownlie et al., 2007; Gill et al., 2014).

Genes involved in iron-dependent activity, including ion transmembrane transport, heme and ATP binding, and iron sulfur cluster formation – co-factors required for ATP production – are significantly upregulated in brown pupa (Figure 8). Thus, *Wolbachia* appears to play an important role in regulating the abundance of iron in the cell, especially prior to the ant's final molt before adulthood. Further supporting this hypothesis, the most significant positive correlations with gene expression in the brown pupa stage were involved in iron-regulating pathways such as oxidative phosphorylation and heme biosynthesis (Table S3). *Wolbachia* ATP

production for its host likely generates damaging reactive oxygen species thereby exacerbating the cell's oxidative stress. However, genes that combat oxidative stress are also differentially expressed in the brown pupa stage, including chaperone proteins, *groEL* and *hscA*. High *groEL* expression has been suggested to be an important adaptation for the intracellular lifestyle of *Wolbachia* as it may act as a buffer mechanism against the accumulation of mildly deleterious mutations in endosymbionts and can protect and maintain appropriate functional protein folding during oxidative stress (Fares et al., 2002; Gutzwiller et al., 2015; Melkani et al., 2004; Moran, 1996). Additionally, the *hscA* chaperone repairs iron-sulfur clusters damaged by oxidants to mitigate continued release of surplus iron (Djaman et al., 2004). Thus, *Wolbachia*'s ability to metabolize iron while responding to oxidative stress may be an important component of its symbiotic relationship with *A. echinatior*.

Up-regulated Wolbachia genes in younger developmental stages are enriched for nucleotide metabolism

Many endosymbionts, including members of the closely related Rickettsia genus, have lost the pathways for de novo purine and pyrimidine biosynthesis and scavenge nucleotides from the host with ATP and ADP translocases (Brownlie et al., 2007; Grote et al., 2017). *Wolbachia* however encodes complete purine and pyrimidine biosynthetic pathways while lacking the ATP/ADP nucleotide translocases used by other endosymbionts. Although *A. echinatior* is not dependent on *Wolbachia* for nucleotide provisioning, these observations suggest *Wolbachia* metabolize nucleotides not only for internal consumption but also to supplement hosts during developmental stages where the demand for rapid replication and cellular division is particularly high (Foster et al., 2005; Grote et al., 2017). Several genes involved in purine metabolism (*purL*,

purE, and *ndk*) are differentially expressed with significantly higher expression in the younger life stages (larva and white pupa) relative to the older life stages (brown pupa and workers; Figure 3). *Wolbachia* rapidly divides during the larval and initial pupal phases requiring amplified expression of replication machinery to increase cell division rates during these early developmental stages.

Expression of Ankyrin repeat domain genes previously implicated in host-microbe interaction

Ankyrin repeat domain (ANK) genes are common in eukaryotic and viral life domains yet generally rare in bacteria (Al-Khodor et al., 2010; Siozios et al., 2013). ANK genes are established type IV secretion system (T4SS) effectors that mediate protein-protein interactions and are known to be involved in an abundance of cell processes including transcriptional regulation, development, signal transduction, sex differentiation, and intracellular trafficking (Sedgwick and Smerdon, 1999). These genes are suggested to play a critical role in the molecular mechanisms that determine *Wolbachia*-host symbiotic interactions and may even play a role in lateral gene transfer events (Dunning Hotopp et al., 2007; Gutzwiller et al., 2015; Iturbe-Ormaetxe et al., 2005; Pan et al., 2008; Sinha et al., 2019; Wu et al., 2004). Of the 26 ANK genes identified in the *WolAcro1 Wolbachia* genome, six were differentially expressed, with notable upregulation observed in the larvae, white pupae, or brown pupae stages. Similar expression results across host developmental stages were also observed in the *wMel Wolbachia* strain infecting *D. melanogaster* (Gutzwiller et al., 2015). Future research focusing on the larvae and pupae stages is merited to study how *Wolbachia* manipulates host biology to maintain persistent infections.

Conclusion

This study provides first insight into the complex relationship between *Wolbachia* and the leaf-cutting ant, *A. echinatior*, a social insect symbiosis that is unexplored yet frequently occurring in nature. The availability of the first ant-derived *Wolbachia* genomes, *WolAcro1* and *wAdent*, will enable further molecular and genetic analyses providing insights into the phylogenetic relationships of *Wolbachia* supergroups. Rather than rely on a single process or pathway, this symbiosis likely relies on a symphony of complex interactions that fluctuate over the development of the host. We identified 127 *Wolbachia WolAcro1* genes exhibiting robust stage-specific differential expression which may serve as candidates for deepening our understanding, and potentially manipulating, the genetic basis of how *Wolbachia* interacts with these eusocial arthropod hosts. Future studies can leverage results from this study to functionally validate the proposed roles *Wolbachia* has in its ant host by, for example, using CRISPR/Cas9 site-specific gene editing systems or RNA*i* gene silencing experiments. These genome and transcriptome analyses provide a baseline for future research of *Wolbachia*-host dynamics that may promote practical applications such as *Wolbachia*-based pest and disease control.

Figures and Tables

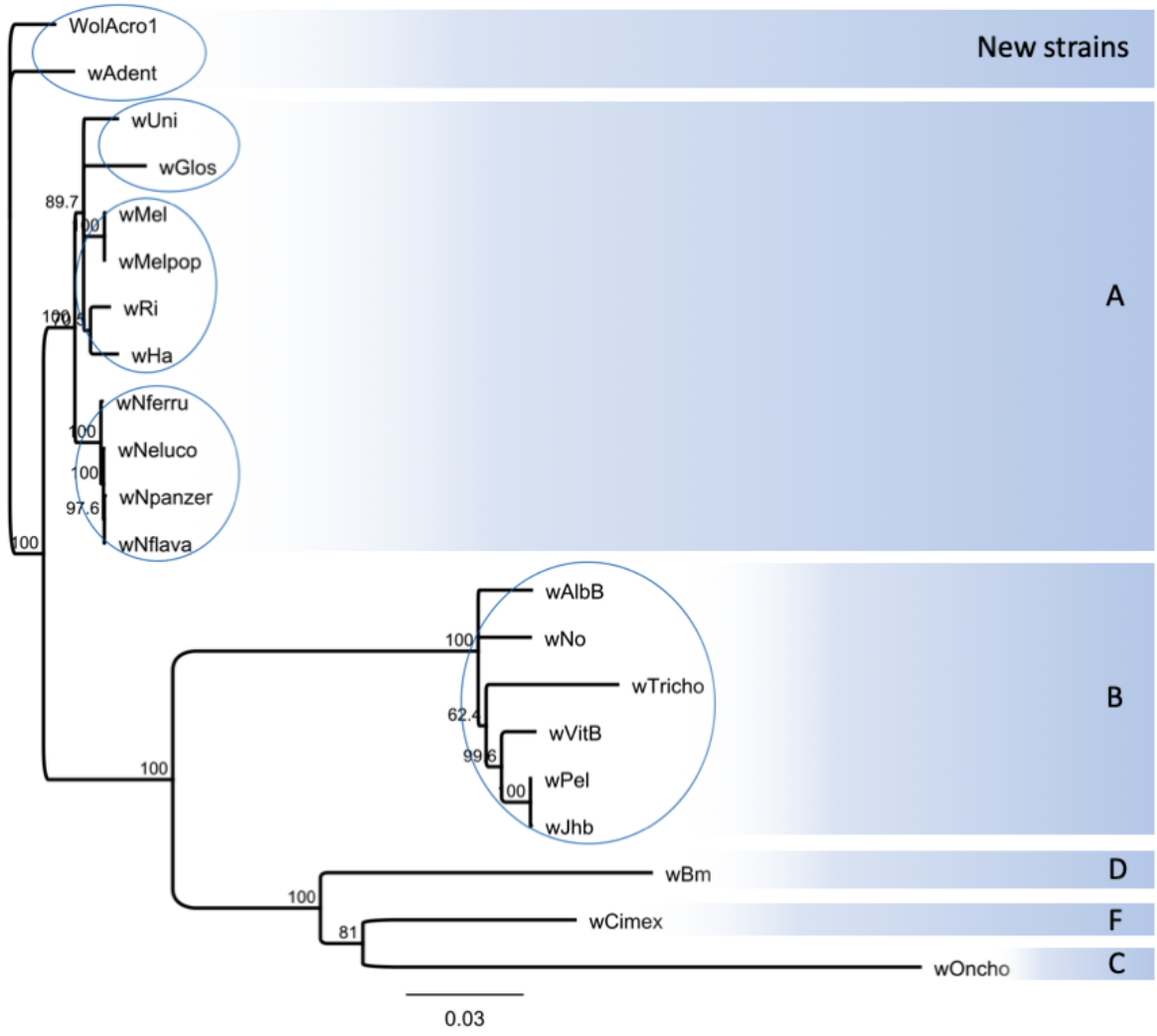


Figure 1: Bayesian phylogeny of 21 *Wolbachia* strains. Bootstrap confidence values from 10,000 replicates are shown as percentages at each node. The scale bar indicates inferred evolutionary changes. Supergroups are shown in colored boxes and labeled in the center right of each box. Blue circles identify closely related strains.

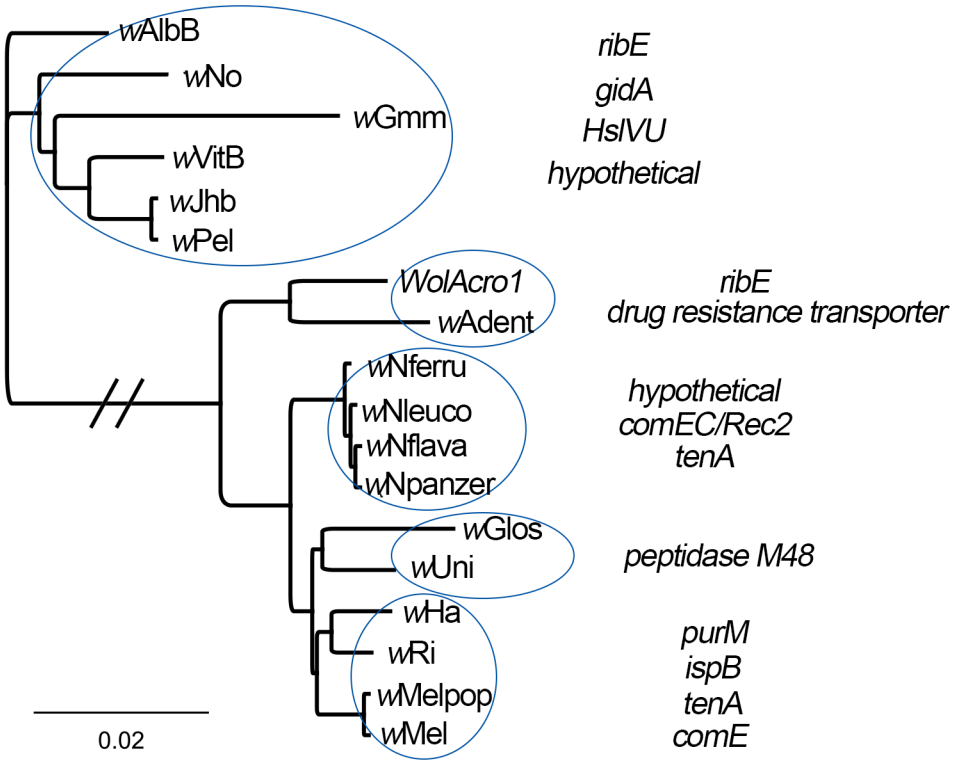
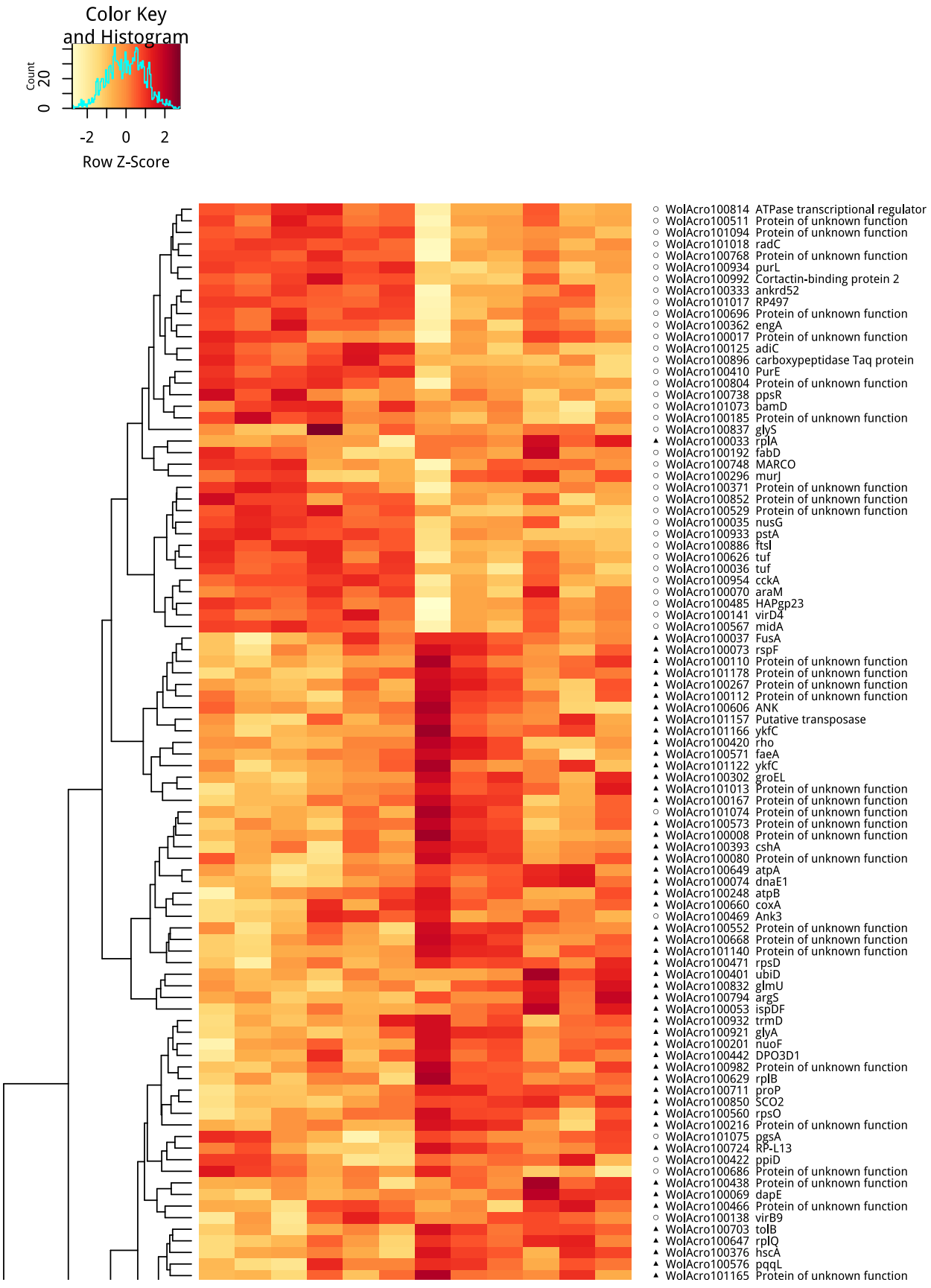


Figure 2: Positive selection analysis and phylogenetic relationship of closely related supergroup A and B *Wolbachia* strains. Genes with signatures of evolution under positive selection are listed to the right of circled closely related strains.


Color Key
and Histogram
Count
0
20
-2
0
2
Row Z-Score
WolAcro100814 ATPase transcriptional regulator
WolAcro100511 Protein of unknown function
WolAcro101094 Protein of unknown function
WolAcro101018 radC
WolAcro100768 Protein of unknown function
WolAcro100934 purL
WolAcro100992 Cortactin-binding protein 2
WolAcro100333 ankrd52
WolAcro101017 RP497
WolAcro100696 Protein of unknown function
WolAcro100362 engA
WolAcro100017 Protein of unknown function
WolAcro100125 adiC
WolAcro100896 carboxypeptidase Taq protein
WolAcro100410 PurE
WolAcro100804 Protein of unknown function
WolAcro100738 ppsR
WolAcro101073 bamD
WolAcro100185 Protein of unknown function
WolAcro100837 glyS
WolAcro100033 rplA
WolAcro100192 fabD
WolAcro100748 MARCO
WolAcro100296 murJ
WolAcro100371 Protein of unknown function
WolAcro100852 Protein of unknown function
WolAcro100529 Protein of unknown function
WolAcro100035 nusG
WolAcro100933 pstA
WolAcro100886 ftsI
WolAcro100626 tuf
WolAcro100036 tuf
WolAcro100954 cckA
WolAcro100070 araM
WolAcro100485 HAPgp23
WolAcro100141 virD4
WolAcro100567 midA
WolAcro100037 FusA
WolAcro100073 rspF
WolAcro100110 Protein of unknown function
WolAcro101178 Protein of unknown function
WolAcro100267 Protein of unknown function
WolAcro100112 Protein of unknown function
WolAcro100606 ANK
WolAcro101157 Putative transposase
WolAcro101166 ykfC
WolAcro100420 rho
WolAcro100571 faeA
WolAcro101122 ykfC
WolAcro100302 groEL
WolAcro101013 Protein of unknown function
WolAcro100167 Protein of unknown function
WolAcro101074 Protein of unknown function
WolAcro100573 Protein of unknown function
WolAcro100008 Protein of unknown function
WolAcro100393 cshA
WolAcro100080 Protein of unknown function
WolAcro100649 atpA
WolAcro100074 dnaE1
WolAcro100248 atpB
WolAcro100660 coxA
WolAcro100469 Ank3
WolAcro100552 Protein of unknown function
WolAcro100668 Protein of unknown function
WolAcro101140 Protein of unknown function
WolAcro100471 rpsD
WolAcro100401 ubiD
WolAcro100832 glmU
WolAcro100794 argS
WolAcro100053 ispDF
WolAcro100932 trmD
WolAcro100921 glyA
WolAcro100201 nuoF
WolAcro100442 DPO3D1
WolAcro100982 Protein of unknown function
WolAcro100629 rplB
WolAcro100711 proP
WolAcro100850 SCO2
WolAcro100560 rpsO
WolAcro100216 Protein of unknown function
WolAcro101075 pgsA
WolAcro100724 RP-L13
WolAcro100422 ppiD
WolAcro100686 Protein of unknown function
WolAcro100438 Protein of unknown function
WolAcro100069 dapE
WolAcro100466 Protein of unknown function
WolAcro100138 virB9
WolAcro100703 tolB
WolAcro100647 rplQ
WolAcro100376 hscA
WolAcro100576 pqqL
WolAcro101165 Protein of unknown function

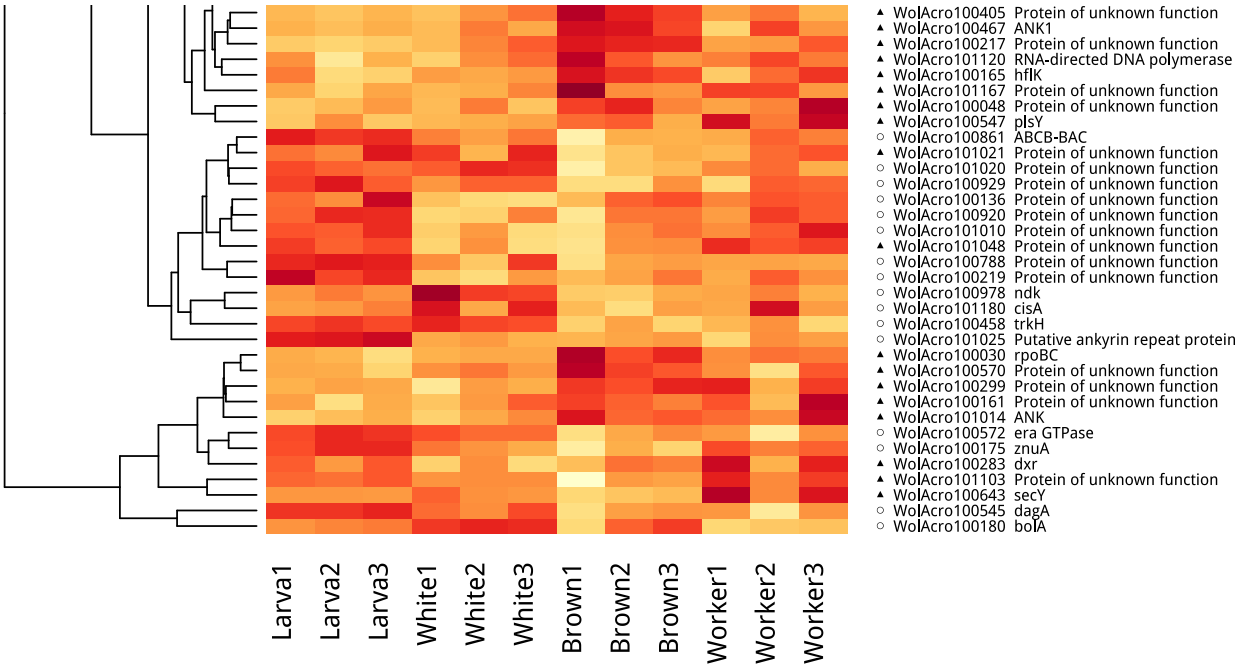


Figure 3: Differentially expressed *Wolbachia* genes with an adjusted p-value <0.05 for at least one life stage comparison based on normalized r log transformed values. Expression values were scaled prior to clustering using a Z-score calculation (red indicating above average expression and yellow indicating below average expression). To the right of the heat map, genes with a higher expression in the earlier life stages (larvae and white pupae) are marked with an open circle while genes more highly expressed in the later life stages (brown pupae and workers) are indicated by a shaded triangle. The right middle column lists the *WolAcro1* gene name used in this study followed by the KEGG gene annotation in the rightmost column.

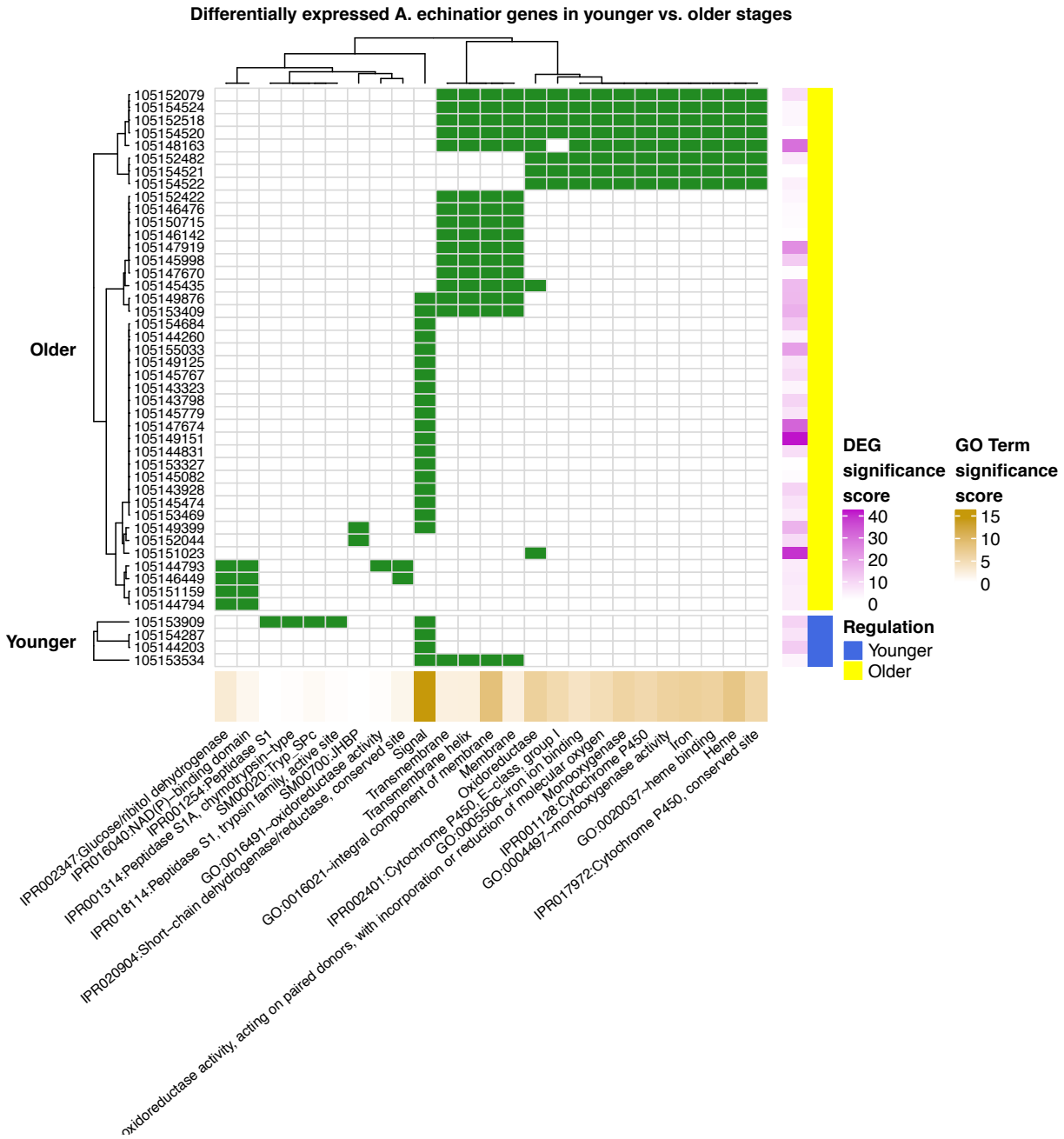


Figure 4: Clustering of DAVID gene ontology enrichments for differentially expressed *A. echinatior* genes across younger vs. older developmental groups (adjusted p-value <0.001, absolute log fold change > 4). *A. echinatior* genes are listed on the left using Entrez Gene IDs. Each green square in the main plot indicates a gene ontology term associated with the gene in that respective row. To the right of the figure is a magenta heatmap showing the significance

score of differentially expressed genes and whether they are upregulated in the younger (blue) or the older (yellow) developmental stages. Below is a brown heatmap showing the significance of each GO term, indicating the relative number of genes that identified with that GO term.

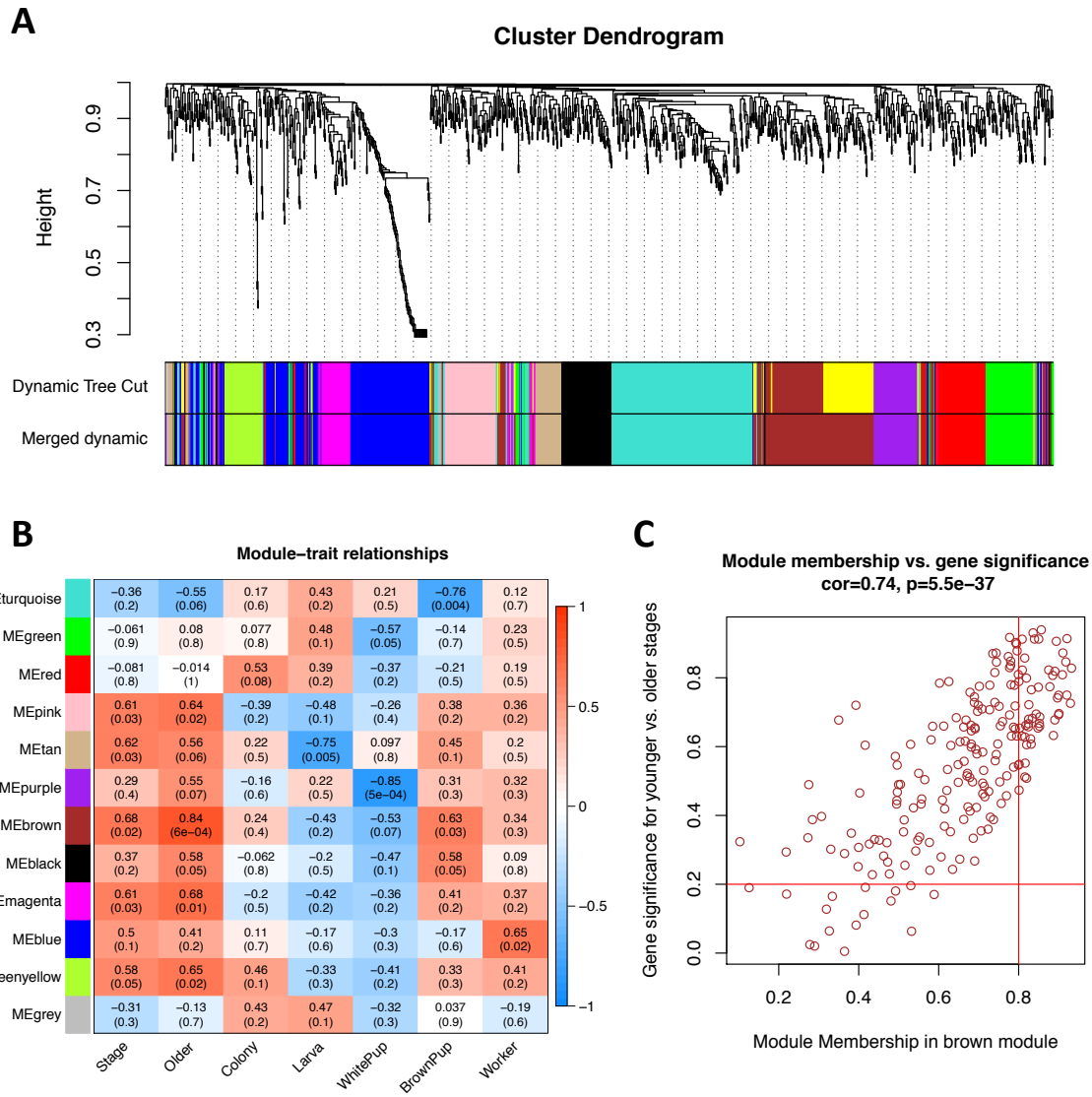


Figure 5: *Wolbachia* WGCNA results using the rlog normalized gene expression data from twelve RNA-seq samples (3 larvae, 3 white pupae, 3 brown pupae, and 3 workers from *A. echinatior* colonies) mapped to the *Wolbachia* genome. (A) Dendrogram of all *Wolbachia* genes clustered on a dissimilarity measure (1-TOM) of normalized expression data. Each vertical line represents a gene and densely clustered branches of the tree reveal interconnected, highly co-expressed networks of genes belonging to the modules shown below. Genes are assigned to colored modules shown on the bottom; genes that are unassigned to a module are shown as gray. "Dynamic Tree Cut" displays the original module assignments and "Merged Dynamic" shows the module assignments after merging those with highly similar expression profiles corresponding to a 70% correlation between module eigengenes. (B) Module-trait association based on the summary profile of each module (eigengenes) to identify modules that are significantly associated with the sample traits. Each row corresponds to a module while each column corresponds to a trait. Each cell contains the test statistic value and its corresponding p-value. The table includes a heatmap representing module eigengene correlation according to the color legend on the right. (C) Scatterplot of the associations of individual genes in the brown module with our trait of interest, "older". Gene significance (GS) is shown on the y-axis representing the absolute value of the correlation between the gene and the older trait, while the x-axis shows module membership (MM) defined as the correlation between the module eigengene and the gene expression profile. Genes in the brown module with a GS > 0.2 (horizontal red line) and MM > 0.8 (vertical red line) reveal hub genes (highly interconnected genes) that are highly significant to the older trait.

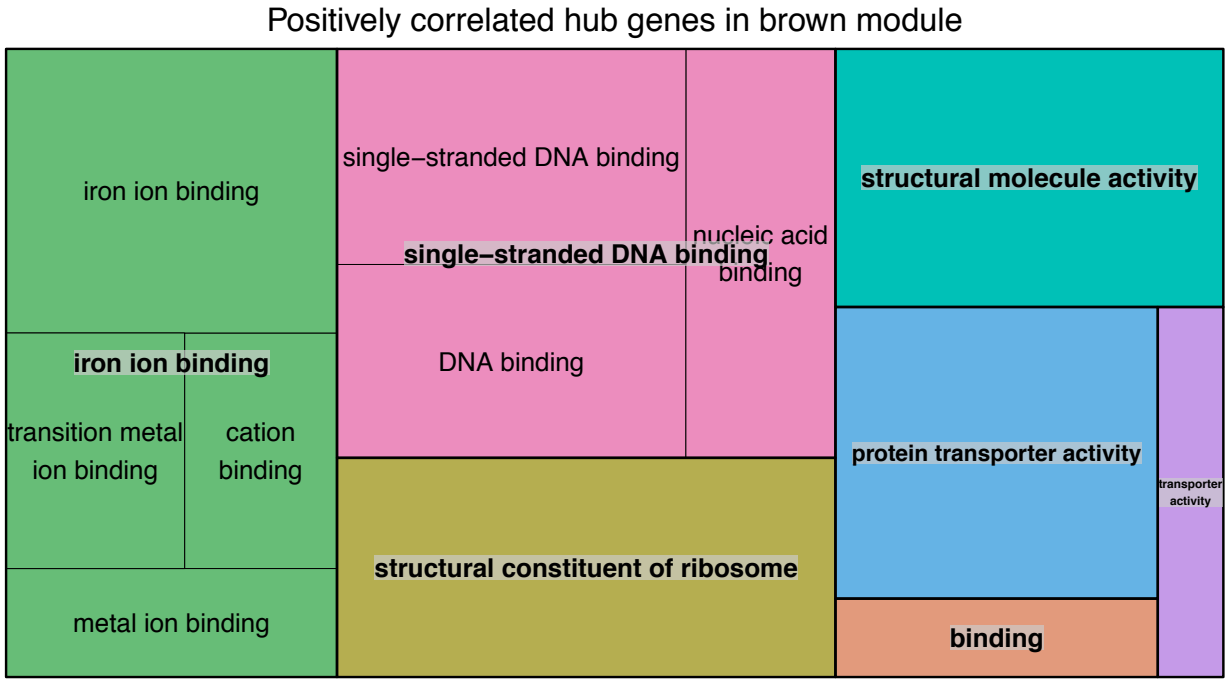


Figure 6: Revigo treemap of statistically significant GO terms (adjusted p-value < 0.05) associated with positively correlated *Wolbachia* hub genes in the brown WGCNA module. WGCNA results reveal *Wolbachia* genes in the brown module are positively correlated with the two older developmental stages, brown pupa and workers (p-value = 0.0006, r2 = 0.84). Each rectangle contains a GO cluster representative and the size of the rectangle reflects the p-value of the GO term.

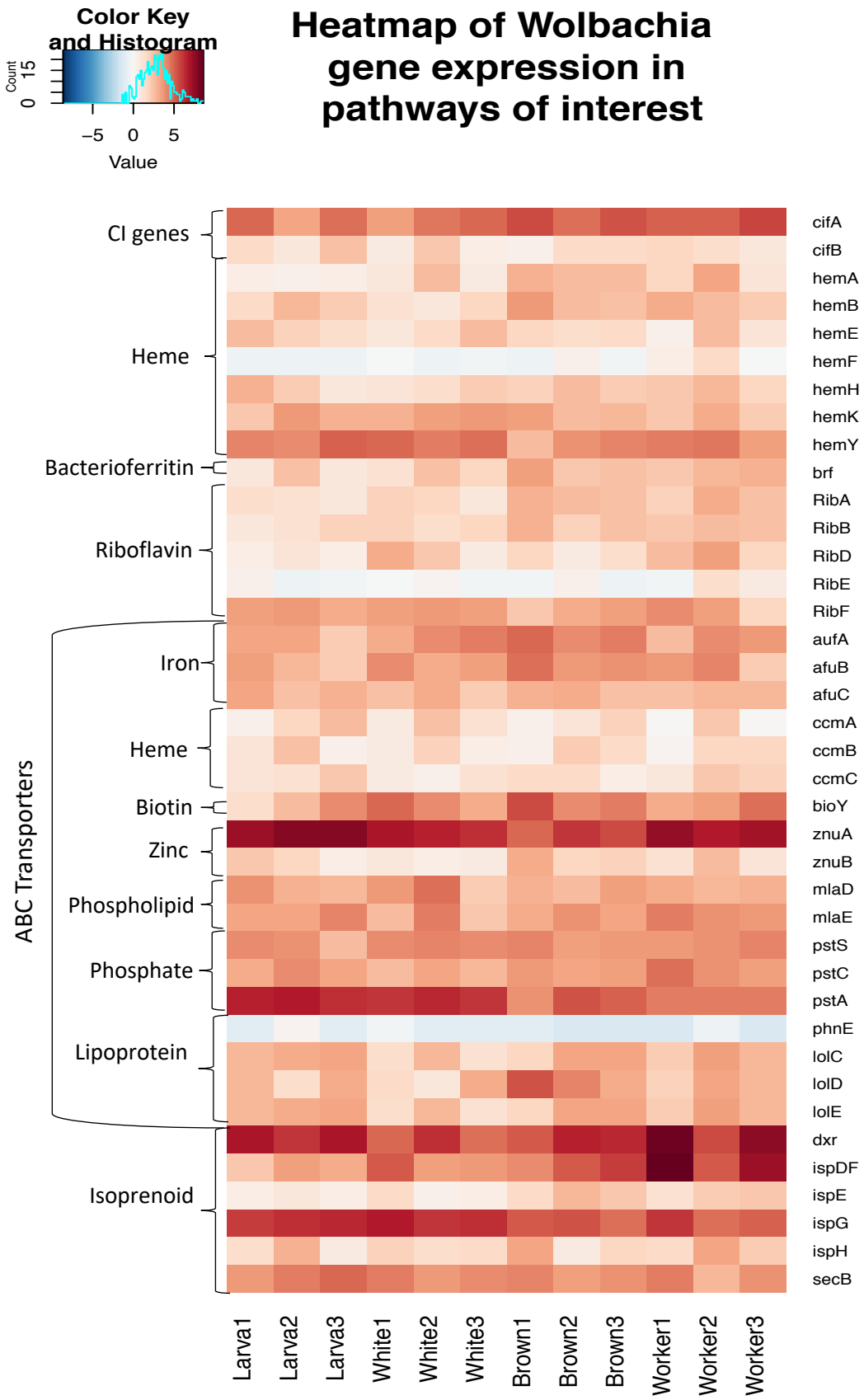

Color Key and Histogram
Count
0
15
−5
0
5
Value
Heatmap of Wolbachia gene expression in pathways of interest
CI genes
Heme
Bacterioferritin
Riboflavin
Iron
Heme
Biotin
Zinc
Phospholipid
Phosphate
Lipoprotein
Isoprenoid
ABC Transporters
cifA
cifB
hemA
hemB
hemE
hemF
hemH
hemK
hemY
brf
RibA
RibB
RibD
RibE
RibF
aufA
afuB
afuC
ccmA
ccmB
ccmC
bioY
znuA
znuB
mlaD
mlaE
pstS
pstC
pstA
phnE
lolC
lolD
lolE
dxr
ispDF
ispE
ispG
ispH
secB
Larva1
Larva2
Larva3
White1
White2
White3
Brown1
Brown2
Brown3
Worker1
Worker2
Worker3

Figure 7: Heatmap of *Wolbachia* gene expression for pathways of interest. Expression is plotted using rlog transformed normalized gene expression values (DESeq2, v1.20.0) across *A. echinatior* developmental stages (from youngest to oldest: larva, white pupa, brown pupa, workers). Pathways of interest include heme, riboflavin, and isoprenoid biosynthesis, ABC transporters for iron, heme, biotin, zinc, phospholipid, phosphate, and lipoprotein, as well as three genes referenced in the discussion section, bacterioferritin, and the two CI-inducing genes, *cifA* and *cifB*.

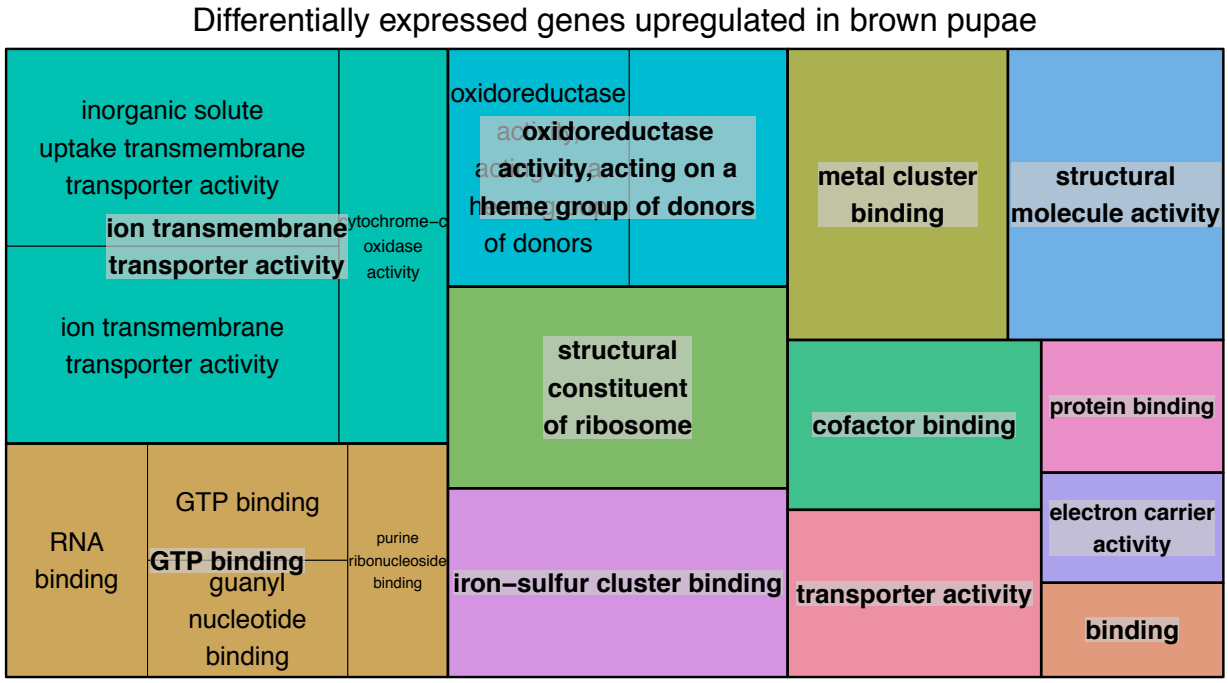


Figure 8: Revigo treemap of statistically significant (adjusted p-value < 0.05) differentially expressed *Wolbachia* genes upregulated in brown pupa samples relative to larvae. This comparison had the largest number of differentially expressed genes of any pairwise comparison across developmental stages. Each rectangle contains a GO cluster representative and the size of the rectangle reflects the p-value of the GO term.

	Larva - W. Pupa	Larva - B. Pupa	Larva - Workers	W. Pupa - B. Pupa	W. Pupa - Workers	B. Pupa - Workers
Total DE *Wolbachia*	3 (0.2%)	91 (7.4%)	36 (2.9%)	35 (2.8%)	15 (1.2%)	13 (1.1%)
***Wolbachia* Up**	1	39	14	23	7	8
***Wolbachia* Down**	2	52	22	12	8	5
Total DE *A. echinatior*	295 (2.4%)	2348 (19%)	3855 (31.2%)	2646 (21.4%)	3998 (32.3%)	2087 (16.9%)
***A. echinatior* Up**	157	968	1490	1198	1632	853
***A. echinatior* Down**	138	1380	2365	1448	2366	1234

Table 1: Differential gene expression results across each pair-wise developmental stage comparison. Results from RNA-seq reads mapped to the *Wolbachia* genome are shown above results from reads mapped to the *A. echinatior* genome. Upregulated genes have a higher expression in the first stage listed in the comparison (the younger stage) while downregulated genes are more highly expressed in the second stage listed (the older stage). Percentages in parentheses show the percent of total *Wolbachia* (1,231) and total *A. echinatior* (12,367) genes that are differentially expressed in each comparison.

	wAdent 00768-RA	wNflava |WP_0650 95437.1	wNeluco |WP_0650 95437.1	wNferru |WP_0650 95437.1	wNpanser |WP_0650 95168.1	wJhb |EEB5 5312	wPel |CAQ5 4402	wAlbB |CCE7 7107	wAlbB |CCE7 7512	wAdent 00728-RA	wMel |WD0631	wMelpop |ERN55636	wRi |ACN9 5317	wRi |ACN9 5711	WolAcro 100876-RA	wHa |AGI9 9744	wNpanser |WP_0650 94606.1	wJhb |EEB5 5247	wJhb |EEB5 5178	wPel |CAQ5 4390	wRi |ACN9 5432	wNflava |WP_0650 95361.1	wNeluco |WP_0650 95361.1	wAdent 00294-RA	WolAcro 101012-RA
wAdent 00768-RA		51.5	51.5	51.5	51.3	50.3	50.3	50	23.9	23.6	29.4	29.4	29.7	29.7	32	30.7	31.2	31.7	26.3	32.2	33.5	30.5	30.5	39.2	31.9
wNflava |WP_0650 95437.1	51.5		100	100	99.5	91.1	91.1	90.1	23.4	23.4	31	31	31.5	31.5	32	30.7	32.5	33.2	26	33.5	32.2	27.4	27.4	37.2	29.4
wNeluco |WP_0650 95437.1	51.5	100		100	99.5	91.1	91.1	90.1	23.4	23.4	31	31	31.5	31.5	32	30.7	32.5	33.2	26	33.5	32.2	27.4	27.4	37.2	29.4
wNferru |WP_0650 95437.1	51.5	100	100		99.5	91.1	91.1	90.1	23.4	23.4	31	31	31.5	31.5	32	30.7	32.5	33.2	26	33.5	32.2	27.4	27.4	37.2	29.4
wNpanser |WP_0650 95168.1	51.3	99.5	99.5	99.5		90.9	90.9	89.8	23.6	23.6	30.7	30.7	31.2	31.2	31.7	30.5	32.2	33	25.7	32.7	32	27.2	27.2	37.2	28.9
wJhb |EEB5 5312	50.3	91.1	91.1	91.1	90.9		100	97	24.1	24.4	30.5	30.5	30.7	30.7	31.7	30.5	32.2	33	25.7	32.7	31.7	27.7	27.7	39.9	28.9
wPel |CAQ5 4402	50.3	91.1	91.1	91.1	90.9	100		97	24.1	24.4	30.5	30.5	30.7	30.7	31.7	30.5	32.2	33	25.7	32.7	31.7	27.7	27.7	39.9	28.9
wAlbB |CCE7 7107	50	90.1	90.1	90.1	89.8	97	97		23.9	23.9	30.2	30.2	30.5	30.5	31.7	30.5	32	33	25.5	32.7	31.5	27.7	27.7	39.2	29.2
wAlbB |CCE7 7512	23.9	23.4	23.4	23.4	23.6	24.1	24.1	23.9		98.5	30.7	30.7	31.2	31.2	34	34	34.8	34.5	28.4	34.5	35.3	40.1	40.1	39.2	36.5
wAdent 00728-RA	23.6	23.4	23.4	23.4	23.6	24.4	24.4	23.9	98.5		30.2	30.2	30.7	30.7	33.5	33.5	34.3	34	28.1	34	35	39.8	39.8	38.5	36.2
wMel |WD0631	29.4	31	31	31	30.7	30.5	30.5	30.2	30.7	30.2		100	99	99	67.8	71.3	70.6	70.6	60.2	70.8	51	46.4	46.4	56.1	51.5
wMelpop |ERN55636	29.4	31	31	31	30.7	30.5	30.5	30.2	30.7	30.2	100		99	99	67.8	71.3	70.6	70.6	60.2	70.8	51	46.4	46.4	56.1	51.5
wRi |ACN9 5317	29.7	31.5	31.5	31.5	31.2	30.7	30.7	30.5	31.2	30.7	99	99		100	68	71.8	71.1	71.1	60.7	71.3	51.5	46.7	46.7	56.1	52
wRi |ACN9 5711	29.7	31.5	31.5	31.5	31.2	30.7	30.7	30.5	31.2	30.7	99	99	100		68	71.8	71.1	71.1	60.7	71.3	51.5	46.7	46.7	56.1	52
WolAcro 100876-RA	32	32	32	32	31.7	31.7	31.7	31.7	34	33.5	67.8	67.8	68	68		84.5	85.8	85.8	62.6	80.2	52	48.7	48.7	58.8	52.6
wHa |AGI9 9744	30.7	30.7	30.7	30.7	30.5	30.5	30.5	30.5	34	33.5	71.3	71.3	71.8	71.8	84.5		82	87.8	64.7	84.3	50.8	48.5	48.5	54.1	52
wNpanser |WP_0650 94606.1	31.2	32.5	32.5	32.5	32.2	32.2	32.2	32	34.8	34.3	70.6	70.6	71.1	71.1	85.8	82		90.9	73.7	90.4	52	48	48	58.1	52.3
wJhb |EEB5 5247	31.7	33.2	33.2	33.2	33	33	33	33	34.5	34	70.6	70.6	71.1	71.1	80.7	87.8	90.9		76.7	95.9	51.5	47.2	47.2	53.4	50.1
wJhb |EEB5 5178	26.3	26	26	26	25.7	25.7	25.7	25.5	28.4	28.1	60.2	60.2	60.7	60.7	62.6	64.7	73.7	76.7		80.9	39.8	36.6	36.6	36.6	41.7
wPel |CAQ5 4390	32.2	33.5	33.5	33.5	33.2	32.7	32.7	32.7	34.5	34	70.8	70.8	71.3	71.3	80.2	84.3	90.4	95.9	80.9		50.5	45.9	45.9	54.1	50.1
wRi |ACN9 5432	33.5	32.2	32.2	32.2	32	31.7	31.7	31.5	35.3	35	51	51	51.5	51.5	52	50.8	52	51.5	39.8	50.5		54.8	54.8	62.2	57.2
wNflava |WP_0650 95361.1	30.5	27.4	27.4	27.4	27.2	27.7	27.7	27.7	40.1	39.8	46.4	46.4	46.7	46.7	48.7	48.5	48	47.2	36.6	45.9	54.8		100	64.2	54.2
wNeluco |WP_0650 95361.1	30.5	27.4	27.4	27.4	27.2	27.7	27.7	27.7	40.1	39.8	46.4	46.4	46.7	46.7	48.7	48.5	48	47.2	36.6	45.9	54.8	100		64.2	54.2
wAdent 00294-RA	39.2	37.2	37.2	37.2	37.2	39.9	39.9	39.2	39.2	38.5	56.1	56.1	56.1	56.1	58.8	54.1	58.1	53.4	28.2	54.1	62.2	64.2	64.2		82.6
WolAcro 101012-RA	31.9	29.4	29.4	29.4	28.9	28.9	28.9	29.2	36.5	36.2	51.5	51.5	52	52	52.6	52	52.3	50.1	41.7	50.1	57.2	54.2	54.2	82.6	

Table 2: *Wolbachia* CI genes in leaf cutter ants. Distance matrices show the more similar genes to the *cifA* and *cifB* genes in darker red. Each box shows percent similarity between CI genes from strains listed across the top row and the leftmost column. Strain abbreviations are listed first

in bold followed by the gene ID. Full strain names and genome accession numbers can be found in Table S1. *A. echinatior* (*WolAcro1*) and *A. dentigerum* (*w*Adent) CI genes are outlined in bold.

Supplemental Results

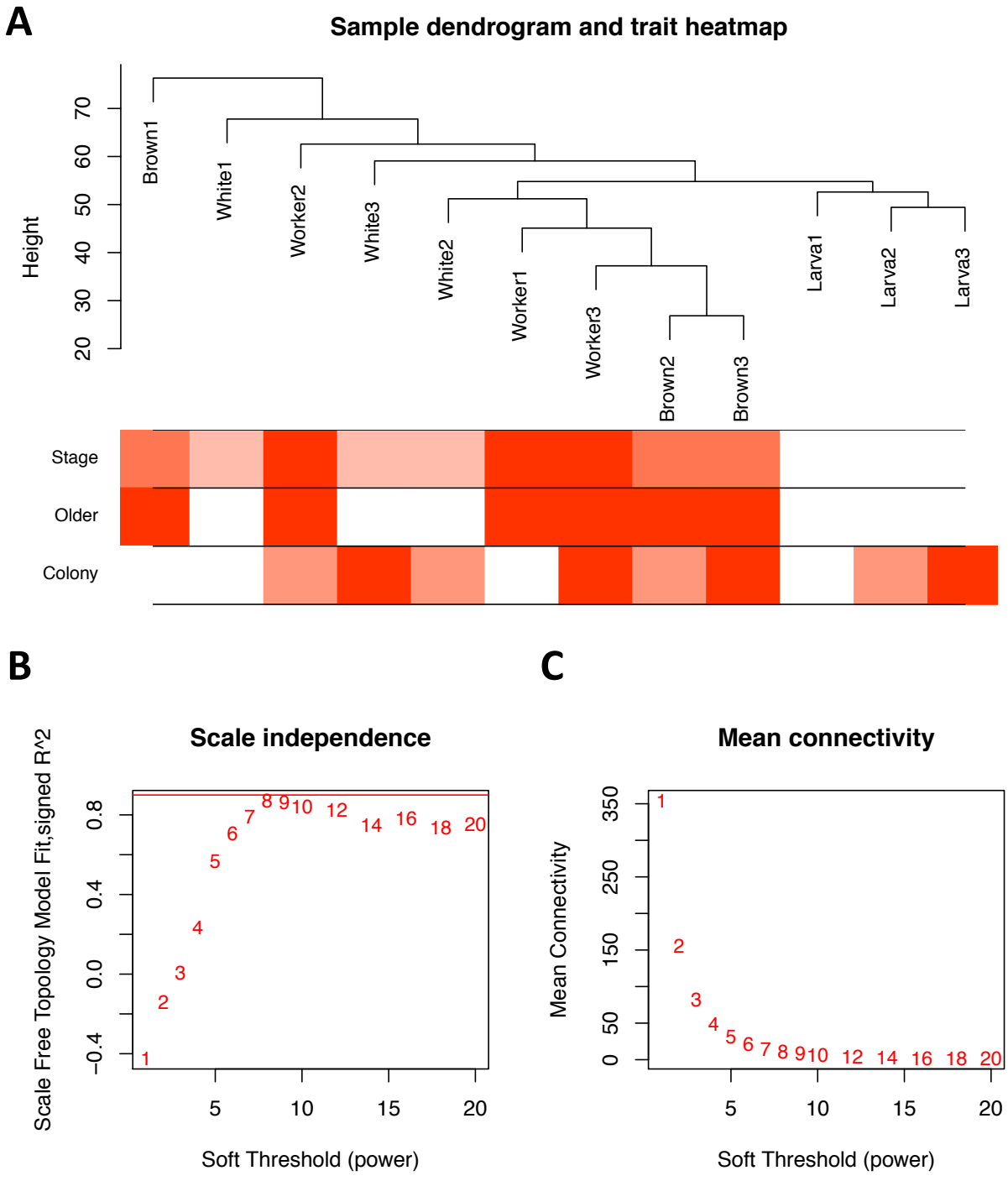


Figure S1: *Wolbachia* WGCNA results using the rlog normalized gene expression data from twelve RNA-seq samples (3 larvae, 3 white pupae, 3 brown pupae, and 3 worker samples from *A. echinatior* colonies) mapped to the *Wolbachia* genome. (A) Clustering dendrogram of samples based on their Euclidean distance with trait heatmap below. This plot is used to detect sample outliers and visualize how the sample traits relate to the sample dendrogram. The trait "stage" refers to the developmental classification of each RNA-seq sample (larva, white pupa, brown pupa, workers) with color intensity increasing with sample age. In the "older" trait classification, the developmental stages are made binary with the two younger developmental stages, larva and white pupa (shown as white), forming one group and the two older stages, brown pupa and workers (red), composing the second group. The trait, "colony" refers to the three ant nests from which the samples were collected. In contrast to the Euclidean distance method for clustering samples in (A), weighted gene clustering requires the use of a soft thresholding power, β. (B) Scale-free fit index analysis (y-axis) shown as a function of candidate soft-thresholding powers (x-axis). (C) Mean connectivity (degree, y-axis) as a function of various candidate soft-thresholding powers (x-axis).

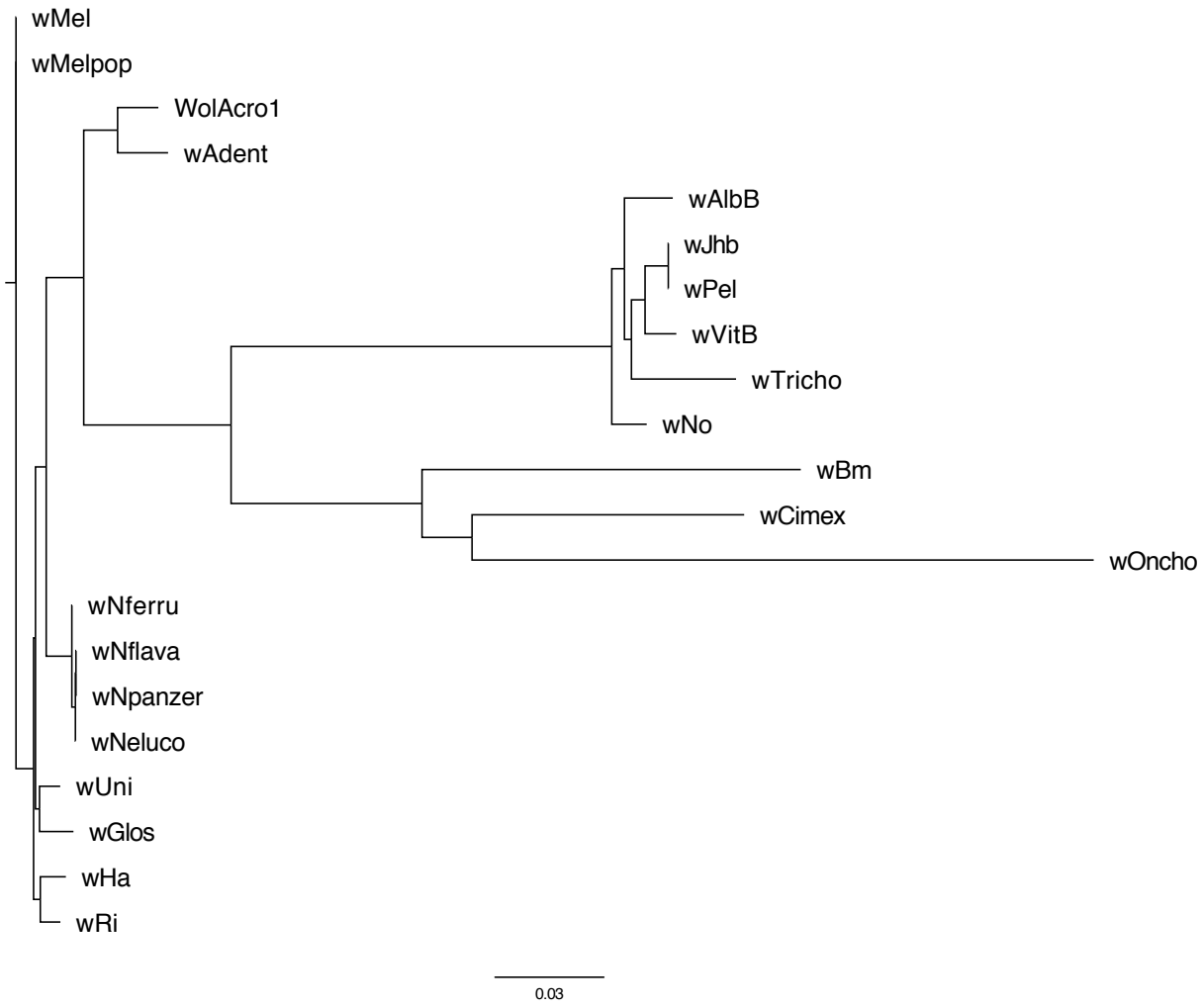


Figure S2: Bayesian phylogeny using the nucleotide sequences of 21 *Wolbachia* strains.

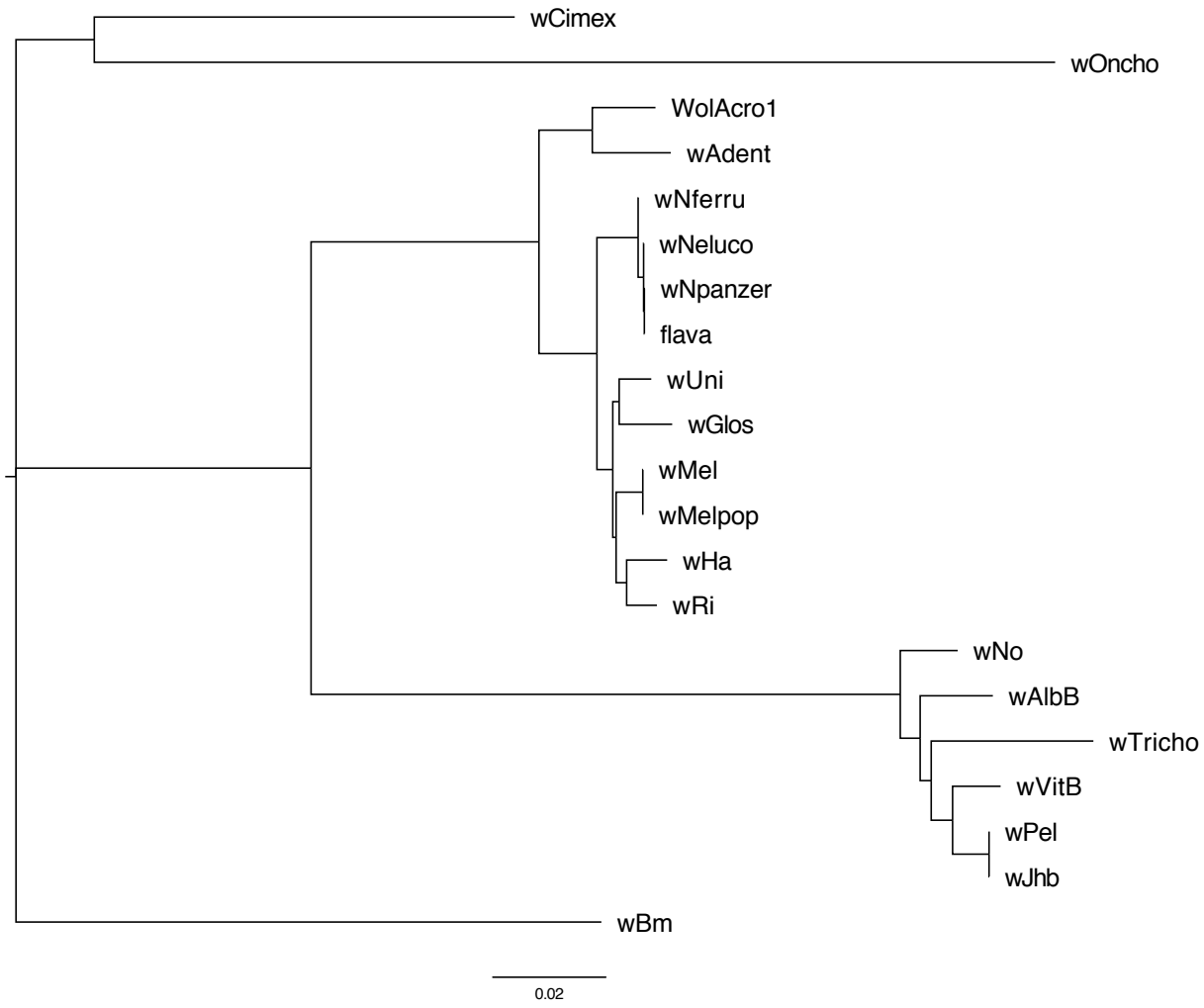


Figure S3: Maximum likelihood phylogeny using the nucleotide sequences of 21 *Wolbachia* strains.

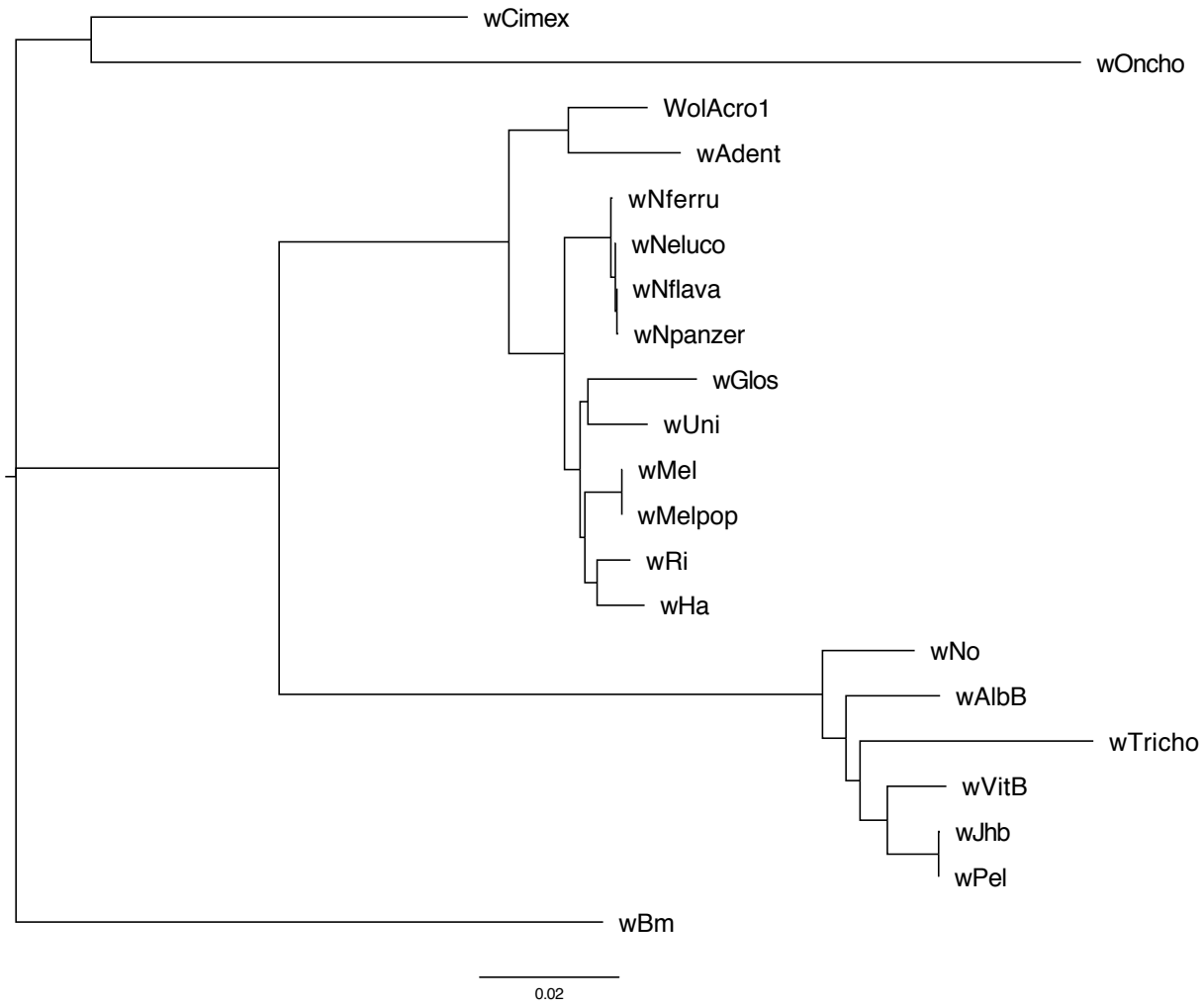


Figure S4: Maximum likelihood phylogeny using the amino acid sequences of 21 *Wolbachia* strains.

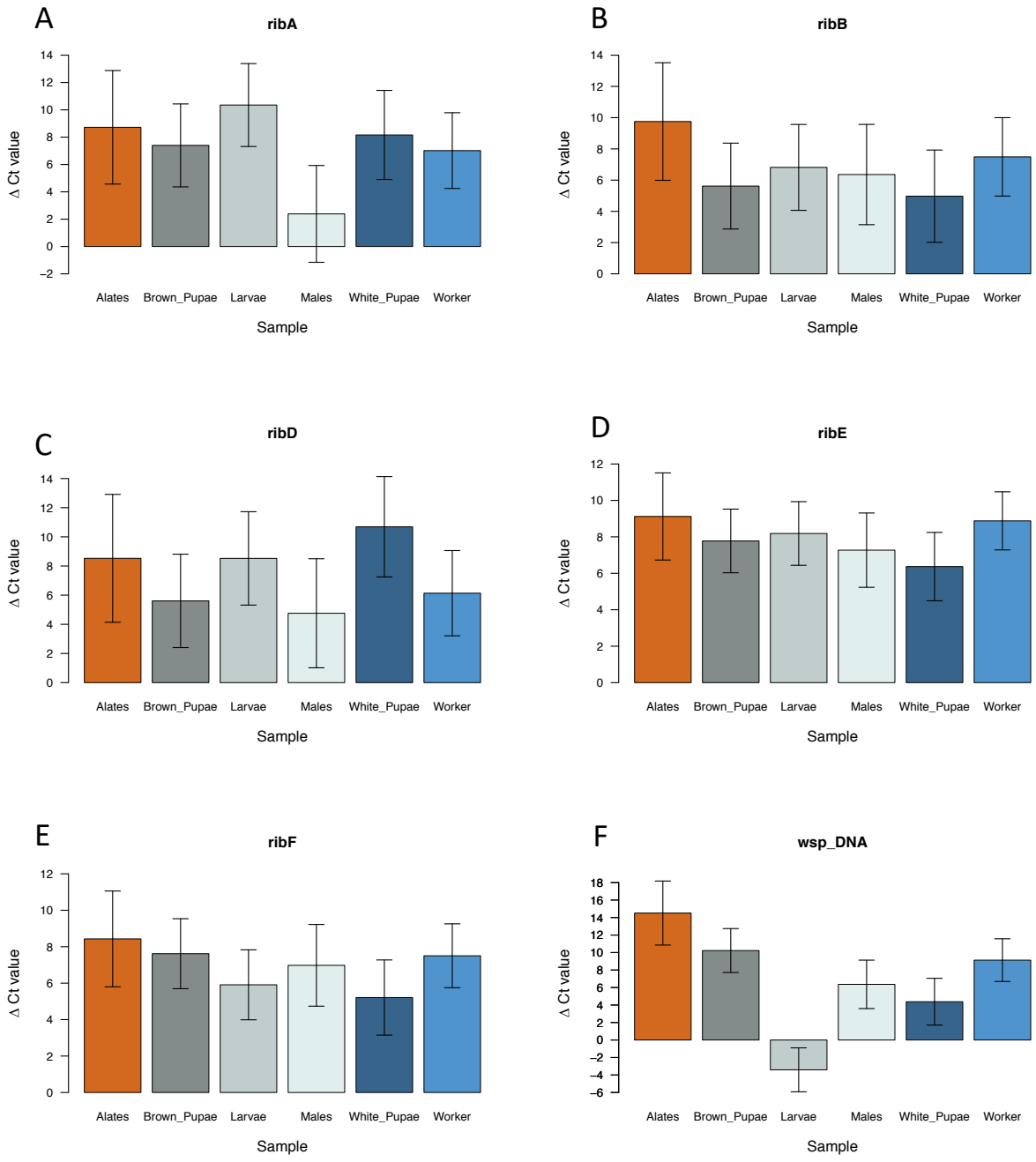


Figure S5: Targeted qPCR results for *Wolbachia* riboflavin synthesis genes in *A. echinatior* hosts. Histograms A-E show the average ΔCT (change in cycle threshold) of riboflavin genes for alates (reproductive females), brown pupae, larvae, males, white pupae, and workers with

standard error bars. Histogram F shows the expression values for the *Wolbachia* outer surface protein, *wsp*, to which all riboflavin genes were normalized.

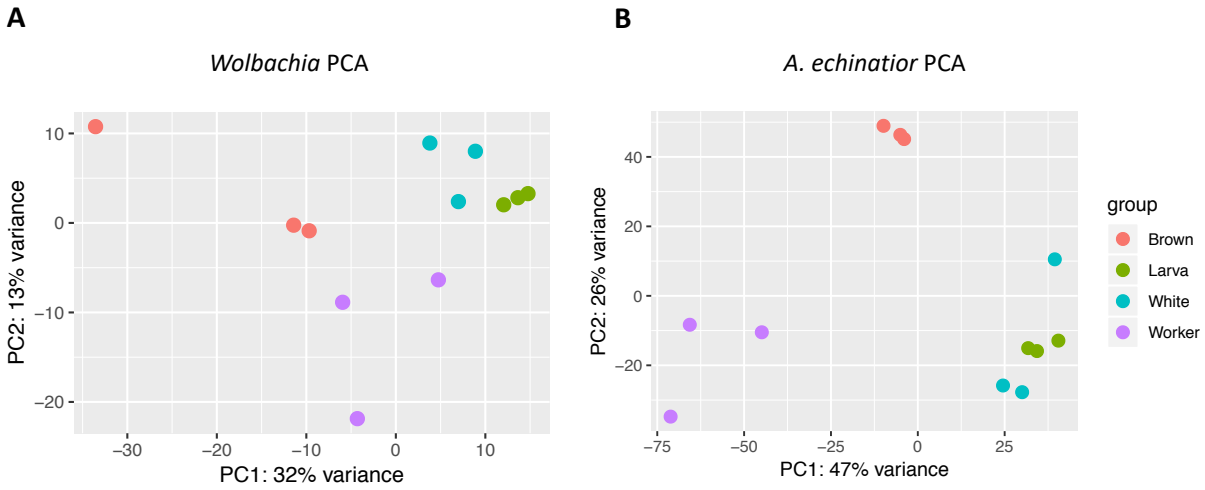


Figure S6: Principle component analyses based on the rlog normalized gene expression values for RNA-Seq samples mapped to (A) the *Wolbachia WolAcro1* genome and (B) the *Acromyrmex echinatior* host genome.

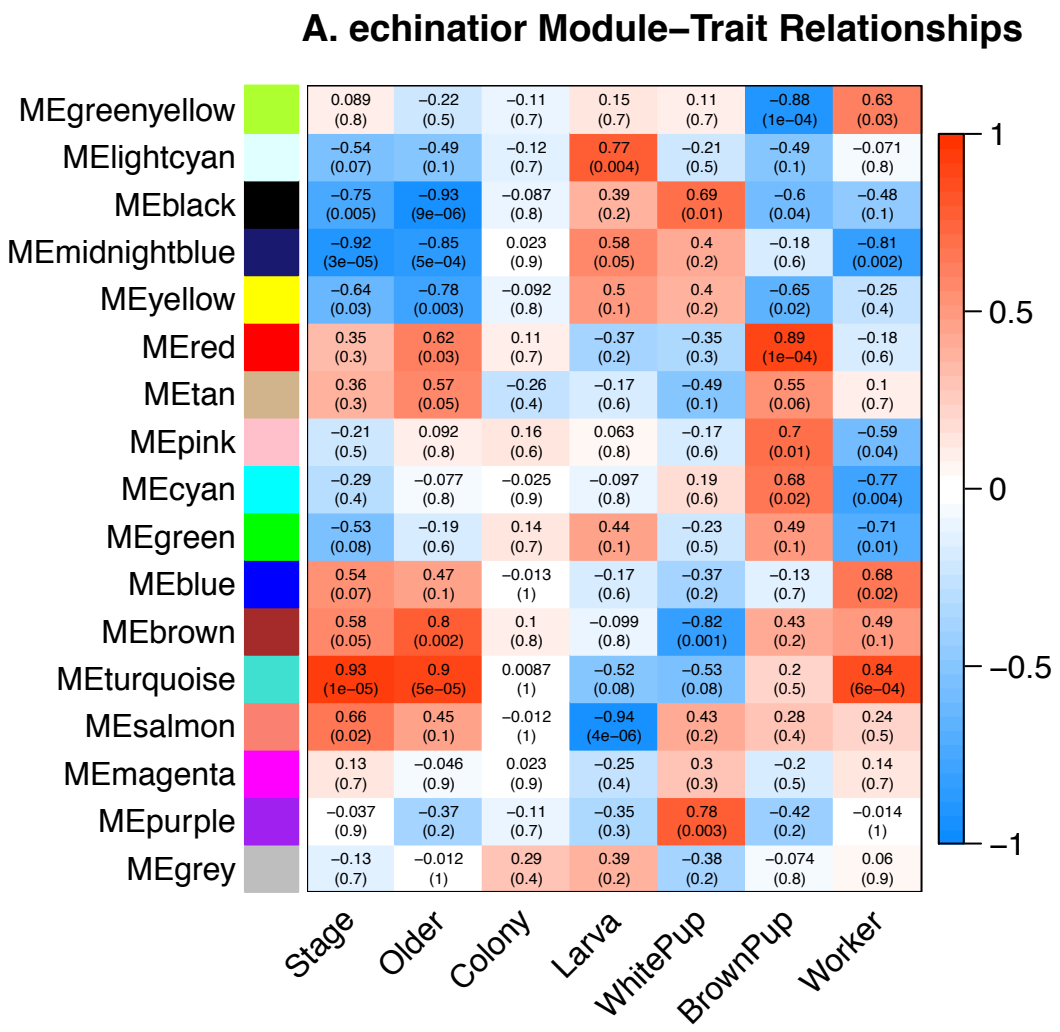


Figure S7: *A. echinatior* WGCNA results using the rlog normalized gene expression data from twelve RNA-seq samples (3 larvae, 3 white pupae, 3 brown pupae, and 3 worker samples from *A. echinatior* colonies) mapped to the *A. echinatior* genome. Module-trait association based on the summary profile of each module (eigengenes) to identify modules that are significantly associated with the sample traits. Each row corresponds to a module while each column

corresponds to a trait. Cells contain the test statistic value and its corresponding p-value. The table includes a heatmap representing module eigengene correlation according to the color legend on the right.

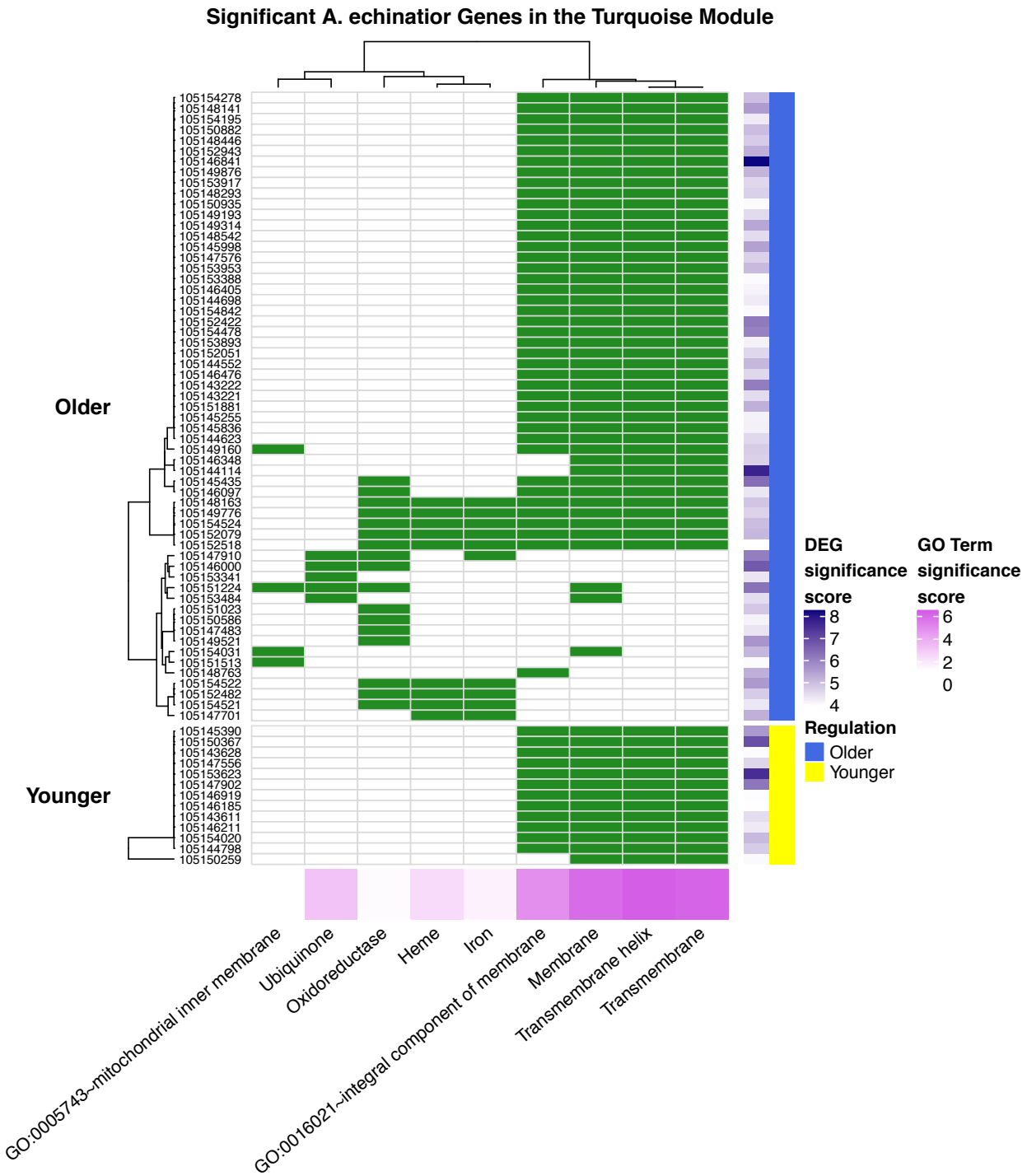

Significant A. echinatior Genes in the Turquoise Module
Older
Younger
105154278
105148141
105154195
105150882
105148446
105152943
105146841
105149876
105153917
105148293
105150935
105149193
105149314
105148542
105145998
105147576
105153953
105153388
105146405
105144698
105154842
105152422
105154478
105153893
105152051
105144552
105146476
105143222
105143221
105151881
105145255
105145836
105144623
105149160
105146348
105144114
105145435
105146097
105148163
105149776
105154524
105152079
105152518
105147910
105146000
105153341
105151224
105153484
105151023
105150586
105147483
105149521
105154031
105151513
105148763
105154522
105152482
105154521
105147701
105145390
105150367
105143628
105147556
105153623
105147902
105146919
105146185
105143611
105146211
105154020
105144798
105150259
DEG significance score
8
7
6
5
4
GO Term significance score
6
4
2
0
Regulation
Older
Younger
GO:0005743~mitochondrial inner membrane
Ubiquinone
Oxidoreductase
Heme
Iron
GO:0016021~integral component of membrane
Membrane
Transmembrane helix
Transmembrane

Figure S8: Clustering of DAVID gene ontology enrichments for statistically significant GO terms (adjusted p-value < 0.05) associated with *A. echinatior* hub genes in the turquoise WGCNA module. *A. echinatior* genes are listed on the right using Entrez Gene IDs. Each green square in the main plot indicates a gene ontology term associated with the gene in that respective row. To the right of the figure is a purple heatmap showing the significance score of differentially expressed genes and whether they are upregulated in the older (blue) or the younger (yellow) developmental stages. Below is a magenta heatmap showing the significance of each GO term, indicating the relative number of genes that identified with that GO term.

Name	Supergroup	Host	Accession	Reference
wUni	A	Muscidifurax uniraptor	PRJNA33275	Klasson et al. (2009)
wGlos	A	Glossina morsitans	PRJNA218057	Brelsfoard et al. (2014)
wMel	A	Drosophila melanogaster	PRJNA272	Wu et al. (2004)
wMelpop	A	Drosophila melanogaster	PRJNA196671	Woolfit et al. (2013)
wRi	A	Drosophila simulans	PRJNA33273	Klasson et al. (2009)
wHa	A	Drosophila simulans	PRJNA176303	Ellegaard et al. (2013)
wNferru	A	Nomada ferruginata	PRJNA322628	Gerth and Bleidorn (2016)
wNeluco	A	Nomada leucophthalm	PRJNA322628	Gerth and Bleidorn (2016)
wNpanzer	A	Nomada panzeri	PRJNA322628	Gerth and Bleidorn (2016)
wNflava	A	Nomada flava	PRJNA322628	Gerth and Bleidorn (2016)
wAlbB	B	Aedes albopictus	PRJEA76855	Mavingui et al. (2012)
wNo	B	Drosophila simulans	PRJNA176302	Ellegaard et al. (2013)
wTricho	B	Trichogramma pretiosum	PRJNA168121	Lindsey et al. (2016)
wVitB	B	Nasonia vitripennis	PRJNA61407	Kent et al. (2011a)
wPel	B	Culex quinquefasciatus	PRJNA30313	Klasson et al. (2008)
wJhb	B	Culex quinquefasciatus	PRJNA32209	Salzberg et al. (2009)
wBm	D	Brugia malayi	PRJNA58107	Foster et al. (2005)
wCimex	F	Cimex lectularius	PRJDB748	Nikoh et al. (2014)
wOncho	C	Onchocerca ochengi	PRJEA171829	Darby et al. (2012)

Table S1: *Wolbachia* strain name, supergroup affiliation, NCBI Bioproject accession number, and reference for all genomes analyzed in this study.

Gene	Primer ID	Sequence (5' - 3')
ribA	ribA_Forward	CCATACGCGTTAGTGGCTGA
	ribA_Reverse	TGAATGACGTGTCCAGTGCA
ribB	ribB_Forward	ATGGTGGGGTTTTAGCACGT
	ribB_Reverse	TCACACCCTACAGCTGCATG
ribD	ribD_Forward	GACGTCCGCATGCAGAGATA
	ribD_Reverse	TGATTTTCGCAGTGCAAGGC
ribE	ribE_Forward	TCGATGGCCACCTAGTTCAG
	ribE_Reverse	CCCATTTAGTGTAACGGAGCCT
ribF	ribF_Forward	AGAGAGTATGTACAAAGGGGCG
	ribF_Reverse	TCTACCTCTACACGCACCCT
wsp	wsp_Forward	GGTGCARCGTATATTAGCACTCC
	wsp_Reverse	GAACCGAAATAACGAGCTCCAG

Table S2: RT-PCR primers used for targeted riboflavin gene expression analyses.

Gene ID	Maker Annotation	GO Annotation	KEGG ID	GS.older	pGS.older
WolAcro100038	RP-S7; small subunit ribosomal protein S7		K02992	0.9394	5.81E-06
WolAcro100481	0	GO:0003677, GO:0004803, GO:0006313	0	0.93059	1.13E-05
WolAcro100157	0	GO:0003677, GO:0004803, GO:0006313	0	0.91721	2.66E-05
WolAcro100886	ftsI; cell division protein FtsI (penicillin-binding protein 3) [EC:3.4.16.4]	GO:0008658	K03587	-0.9138	3.24E-05
WolAcro100529	0		0	-0.9125	3.48E-05
WolAcro100410	purE; 5-(carboxyamino)imidazole ribonucleotide mutase [EC:5.4.99.18]	GO:0006189	K01588	-0.9125	3.48E-05
WolAcro100934	purL; phosphoribosylformylglycinamidine synthase [EC:6.3.5.3]		K01952	-0.9122	3.55E-05
WolAcro100458	trkH; trk system potassium uptake protein	GO:0006812, GO:0008324, GO:0055085	K03498	-0.9061	4.91E-05
WolAcro100933	pstA; phosphate transport system permease protein	GO:0006810, GO:0016020	K02038	-0.9025	5.89E-05
WolAcro100640	RP-L18; large subunit ribosomal protein L18	GO:0003735, GO:0005622, GO:0005840, GO:0006412	K02881	0.89839	7.18E-05
WolAcro100896	E3.4.17.19; carboxypeptidase Taq [EC:3.4.17.19]	GO:0004181, GO:0006508	K01299	-0.8783	0.00017
WolAcro100344	yajC; preprotein translocase subunit YajC		K03210	0.87224	0.00022
WolAcro100930	purM; phosphoribosylformylglycinamidine cyclo-ligase [EC:6.3.3.1]		K01933	0.87197	0.00022
WolAcro100125	TC.APA; basic amino acid/polyamine antiporter, APA family	GO:0003333, GO:0015171, GO:0016020	K03294	-0.8702	0.00023
WolAcro100448	radC; DNA repair protein RadC		K03630	0.86159	0.00032
WolAcro100600	nuoJ; NADH-quinone oxidoreductase subunit J [EC:1.6.5.3]	GO:0008137, GO:0055114	K00339	-0.8574	0.00036
WolAcro100708	dpo; DNA polymerase bacteriophage-type [EC:2.7.7.7]		K02334	0.84826	0.00049
WolAcro101127	tatA; sec-independent protein translocase protein TatA	GO:0008565, GO:0015031	K03116	0.84615	0.00052
WolAcro100874	E2.7.7.49; RNA-directed DNA polymerase [EC:2.7.7.49]		K00986	0.84566	0.00053
WolAcro101095	0	GO:0003677, GO:0004803, GO:0006313	0	0.8358	0.00071
WolAcro100711	proP; MFS transporter, MHS family, proline/betaine transporter	GO:0016021, GO:0022857, GO:0055085	K03762	0.82957	0.00084
WolAcro100714	RP-L25; large subunit ribosomal protein L25	GO:0003735, GO:0005840, GO:0006412, GO:0008097	K02897	0.82797	0.00088
WolAcro100412	rpoZ; DNA-directed RNA polymerase subunit omega [EC:2.7.7.6]	GO:0003677, GO:0003899, GO:0006351	K03060	0.82277	0.00101
WolAcro100042	gcpE; (E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase [EC:1.17.7.1 1.17.7.3]	GO:0016114, GO:0046429, GO:0055114	K03526	-0.8225	0.00102
WolAcro100992	0		0	-0.8189	0.00112
WolAcro100521	ctrA; two-component system, cell cycle response regulator CtrA	GO:0000160, GO:0003677, GO:0006355	K13584	0.81533	0.00123
WolAcro100036	tuf; elongation factor Tu	GO:0005525	K02358	-0.8094	0.00142
WolAcro101014	ANK; ankyrin	GO:0005515	K10380	0.80881	0.00145
WolAcro100850	SCO1_2; protein SCO1/2		K07152	0.80702	0.00151
WolAcro101140	0		0	0.7991	0.00182
WolAcro100378	iscU; nitrogen fixation protein NifU and related proteins	GO:0005506, GO:0016226, GO:0051536	K04488	0.78917	0.00227
WolAcro100794	RARS; arginyl-tRNA synthetase [EC:6.1.1.19]	GO:0000166, GO:0004814, GO:0005524, GO:0005737, GO:0006420	K01887	0.78909	0.00228
WolAcro100791	0		0	0.78437	0.00252
WolAcro100035	nusG; transcriptional antiterminator NusG	GO:0006355	K02601	-0.78	0.00277
WolAcro100641	RP-S5; small subunit ribosomal protein S5	GO:0003723, GO:0003735, GO:0005840, GO:0006412	K02988	0.77843	0.00286
WolAcro100249	0		0	0.77831	0.00287
WolAcro100155	0	GO:0009055, GO:0051536	0	0.77517	0.00306
WolAcro100110	0		0	0.77387	0.00314
WolAcro100150	ssb; single-strand DNA-binding protein	GO:0003697	K03111	0.77261	0.00322
WolAcro100299	0		0	0.7652	0.00373
WolAcro101087	lysK; lysyl-tRNA synthetase, class I [EC:6.1.1.6]		K04566	-0.7578	0.0043
WolAcro100921	glyA; glycine hydroxymethyltransferase [EC:2.1.2.1]	GO:0016740	K00600	0.7497	0.00499
WolAcro100309	secF; preprotein translocase subunit SecF		K03074	-0.7492	0.00504

WolAcro100547	plsY; glycerol-3-phosphate acyltransferase PlsY [EC:2.3.1.15]	GO:0005886, GO:0008654, GO:0043772	K08591	0.74454	0.00548
WolAcro100326	0	GO:0019028	0	-0.7415	0.00578
WolAcro100048	0		0	0.7285	0.00721
WolAcro100941	0		0	-0.7282	0.00724
WolAcro100636	RP-L5; large subunit ribosomal protein L5		K02931	0.72597	0.00751
WolAcro100022	0		0	0.72183	0.00804
WolAcro100053	ispDF; 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase / 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase [EC:2.7.7.60 4.6.1.12]	GO:0008299, GO:0008685, GO:0016114, GO:0050518	K12506	0.72015	0.00826
WolAcro100100	rnd; ribonuclease D [EC:3.1.13.5]	GO:0003676, GO:0006139, GO:0008408	K03684	0.72008	0.00827
WolAcro100224	IMPDH; IMP dehydrogenase [EC:1.1.1.205]	GO:0003824, GO:0055114	K00088	0.71809	0.00853
WolAcro100106	rimM; 16S rRNA processing protein RimM	GO:0006364	K02860	0.71725	0.00865
WolAcro100259	YARS; tyrosyl-tRNA synthetase [EC:6.1.1.1]	GO:0000166, GO:0003723, GO:0004812, GO:0005524, GO:0006418	K01866	0.70923	0.0098
WolAcro100390	pmbA; PmbA protein	GO:0006508, GO:0008237	K03592	-0.707	0.01013
WolAcro100845	putA; RHH-type transcriptional regulator, proline utilization regulon repressor / proline dehydrogenase / delta 1-pyrroline-5-carboxylate dehydrogenase [EC:1.5.5.2 1.2.1.88]	GO:0003842, GO:0008152, GO:0016491, GO:0055114	K13821	0.70237	0.01087
WolAcro100707	0		0	0.69608	0.01192
WolAcro100217	0		0	0.69454	0.01219
WolAcro100536	gyrA; DNA gyrase subunit A [EC:5.99.1.3]	GO:0003677, GO:0003916, GO:0003918, GO:0005524, GO:0005694, GO:0006265	K02469	0.69227	0.0126
WolAcro100114	bfr; bacterioferritin [EC:1.16.3.1]	GO:0006879, GO:0008199	K03594	0.69199	0.01265
WolAcro100635	RP-L14; large subunit ribosomal protein L14	GO:0003735, GO:0005840, GO:0006412	K02874	0.68638	0.0137
WolAcro100601	0		0	0.68406	0.01415

WolAcro100978	ndk; nucleoside-diphosphate kinase [EC:2.7.4.6]	GO:0004550, GO:0006165, GO:0006183, GO:0006228, GO:0006241	K00940	-0.6823	0.01451
WolAcro100630	RP-S19; small subunit ribosomal protein S19	GO:0003735, GO:0005840, GO:0006412	K02965	0.67962	0.01505
WolAcro100470	E2.2.1.1; transketolase [EC:2.2.1.1]		K00615	0.67749	0.01549
WolAcro100272	pleD; two-component system, cell cycle response regulator [EC:2.7.7.65]	GO:0000160	K02488	0.67712	0.01557
WolAcro100033	RP-L1; large subunit ribosomal protein L1		K02863	0.67705	0.01558
WolAcro100840	hemB; porphobilinogen synthase [EC:4.2.1.24]	GO:0004655, GO:0033014, GO:0046872	K01698	0.67265	0.01654
WolAcro100025	0	GO:0003677, GO:0004803, GO:0006313	0	0.66763	0.01767
WolAcro101162	0		0	0.66644	0.01795
WolAcro100431	0		0	0.66215	0.01898
WolAcro100673	PRPS; ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	GO:0000287, GO:0004749, GO:0009165	K00948	-0.6615	0.01914
WolAcro100310	ccmF; cytochrome c-type biogenesis protein CcmF	GO:0017004, GO:0020037	K02198	-0.6604	0.01942
WolAcro100915	0		0	-0.6597	0.01959
WolAcro100151	dapA; 4-hydroxy-tetrahydrodipicolinate synthase [EC:4.3.3.7]	GO:0008152, GO:0016829	K01714	0.65906	0.01975
WolAcro100594	0		0	0.65901	0.01976
WolAcro100795	0		0	0.65771	0.02009
WolAcro100165	hflK; membrane protease subunit HflK [EC:3.4.-.-]		K04088	0.65705	0.02026
WolAcro100099	trxA; thioredoxin 1	GO:0045454	K03671	0.65646	0.02041
WolAcro100724	RP-L13; large subunit ribosomal protein L13	GO:0003735, GO:0005840, GO:0006412	K02871	0.65506	0.02078
WolAcro101178	0		0	0.6539	0.02108
WolAcro100072	RP-L9; large subunit ribosomal protein L9		K02939	0.65382	0.0211
WolAcro100394	clpA; ATP-dependent Clp protease ATP-binding subunit ClpA	GO:0005524	K03694	0.65151	0.02172
WolAcro100876	0		0	0.65118	0.02181
WolAcro100499	0		0	0.6458	0.0233
WolAcro100006	recR; recombination protein RecR		K06187	-0.6324	0.02735
WolAcro100548	0		0	0.6309	0.02782
WolAcro100357	trxB; thioredoxin reductase (NADPH) [EC:1.8.1.9]	GO:0016491, GO:0055114	K00384	0.63082	0.02785

WolAcro100851	gyrB; DNA gyrase subunit B [EC:5.99.1.3]	GO:0003677, GO:0003918, GO:0005524, GO:0006265	K02470	0.63047	0.02796
WolAcro101120	E2.7.7.49; RNA-directed DNA polymerase [EC:2.7.7.49]		K00986	0.62788	0.02881
WolAcro100405	0		0	0.62255	0.03061
WolAcro100897	0	GO:0003677, GO:0004803, GO:0006313	0	0.61979	0.03158
WolAcro101083	MARS; methionyl-tRNA synthetase [EC:6.1.1.10]	GO:0000166, GO:0004812, GO:0005524, GO:0006418	K01874	0.61581	0.03302
WolAcro100789	ABCB-BAC; ATP-binding cassette, subfamily B, bacterial	GO:0005524, GO:0006810, GO:0016021, GO:0016887, GO:0042626, GO:0055085	K06147	0.61428	0.03358
WolAcro100584	purF; amidophosphoribosyltransferase [EC:2.4.2.14]	GO:0009116	K00764	0.60769	0.03608
WolAcro100269	E1.6.5.3; NADH dehydrogenase [EC:1.6.5.3]	GO:0003824, GO:0050662	K00329	0.60737	0.0362
WolAcro100883	0		0	0.6071	0.03631
WolAcro100180	bolA; BolA protein		K05527	-0.6066	0.03651
WolAcro100080	0		0	0.6055	0.03694
WolAcro100428	pyrG; CTP synthase [EC:6.3.4.2]	GO:0003883, GO:0006221	K01937	-0.604	0.03752
WolAcro100832	glmU; bifunctional UDP-N-acetylglucosamine pyrophosphorylase / Glucosamine-1-phosphate N-acetyltransferase [EC:2.7.7.23 2.3.1.157]		K04042	0.60334	0.0378
WolAcro100160	0		0	0.59867	0.03971
WolAcro100176	0		0	0.59223	0.04247
WolAcro100599	nuoL; NADH-quinone oxidoreductase subunit L [EC:1.6.5.3]	GO:0008137, GO:0042773, GO:0055114	K00341	-0.5842	0.04609
WolAcro100161	0		0	0.57876	0.04865
WolAcro100970	acpP; acyl carrier protein		K02078	0.57714	0.04943
WolAcro100401	ubiD; 4-hydroxy-3-polyprenylbenzoate decarboxylase [EC:4.1.1.98]	GO:0016831	K03182	0.57156	0.0522
WolAcro100112	0		0	0.55886	0.05891
WolAcro100131	nusA; N utilization substance protein A	GO:0003700, GO:0003723, GO:0031554	K02600	0.55741	0.05971

WolAcro100193	RP-L31; large subunit ribosomal protein L31	GO:0003735, GO:0005622, GO:0005840, GO:0006412	K02909	0.55198	0.06278
WolAcro100854	0		0	0.54867	0.06471
WolAcro100070	araM; glycerol-1-phosphate dehydrogenase [NAD(P)+] [EC:1.1.1.261]		K00096	-0.5476	0.06534
WolAcro100898	ppdK; pyruvate, orthophosphate dikinase [EC:2.7.9.1]	GO:0005524, GO:0016301, GO:0016310, GO:0016772	K01006	-0.5462	0.06615
WolAcro100118	tmk; dTMP kinase [EC:2.7.4.9]		K00943	-0.5462	0.06618
WolAcro100638	RP-S8; small subunit ribosomal protein S8	GO:0003735, GO:0005840, GO:0006412	K02994	0.54053	0.06961
WolAcro100179	dnaK; molecular chaperone DnaK		K04043	-0.5336	0.07398
WolAcro101130	0		0	0.53093	0.07571
WolAcro100332	0		0	-0.5226	0.08133
WolAcro100452	0		0	0.52223	0.08155
WolAcro100338	E2.7.7.3A; pantetheine-phosphate adenylyltransferase [EC:2.7.7.3]	GO:0003824, GO:0009058	K00954	0.51516	0.08652
WolAcro101079	vexE; membrane fusion protein, multidrug efflux system		K18990	0.51482	0.08676
WolAcro100393	0	GO:0003676, GO:0005524	0	0.51181	0.08895
WolAcro100890	0		0	-0.5104	0.09001
WolAcro100044	0		0	0.5081	0.09169
WolAcro100936	lysC; aspartate kinase [EC:2.7.2.4]		K00928	-0.5039	0.09484
WolAcro100343	RP-S2; small subunit ribosomal protein S2	GO:0003735, GO:0005622, GO:0005840, GO:0006412	K02967	0.49756	0.09977
WolAcro100026	purD; phosphoribosylamine--glycine ligase [EC:6.3.4.13]	GO:0004637, GO:0009113	K01945	0.4974	0.09989
WolAcro100029	dusB; tRNA-dihydrouridine synthase B [EC:1.-.-.-]	GO:0008033, GO:0017150, GO:0050660, GO:0055114	K05540	0.49627	0.10079
WolAcro100990	0		0	0.49395	0.10264
WolAcro100667	pcnB; poly(A) polymerase [EC:2.7.7.19]	GO:0003723, GO:0006396, GO:0016779	K00970	-0.4902	0.10565
WolAcro100589	virB9; type IV secretion system protein VirB9		K03204	0.48966	0.10613

WolAcro100540	SIG3.3.1; RNA polymerase sigma-32 factor	GO:0003677, GO:0003700, GO:0006352, GO:0006355, GO:0016987	K03089	0.48952	0.10624
WolAcro100825	coxC; cytochrome c oxidase subunit III [EC:1.9.3.1]	GO:0015002, GO:0016020	K02276	0.4845	0.11043
WolAcro100169	cutA; periplasmic divalent cation tolerance protein	GO:0010038	K03926	-0.4734	0.12005
WolAcro101084	0		0	0.47273	0.12066
WolAcro100443	0		0	0.46982	0.12328
WolAcro101214	0	GO:0003677, GO:0004803, GO:0006313	0	0.46773	0.12519
WolAcro100839	0		0	0.46528	0.12745
WolAcro101138	0	GO:0000150, GO:0003677, GO:0006310	0	-0.4631	0.12948
WolAcro100943	folD; methylenetetrahydrofolate dehydrogenase (NADP+) / methenyltetrahydrofolate cyclohydrolase [EC:1.5.1.5 3.5.4.9]	GO:0003824, GO:0004488, GO:0009396, GO:0055114	K01491	-0.4572	0.1351
WolAcro100677	mutS; DNA mismatch repair protein MutS	GO:0005524, GO:0006298, GO:0030983	K03555	-0.4543	0.13786
WolAcro100170	0		0	0.44793	0.14421
WolAcro100020	0		0	-0.4449	0.14726
WolAcro100678	0	GO:0005515	0	0.44325	0.14896
WolAcro100977	0		0	0.43486	0.15773
WolAcro100028	0		0	0.43316	0.15954
WolAcro100948	carB; carbamoyl-phosphate synthase large subunit [EC:6.3.5.5]	GO:0005524	K01955	-0.433	0.15969
WolAcro101099	0	GO:0003677, GO:0004803, GO:0006313	0	-0.4249	0.1686
WolAcro101064	0		0	-0.4232	0.1705
WolAcro101142	0		0	0.42209	0.17168
WolAcro100084	HARS; histidyl-tRNA synthetase [EC:6.1.1.21]		K01892	0.41919	0.17496
WolAcro100235	RP-L35; large subunit ribosomal protein L35	GO:0003735, GO:0005622, GO:0005840, GO:0006412	K02916	0.41259	0.18256
WolAcro100237	MTFMT; methionyl-tRNA formyltransferase [EC:2.1.2.9]	GO:0009058, GO:0016742	K00604	-0.397	0.20127

WolAcro100058	RP-S20; small subunit ribosomal protein S20	GO:0003723, GO:0003735, GO:0005622, GO:0005840, GO:0006412	K02968	0.39488	0.20395
WolAcro100297	0		0	-0.3883	0.21227
WolAcro100725	RP-S9; small subunit ribosomal protein S9	GO:0003735, GO:0005840, GO:0006412	K02996	0.38777	0.21295
WolAcro101154	0		0	0.38771	0.21302
WolAcro100355	recF; DNA replication and repair protein RecF		K03629	0.38699	0.21395
WolAcro100790	0		0	0.36864	0.23835
WolAcro100001	0		0	0.36007	0.25027
WolAcro100973	hslV; ATP-dependent HslUV protease, peptidase subunit HslV [EC:3.4.25.2]	GO:0004298, GO:0005839, GO:0051603	K01419	0.3551	0.25735
WolAcro100824	ruvC; crossover junction endodeoxyribonuclease RuvC [EC:3.1.22.4]	GO:0004520, GO:0006281, GO:0006310	K01159	-0.3526	0.26089
WolAcro100059	CS; citrate synthase [EC:2.3.3.1]	GO:0046912	K01647	0.35224	0.26147
WolAcro100305	dut; dUTP pyrophosphatase [EC:3.6.1.23]	GO:0016787, GO:0046080	K01520	-0.3485	0.26698
WolAcro100493	0		0	-0.3375	0.28336
WolAcro101161	0	GO:0043565	0	0.33513	0.28693
WolAcro101200	0		0	0.33375	0.28906
WolAcro100322	0		0	0.33017	0.29457
WolAcro100543	recA; recombination protein RecA	GO:0003697, GO:0005524, GO:0006281	K03553	0.32811	0.29778
WolAcro100027	SARS; seryl-tRNA synthetase [EC:6.1.1.11]	GO:0000166, GO:0004812, GO:0005524, GO:0006418	K01875	-0.3279	0.29806
WolAcro100686	0		0	-0.3237	0.30478
WolAcro100397	rnc; ribonuclease III [EC:3.1.26.3]	GO:0004525, GO:0006396	K03685	-0.3215	0.30812
WolAcro100674	0		0	-0.3207	0.3094
WolAcro101180	0	GO:0000150, GO:0003677, GO:0006310	0	-0.317	0.31544
WolAcro100146	0	GO:0005524, GO:0006810, GO:0016021, GO:0016887, GO:0042626, GO:0055085	0	-0.3071	0.33156
WolAcro100848	0		0	-0.3018	0.34039

WolAcro100463	0		0	0.29677	0.34891
WolAcro100637	RP-S14; small subunit ribosomal protein S14	GO:0003735, GO:0005622, GO:0005840, GO:0006412	K02954	-0.2938	0.35397
WolAcro100241	0		0	0.28928	0.36178
WolAcro100082	CYTB; ubiquinol-cytochrome c reductase cytochrome b subunit	GO:0009055, GO:0016020, GO:0016491	K00412	0.28219	0.37419
WolAcro100931	purC; phosphoribosylaminoimidazole-succinocarboxamide synthase [EC:6.3.2.6]		K01923	-0.2795	0.37902
WolAcro100284	0		0	0.27294	0.39071
WolAcro100813	0		0	0.2677	0.40022
WolAcro100588	mlaD; phospholipid/cholesterol/gamma-HCH transport system substrate-binding protein		K02067	-0.2646	0.40599
WolAcro101088	virB3; type IV secretion system protein VirB3		K03198	0.25512	0.42355
WolAcro100628	RP-L4; large subunit ribosomal protein L4	GO:0003735, GO:0005840, GO:0006412	K02926	0.25115	0.43105
WolAcro100484	0		0	-0.2433	0.44608
WolAcro100792	0		0	0.24063	0.45123
WolAcro100745	0	GO:0043565	0	-0.2299	0.47215
WolAcro100616	cybB; cytochrome b561	GO:0009055, GO:0016021	K12262	0.22791	0.4762
WolAcro100849	DPO3E; DNA polymerase III subunit epsilon [EC:2.7.7.7]		K02342	0.19598	0.54157
WolAcro100301	0		0	-0.191	0.55206
WolAcro101156	0	GO:0043565	0	0.19007	0.55406
WolAcro100869	0		0	0.18041	0.57474
WolAcro100242	PCCA; propionyl-CoA carboxylase alpha chain [EC:6.4.1.3]	GO:0005524	K01965	0.17143	0.59423
WolAcro100173	dksA; DnaK suppressor protein	GO:0008270	K06204	0.17027	0.59676
WolAcro100406	0		0	0.16503	0.60827
WolAcro100445	0		0	-0.1514	0.63849
WolAcro100913	0		0	-0.1275	0.69284
WolAcro100439	putative endonuclease		K07461	0.11197	0.72899
WolAcro100820	0	GO:0003677, GO:0004803, GO:0006313	0	0.08131	0.80165
WolAcro100879	0	GO:0003677, GO:0004803, GO:0006313	0	0.06434	0.84254
WolAcro100055	0		0	-0.0634	0.8448
WolAcro100549	czcD; cobalt-zinc-cadmium efflux system protein	GO:0006812, GO:0008324, GO:0016021, GO:0055085	K16264	-0.0247	0.9392
WolAcro100090	SAM50; outer membrane protein insertion porin family	GO:0019867	K07277	0.02056	0.94943
WolAcro100265	nuoA; NADH-quinone oxidoreductase subunit A [EC:1.6.5.3]	GO:0008137, GO:0055114	K00330	-0.0052	0.98727

Table S3: WGCNA analysis of *Wolbachia* genes in the brown module. Gene IDs are listed first followed by their MAKER annotation, gene ontology annotations, KEGG identifiers, and gene significance and p-value associated with the "older" sample trait (larva and white pupa compared to brown pupa and workers).

References

Abascal, F., Zardoya, R., Posada, D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21: 2104-5. doi: 10.1093/bioinformatics/bti263.

Al-Khodor, S., Price, C. T., Kalia, A. Abu Kwaik, Y. 2010. Functional diversity of ankyrin repeats in microbial proteins. Trends Microbiol. 18: 132-139. doi: 10.1016/j.tim.2009.11.004.

Alexa, A., Rahnenfuhrer, J. (2018). topGO: Enrichment analysis for gene ontology. R package version 2.34.0.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Anders, S., Huber, W. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11: R106. doi: 10.1186/gb-2010-11-10-r106.

Anders, S., Ply, P. T., Huber, W. 2015. HT-Seq—a python framework to work with high-throughput sequencing data. *Bioinformatics*. 31(2):166-9. doi: 10.1093/bioinformatics/btu638.

Andersen, S. B., Boye, M., Nash, D. R., and Boomsma, J. J. 2012. Dynamic *Wolbachia* prevalence in Acromyrmex leaf-cutting ants: potential for a nutritional symbiosis. J. Evol. Biol. 25: 1340–1350. doi:10.1111/J.1420-9101.2012.02521.X.

Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Pontén, T., Alsmark, U. C., Podowski, R. M., Näslund, A. K., Eriksson, A. S., Winkler, H. H., Kurland, C. G. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* 396(6707): 133.40.

Beckmann, J. F., Ronau, J. A., Hochstrasser, M. 2017. A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. Nat. Microbiol. 1: 17007. doi: 10.1038/nmicrobiol.2017.7.

Bordenstein, S. R., Paraskevopoulos, C., Dunning Hotopp, J. C., Sapountzis, P., Lo, N., Bandi, C., Tettelin, H., Werren, J. H., Boourtzis, K. 2009. Parasitism and mutualism in *Wolbachia*: what the phylogenomic trees can and cannot say. *Mol. Biol. Evol.* 26(1): 231-41. doi: 10.1093/molbev/msn243.

Borowiec, M. 2015. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4:e1660. doi:10.7717/peerj.1660.

Bourtzis, K. 2008. *Wolbachia*-based technologies for insect pest population control. *Adv. Exp. Med. Biol.* 627: 104-13. doi: 10.1007/978-0-387-78225-6_9.

Bouwma, A. M., Shoemaker, D. 2011. *Wolbachia* wSinvictaA infections in natural populations of the fire ant Solenopsis invicta: testing for phenotypic effects. J. Insect Sci. 11: 11. doi: 10.1673/031.011.0111.

Brownlie, J. C., Adamski, M., Slatko, B., McGraw, E. A. 2007. Diversifying selection and host adaptation in two endosymbiont genomes. BMC Evol. Biol. 7:68. doi: 10.1186/1471-2148-7-68.

Brownlie, J. C., Cass, B. N., Riegler, M., Witsenburg, J. J., Iturbe-Ormaetxe, I., McGraw, E. A., O'Neill, S. L. 2009. Evidence for metabolic provisioning by a common invertebrate endosymbiont, *Wolbachia pipientis*, during periods of nutritional stress. *PLoS Pathog.* 5(4):e1000368. doi: 10.1371/journal.ppat.1000368.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berk-Lyons, D., Lozupone, C. A., Fierer, N., Knight, R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences upper sample. Proc. Natl. Acad. Sci. USA. 1: 4516-22. doi: 10.1073/pnas.1000080107.

Carver, T., Böhme, U., Otto, T. D., Parkhill, J., Berriman, M. 2010. BamView: viewing mapped read alignment data in the context of the reference sequence. Bioinformatics 26: 676-677. doi:10.1093/bioinformatics/btq010

Carver, T., Harris, S. R., Berriman, M., Parkhill, J., McQuillan, J. A. 2012. Atemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28: 464-469. doi:10.1093/bioinformatics/btr703.

Comandatore, F., Cordaux, R., Bandi, C., Blaxter, M., Darby, A., Makepeace, B. L., Montagna, M., Sassera, D. 2015. Supergroup C *Wolbachia*, mutualist symbionts of filarial nematodes have a distinct genome structure. *Open Biol.* 5(12): 150099. doi: 10.1098/rsob.150099.

Comandatore, F., Sassera, D., Montagna, M., Kumar, S., Koutsovoulos, G., Thomas, G., Repton, C., Babayan, S. A., Gray, N., Cordaux, R., Darby, A., Makepeace, B., Blaxter, M. 2013. Phylogenomics and analysis of shared genes suggest a single transition to mutualism in *Wolbachia* of nematodes. Genome Biol. Evol. 5: 1668-74. doi: 10.1093/gbe/evt125.

Djaman, O., Outten, F. W., Imlay, J. A. 2004. Repair of oxidized iron-sulfur clusters in Escherichia coli. J. Biol. Chem. 279: 44590-44599. doi: 10.1074/jbc.M406487200.

Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O., Arnold, F. H. 2005. Why highly expressed proteins evolve slowly. *PNAS* 102(40): 14338-14343. doi: 10.1073/pnas.0504070102.

Dunning Hotopp, J. C., Clark, M. E., Oliveira, D. C., Foster, J. M. Fischer, P. Muñoz Torres, M. C., Giebel, J. D., Kumar, N., Ishmael, N., Wang, S. et al. 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317(5845): 1753-6. doi: 10.1126/science.1142490.

Emms, D. M., Kelly, S. 2015. Orthofinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16: 157. doi: 10.1186/s13059-015-0721-2.

Engelstädter, J., Jurst, G. D. D. 2009. The ecology and evolution of microbes that manipulate host reproduction. *Annu. Rev. Ecol. Evol. Syst.* 40: 127-149. doi: 10.1146/annurev.ecolsys.110308.120206.

Fallon, A. M., Baldridge, G. D., Carroll, E. M., Kurtz, C. M. 2014. Depletion of host cell riboflavin reduces *Wolbachia* levels in cultured mosquito cells. *In Vitro Cell Dev. Anim.* 50(8): 707-713. doi: 10.1007/s11626-014-9758-x.

Fares, M. A., Barrio, E., Sabater-Muñoz, B., Moya, A. 2002. The evolution of the heat-shock protein GroEL form Buchnera, the primary endosymbiont of aphids, is governed by positive selection. Mol. Biol. Evol. 19(7): 1162-1170. doi: 10.1093/oxfordjournals.molbev.a004174.

Foster, J., Ganatra, M., Kamal, I., Ware, J., Makarova, K., Ivanova, N., Bhattacharyya, A., Kapatral, V., Kumar, S., Posfai, J. et al. 2005. The *Wolbachia* genome of *Brugia malayi*:

endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* 3(4): e121. doi: 10.1371/journal.pbio.0030121.

Frost, C. L., Fernández-Marín, H., Smith, J. E., Hughes, W. O. 2010. Multiple gains and losses of *Wolbachia* symbionts across a tribe of fungus-growing ants. *Mol. Ecol.* 19(18):4077-85. doi: 10.1111/j.1365-294X.2010.04764.x

Frost, C. L., Pollock, S. W., Smith, J. E., and Hughes, W. O. H. 2014. *Wolbachia* in the flesh: symbiont intensities in germ-line and somatic tissues challenge the conventional view of *Wolbachia* transmission routes. PloS One 9,:e95122. doi:10.1371/journal.pone.0095122.

Gerth, M., Bleidorn, C. 2016. Comparative genomics provides a timeframe for *Wolbachia* evolution and exposes a recent biotin synthesis operon transfer. Nat. Microbiol. 2, 16241. doi: 10.1038/nmicrobiol.2016.241.

Gerth, M., Gansauge, M., Weigert, A., Bleidorn, C. 2014. Phylogenomic analyses uncover origin and spread of the *Wolbachia* pandemic. Nat. Com. 5: 5117. doi:10.1038/ncomms6117.

Gill, A. C., Darby, A. C., Makepeace, B. L. 2014. Iron necessity: the secret of *Wolbachia*'s success? PLoS Negl. Trop. Dis. 8(10): e3224. doi: 10.1371/journal.pntd.0003224.

Grote, A., Voronin, D., Ding, T., Twaddle, A., Unnasch, T. R., Lustigman, S., Ghedin, E. 2017. Defining Brugia malayi and *Wolbachia* symbiosis by stage-specific dual RNA-seq. PLoS Negl. Trop. Dis. 11(3): e0005357. doi: 10.1371/journal.pntd.0005357.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59: 2422-24. doi: 10.1093/sysbio/syq010.

Gutzwiller, F., Carmo, C. R., Miller, D. E., Rice, D. W., Newton, I. L. G., Hawley, R. S., Teixeria, L., Bergman, C. M. 2015. Dynamics of *Wolbachia* pipientis gene expression across the Drosophila melanogaster life cycle. G3 (Bathesda) 5(12): 2843-2856. doi: 10.1534/g3.115.021931.

Huang, D. W., Sherman, B. T., Lempicki, R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4(1): 44-57. doi: 10.1038/nprot.2008.211.

Huerta-Cepas, J., Serra, F., Bork, P. 2016. ETE 3: reconstruction, analysis and visualization of phylogenomic data. Mol. Biol. Evol. 6: 1635-1638. doi: 10.1093/molbev/msw046.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., Rattei, T., Mende, D. R., Sunagawa, S., Kuhn, M., Jensen, L. J., Mering, C., Bork, P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44: D286-D293. doi: 10.1093/nar/gkv1248.

Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M., Otto, T. D. 2013. REAPR: a universal tool for genome assembly evaluation. Genome Biol. 14: R47. doi:10.1186/gb-2013-14-5-r47.

Iturbe-Ormaetxe, I., Burke, G. R., Riegler, M., O'Neill, S. L. 2005. Distribution, expression, and motif variability of ankyrin domain genes in *Wolbachia* pipientis. Gen. Mol. Biol. 187(15): 5136-5145. doi:10.1128/JB.187.15.5136–5145.2005.

Iturbe-Ormaetxe, I., Woolfit, M., Rancès, E., Duplouy, A., O'Neill, S. L. 2011. A simple protocol to obtain highly pure *Wolbachia* endosymbiotic DNA for genome sequencing. J. Microbiol. Methods 84: 134-136. doi:10.1016/j.mimet.2010.10.019.

Jones, D. T., Taylor, W. R., Thornton, J. M. 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8: 275-82.

Kanehisa, M., Goto, S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28: 27-30.

Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., Hirakawa, M. 2010. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res. 38: D355-60. doi: 10.1093/nar/gkp896.

Keller, L., C. Liautard, M. Reuter, W. D. Brown, L. Sundström, and M. Chapuisat. 2001. Sex ratio and *Wolbachia* infection in the ant Formica exsecta. Heredity 87:227-233.

Klasson, L., Walker, T., Sebaihia, M., Sanders, M. J., Quail, M. A., Lord, A., Sanders, S., Earl, J., O'Neill, S. L., Thomson, N., Sinkins, S. P., Parkhill, J. 2008. Genome evolution of *Wolbachia* strain wPip from the Culex pipiens group. Mol. Biol. Evol. 25(9): 1877-1887. doi: 10.1093/molbev/msn133.

Klasson, L., Westberg, J., Sapountzis, P., Näslund, K., Lutnaes, Y., Darby, A. C., et al. 2009. The mosaic genome structure of the *Wolbachia* wRi strain infecting Drosophila simulans. Proc. Natl. Acad. Sci. U. S. A. 106: 5725–5730. doi:10.1073/pnas.0810753106.

Kosiol, C., Vinař, T., da Fonseca, R. R., Hubisz, M. J., Bustamante, C. D., Nielsen, R., Siepel, A. 2008. Patterns of positive selection in six mammalian genomes. *PLoS Genet.* 4(8): e1000144. doi: 10.1371/journal/pgen.1000144.

Kremer, N., Voronin, D., Charif, D., Mavingui, P., Mollereau, B., Vavre, F. 2009. *Wolbachia* interferes with ferritin expression and iron metabolism in insects. *PLoS Pathog.* 5(10): e1000630. doi: 10.1371/journal.ppat.1000630.

Krueger, F. 2015. Trim galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

Langfelder, P., Horvath, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559.

Larracuente, A. M., Sackton, T. B., Greenberg, A. J., Wong, A., Singh, N. D., Sturgill, D., Zhang, Y., Oliver, B., Clark, A. G. 2008. Evolution of protein-coding genes in *Drosophila*. *Trends Genet*. 24(3): 11-1234. doi: 10.1016/j.tig.2007.12.001.

Le, S. Q., Gascuel, O. 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25: 1307-20. doi: 10.1093/molbev/msn067.

LePage, D. P., Metcalf, J. A., Bordenstein, S. R., On, J., Permutter, J. I., Shropshire, D., Layton, E. M., Funkhouser-Jones, L. J., Beckmann, J. F., Bordenstein, S. R. 2017. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. Nat. 543: 243-247. doi: 10.1038/nature21391.

Lindsey, A. R., Rice, D. W., Bordenstein, S. R., Brooks, A. W., Bordenstein, S. R., Newton, I. L. G. 2018. Evolutionary genetics of cytoplasmic incompatibility genes *cifA* and *cifB* in prophage WO of *Wolbachia*. Genome Biol. Evol. 10: 434-451. doi:10.1093/gbe/evy012.

Lindsey, A. R. I., Werren, J. H., Richards, S., Stouthamer, R. 2016. Comparative genomics of a parthenogenesis-inducing *Wolbachia* symbiont. *G3* 6(7): 2113-2123. doi: 10.1534/g3.166.028449.

Liu, H., Bao, W., Lin, M., Rikihisa, Y. 2012. *Erlichia* type IV secretion effector ECH0825 is translocated to mitochondria and curbs ROS and apoptosis by upregulating host MnSOD. *Cell Microbiol*. 14(7):1037-1050. doi: 10.1111/j.1462-5822.2012.01775.x.

Melkani, G. C., Zardeneta, G., Mendoza, J. A. 2004. Oxidized GroEL can function as a chaperonin. Front. Biosci. 9: 724-731.

Moran, N. A. 1996. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. Proc. Natl. Acad. Sci. USA 93(7): 2873-2878.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., Kanehisa, M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35: W182-W185. doi: 10.1093/nar/gkm321.

Moriyama, M., Nikoh, N., Hosokawa, T., Fukatsu, T. 2015. Riboflavin provisioning underlies *Wolbachia*'s fitness contribution to its insect host. *MBio.* 6(6): e01732-15. doi: 10.1128/mBio.01732-15.

Newton, I. L. G., Clark, M. E., Kent, B. N., Bordenstein, S. R., Qu, J., Richards, S., Kelkar, Y. D., Werren, J. H. 2016. Comparative genomics of two closely related *Wolbachia* with different reproductive effects on hosts. *Genome Biol. Evol.* 8(5): 1526-1542. doi: 10.1093/gbe/evw096.

Pan, X., Lührmann, A., Satoh, A., Laskowski-Arce, M. A., Roy, C. R. 2008. Ankyrin repeat proteins comprise a diverse family of bacterial Type IV effectors. *Science* 320(5883): 1651-1654. doi: 10.1126/science.1158160.

Pontieri, L., Schmidt, A. M., Singh, R., Pedersen, J. S., Linksvayer, T. A. 2017. Artificial selection on ant female caste ratio uncovers a link between female-biased sex ratios and infection by *Wolbachia* endosymbionts. *J. Evol. Biol.* 30(2): 225-234. doi: 10.1111/jeb.13012.

Rambaut, A. (2016). FigTree version 1.4.3 [computer program] http://tree.bio.ed.ac.uk/software/figtree/.

Rikihisa, Y., Lin, M. 2010. *Anaplasma phagocytophilum* and *Erlichia chaffeensis* Type IV secretion and Ank proteins. *Curr. Opin. Microbiol.* 13(1): 59-66. doi: 10.1016/j.mib.2009.12.008.

Russell, J. A. 2012. The ants (Hymenotera: Formicidae) are unique and enigmatic hosts of prevalent *Wolbachia* (Alphaproteobacteria) symbionts. *Myrmecological News* 16: 7-23.

Sapountzis, P., Zhukova, M., Hansen, L. H., Sørensen, S. J., Schiøtt, M., and Boomsma, J. J. (2015). Acromyrmex leaf-cutting ants have simple gut microbiota with nitrogen-fixing potential. Appl. Environ. Microbiol. 81: 5527–5537. doi:10.1128/AEM.00961-15.

Schaible, U. E., Kaufmann, S. H. 2004. Iron and microbial infection. *Nat. Rev. Microbiol.* 2(12): 946-53. doi: 10.1038/nrmicro1046.

Sedgwick, S. G., Smerdon, S. J. 1999. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci.* 24(8): 311-6.

Sinha, A., Li, A., Sun, L., Carlow, C. K. S. 2019. Complete genome sequence of the *Wolbachia* wAlbB endosymbiont of Aedes albopictus. Genome Biol. Evol. 11(3): 706-720. doi: 10.1093/gbe/evz025.

Siozios, S., Ioannidis, P., Klasson, L., Andersson, S. G. E., Braig, H. R., Bourtzis, K. 2013. The diversity and evolution of *Wolbachia* ankyrin repeat domain genes. PLoS One 8(2): e55390. doi: 10.1371/journal.pone.0055390.

Supek, F., Bošnjak, M., Škunca, N., Šmuc, T. 2011. REVIGO summarizes and visualizes long lists of Gene Ontology terms. PLoS ONE 6(7): e21800. doi:10.1371/journal.pone.0021800.

Susin, M. F., Baldini, R. L., Gueiros-Filho, F., Gomes, S. L. 2006. GroES/GroEL and DnaK/DnaJ have distinct roles in stress responses and during cell cycle progression in Caulobacter crescentus. J. Bacteriol. 188(23): 8044-8053. doi: 10.1128/JB.00824-06.

Suyama, M., Torrents, D., Bork, P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34: W609-W612. doi: 10.1093/nar/gkl315.

Trapnell, C., Hendrickson, D., Sauvageau, M., Goff, L., Rinn, J. L., Pachter, L. 2012. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31: 46-53. doi: 10.1038/nbt.2450.

Walker, T., Johnson, P. H., Moreira, L. A., Iturbe-Ormaetxe, I., Frentiu, F. D., McMeniman, C. J., Leong, Y. S., Dong, Y., Axford, J., Kriesner, P., Lloyd, A. L., Ritchie, S. A., O'Neill, S. L. Hoffmann, A. A. 2011. The *wMel Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature* 246(7361): 450-3. doi: 10.1038/nature10355.

Wenseleers, T., Sundström, L., Billen, J. 2002. Deleterious *Wolbachia* in the ant *Formica truncorum*. *Proc. Biol. Sci.* 269(1491): 623-9. doi: 10.1098/rspb.2001.1927.

Werren, J. H., Baldo, L., Clark, M. E. 2008. *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* 6(10): 741-51. doi: 10.1038/nrmicro1969.

Werren, J. H., Windsor, D. M. 2000. *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? *Proc. R. Soc. Lond. Ser. B. Biol. Sci.* 267: 1277-85. doi: 10.1098/rspb.2000.1139.

Wu, M., Sun, L. V. Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J. C., McGraw, E. A., Martin, W., Esser, C., Ahmadinejad, N., et al. 2004. Phylogenomics of the reproductive parasite *Wolbachia* pipientis wMel: a streamlined genome overrun by mobile genetic elements. PLoS Biol. 2(3): E69. doi: 10.1371/journal.pbio.0020069.

Yang, Z., Neilsen, R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol. Biol. Evol. 19: 908-17. doi:10.1093/oxfordjournals.molbev.a004148.

Yen, J. H., Barr, R. 1971. New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens* L. *Nature* 232: 657-658.

Zhang, B., Horvath, S. 2005. A general framework for weighted gene co-expression network analysis. Stat. Appl. Genet. Mol. 4: 1544-6115. doi:10.2202/1544-6115.1128.

Zhang, J., Neilsen, R., Yang, Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. 22: 2472-9. doi: 10.1093/molbev/msi237.